

DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF BUSINESS ADMINISTRATION
BUSINESS ADMINISTRATION PROGRAM
DOCTORAL THESIS
Doctor of Philosophy (Ph.D.)

**EXCELLING CUSTOMER COMPLAINT MANAGEMENT THROUGH
TEXT ANALYTICS: COMPLAINT CLASSIFICATION IN SOCIAL
MEDIA AND COMPLAINT TYPE PREDICTION**

Birce DOBRUCALI

Supervisor
Prof. Dr. Burcu İLTER

İZMİR - 2021

DOCTORAL THESIS
APPROVAL PAGE

University : Dokuz Eylul University
Graduate School : Graduate School of Social Sciences
Name and Surname : Birce Dobrucalı
Title of the Thesis : Excelling Customer Complaint Management Through Text Analytics:
Complaint Classification in Social Media and Complaint Type Prediction
Defence Date : 12.03.2021
Supervisor : Professor Doc.Burcu İLTER

EXAMINING COMMITTEE MEMBERS

<u>Title,Name and Surname</u>	<u>University</u>	<u>Signature</u>
Professor Doc.Burcu İLTER	- Dokuz Eylül Üniversitesi	
Associate Prof.Güzin ÖZDAĞOĞLU	- Dokuz Eylül Üniversitesi	
Professor Doc.Burcu Selin YILMAZ	- Dokuz Eylül Üniversitesi	
Associate Prof.Aykan CANDEMİR	- Ege Üniversitesi	
Professor Doc.Coşkun KÜÇÜKÖZMEN	- İzmir Ekonomi Üniversitesi	

Unanimity ()

Majority of votes ()

The thesis titled as "**Excelling Customer Complaint Management Through Text Analytics: Complaint Classification in Social Media and Complaint Type Prediction**" prepared and presented by Birce Dobrucalı is accepted and approved.

Prof. Dr. Asuman ALTAY
Müdür

DECLARATION

I hereby declare that this doctoral thesis titled as "Excelling Customer Complaint Management Through Text Analytics: Complaint Classification in Social Media and Complaint Type Prediction" has been written by myself in accordance with the academic rules and ethical conduct. I also declare that all materials benefited in this thesis consist of the mentioned resources in the reference list. I verify all these with my honour.

Date 12/03/2021

Birce DOBRUCALI

Signature

ABSTRACT

Doctoral Thesis

Doctor of Philosophy (Ph.D.)

**Excelling Customer Complaint Management Through Text Analytics: Complaint
Classification in Social Media and Complaint Type Prediction**

Birce DOBRUCALI

Dokuz Eylül University

Graduate School of Social Sciences

Department of Business Administration

Business Administration Program

The ubiquity, ease of use, and speed of social media gave a tremendous rise to sharing consumption experiences online. The upsurge in electronic word-of-mouth created both an advantageous and dangerous field for the companies. Whereas an effectively handled online complaint may result in winning the complainant back, a poorly handled complaint may dangerously backfire on the company. Thus, excelling in online complaint management became a priority for companies. Therefore, understanding online complaints, and identifying typologies of online complaints and complainants are vital for surviving in the digital landscape. Initiated from this fact, this thesis aims to both identify content-based and interaction-based online consumer complaint types and predict complaint types according to the complaint magnitude rooted in complainants' personality traits, emotion, Twitter usage activity, as well as complaint's sentiment polarity, and feedback frequency. Accordingly, 297 thousand complaint tweets, features of over 220 thousand consumer profiles and more than 24 million user tweets were gathered from Twitter. The obtained data was

analyzed via two-step machine learning approach, where the first step is the determination of the content features through text analytics. In the second step, the relationships between the clusters and the independent variables were revealed through machine learning. The validity of the revealed relationships was proven by the prediction power of the model, where the model's overall accuracy rate reaches 61.7% for the content-based clusters and 70.5% for the interaction-based clusters. This thesis contributes to both literature and practice by identifying types of online complaints that reveal diverse customer complaint behavior on Twitter, proposing a set of content and profile features that can be utilized for predicting complaint type, revealing the relationship between content features, profile features, and online complaint type, and proposing a novel and flexible approach for excelling online complaint through prioritization.

Keywords: Online Complaint, Complaint Handling, Social CRM, Social Media Analytics, Machine Learning, Text Analytics, Twitter

ÖZET

Doktora Tezi

**Müşteri Şikayet Yönetiminin Metin Analitiği Yöntemiyle Mükemmelleştirilmesi:
Sosyal Medyadaki Şikayetlerin Sınıflandırılması ve Şikayet Tipi Tahminlemesi**

Birce DOBRUCALI

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

İngilizce İşletme Anabilim Dalı

İngilizce İşletme Yönetimi Programı

Sosyal medyanın her yerden erişilebilir, hızlı ve kolay kullanılır olması; tüketici deneyimlerinin çevrimiçi olarak paylaşılmasında büyük bir artışa sebep olmuştur. Elektronik ağızdan ağıza iletişimde görülen bu artış, firmalar için hem avantajlı hem de tehlikeli bir alan yaratmaktadır. Etkili yönetilen bir çevrimiçi şikayet, şikayet sahibinin geri kazanılmasına olanak tanırken, kötü yönetilen bir şikayet ise şirketi zarara uğratabilmektedir. Bu nedenle, çevrimiçi şikayet yönetiminde üstünlük sağlanması, şirketler için bir öncelik haline gelmiştir. Bu doğrultuda, çevrimiçi şikayetlerin anlaşılması, şikayetlerin ve şikayetçilerin tipolojilerinin belirlenmesi, dijital ortamda hayatta kalmak için büyük önem taşır. Buradan yola çıkarak, bu tez hem içerik hem de etkileşim bazında çevrimiçi tüketici şikayet türlerini belirlemeyi amaçlamaktadır. Bu amaçla, şikayetçilerin kişilik özellikleri, şikayetlerine yansıttıkları hisleri, Twitter kullanım aktiviteleri ile söz konusu şikayetlerin duygu kutupluluğu ve etkileşim hacmi baz alınarak, önemlerine göre şikayet türlerinin tahminlenmesi amaçlanmaktadır. Bu bağlamda, Twitter'dan 297 bin şikayet tweet'i, 220 binin üzerinde tüketici

profilinin özellikleri ve bu profil sahipleri tarafından paylaşılmış 24 milyonu aşkın tweet toplanmış ve iki aşamalı makine öğrenimi yaklaşımı ile analiz edilmiştir. İlk adımda içerik özellikleri metin analitiği yoluyla belirlenmiş, ikinci adımda ise kümeler ve bağımsız değişkenler arasındaki ilişkiler makine öğrenimi yoluyla ortaya çıkarılmıştır. Ortaya çıkan ilişkilerin geçerliliği, modelin genel doğruluk oranının içeriğe dayalı kümeler için %61.7'e ve etkileşime dayalı kümeler için %70.5'e ulaşması ile de desteklenmiştir. Bu tez, Twitter'da farklı müşteri şikayeti davranışlarını ortaya çıkaran çevrimiçi şikayet türlerini belirleyerek, şikayet türünü tahmin etmek için kullanılabilecek bir dizi içerik ve profil özelliği önererek, ve aynı zamanda içerik özellikleri ile profil özellikleri arasındaki ilişkiyi ortaya çıkararak hem literatüre hem de uygulamaya katkıda bulunmakta ve çevrimiçi şikayet türü ve önceliklendirme yoluyla çevrimiçi şikayet yönetiminin mükemmelleştirilmesi için yeni ve esnek bir yaklaşım önermektedir.

Anahtar Kelimeler: Çevrimiçi Şikayet, Şikayet Yönetimi, Sosyal MİY, Sosyal Medya Analizi, Makine Öğrenimi, Metin Analizi, Twitter

EXCELLING CUSTOMER COMPLAINT MANAGEMENT THROUGH TEXT ANALYTICS: COMPLAINT CLASSIFICATION IN SOCIAL MEDIA AND COMPLAINT TYPE PREDICTION

TABLE OF CONTENTS

THESIS APPROVAL PAGE	II
DECLARATION	III
ABSTRACT	IV
ÖZET	VI
TABLE OF CONTENTS	VIII
ABBREVIATIONS	XIV
LIST OF TABLES	XVI
LIST OF FIGURES	XVIII
INTRODUCTION	1

CHAPTER ONE

COMPLAINING BEHAVIOR AND COMPLAINT HANDLING

1.1.	THE CONCEPT OF CUSTOMER COMPLAINT	7
1.2.	MODELS ON CUSTOMER COMPLAINT BEHAVIOR	9
1.3.	FACTORS THAT INFLUENCE COMPLAINING BEHAVIOR	13
1.3.1.	Personal Factors	14

1.3.1.1.	Personality Traits	17
1.3.1.2.	Emotions	21
1.3.1.3.	Sentiment Polarity	23
1.3.2.	Non-personal Factors	24
1.4.	ONLINE COMPLAINING BEHAVIOR	26
1.4.1.	The Importance of Online Complaints	27
1.4.2.	Online Complaint Channels	29
1.4.2.1.	Online Complaint Sites	30
1.4.2.2.	Social Media Platforms	31
1.5.	ORGANIZATIONAL COMPLAINT MANAGEMENT	33
1.5.1.	The Importance of Effective Organizational Complaint Management	36

CHAPTER TWO

SOCIAL MEDIA AND SOCIAL CUSTOMER RELATIONSHIP MANAGEMENT (CRM)

2.1.	DEFINITION OF SOCIAL MEDIA	39
2.2.	HISTORICAL DEVELOPMENT OF SOCIAL MEDIA	41
2.3.	DIFFERENCES BETWEEN SOCIAL MEDIA AND TRADITIONAL MEDIA	43
2.4.	SOCIAL MEDIA PLATFORMS	44
2.4.1.	Blogs	45
2.4.2.	Microblogs	46
2.4.3.	Social Networking Sites (SNS)	47
2.4.4.	Content Sharing Platforms	48
2.4.5.	Social News and Bookmarking	49

2.4.6.	Rating and Review Sites	50
2.4.7.	Forums	50
2.4.8.	Virtual Worlds	51
2.5.	SOCIAL CRM CONCEPT	52
2.6.	DIFFERENCES BETWEEN TRADITIONAL CRM AND SOCIAL CRM	56
2.7.	THE IMPORTANCE OF SOCIAL CRM FOR COMPANIES	59

CHAPTER THREE

MACHINE LEARNING CONCEPT AND APPLICATIONS IN MARKETING

3.1.	DEFINITION OF MACHINE LEARNING	62
3.2.	DEVELOPMENT OF MACHINE LEARNING TECHNIQUES	63
3.3.	TYPES OF MACHINE LEARNING	65
3.3.1.	Supervised Learning	65
3.3.2.	Unsupervised Learning	66
3.3.3.	Semi-Supervised Learning	66
3.3.4.	Reinforcement Learning	67
3.4.	MACHINE LEARNING AND TEXT MINING	67
3.4.1.	Information Extraction	69
3.4.2.	Information Retrieval	70
3.4.3.	Clustering	70
3.4.4.	Summarization	70
3.5.	MACHINE LEARNING AND TEXT MINING ALGORITHMS	71
3.5.1.	Support Vector Machine (SVM)	72

3.5.2.	Linear Discriminant Analysis	73
3.5.3.	Naïve Bayes	73
3.5.4.	Nearest Neighbor	74
3.5.5.	Linear Regression	74
3.5.6.	Ensemble Models	74
3.5.7.	Decision Trees	75
3.5.8.	Neural Networks	75
3.5.9.	K-Means Clustering	76
3.5.10.	Fuzzy C-Means	76
3.5.11.	Hidden Markov Models	77
3.6.	LITERATURE REVIEW ON MACHINE LEARNING AND TEXT MINING APPLICATIONS IN SOCIAL MEDIA ANALYTICS	77
3.6.1.	Review Sites	78
3.6.2.	Microblogs	81
3.6.3.	Social Networking Sites	82
3.6.4.	Visual Content Communities	83
3.6.5.	Cross-Media Studies	84

CHAPTER FOUR

RESEARCH FRAMEWORK, METHODOLOGY, AND FINDINGS

4.1.	RESEARCH FRAMEWORK	86
4.1.1.	Significance and Aim of the Study	86
4.1.2.	Research Model	88
4.2.	RESEARCH METHOD	91

4.2.1.	Research Design	91
4.2.1.1.	Independent and Dependent Variables	91
4.2.1.2.	Selection of the Research Platform	92
4.2.1.3.	Sampling Frame	93
4.2.1.4.	Data Collection	94
4.3.	DATA ANALYSIS	98
4.3.1.	Descriptive Statistics	98
4.3.2.	Classification and Clustering	99
4.3.2.1.	Content-based Classification	99
4.3.2.1.1.	<i>N</i> -Grams	100
4.3.2.1.2.	Dictionary Establishment	101
4.3.2.1.3.	Results of Content-Based Classification	103
4.3.2.1.3.1.	Legal Threat	103
4.3.2.1.3.2.	Call for Boycott	105
4.3.2.1.3.3.	Brand Switch Warning	106
4.3.2.2.	Interaction-based Clustering	107
4.3.2.2.1.	Two-Step Cluster Analysis	108
4.3.2.2.2.	Results of Interaction-based Clustering	110
4.3.2.2.2.1.	Spiteful Complaints	113
4.3.2.2.2.2.	Pervasive Complaints	114
4.3.2.2.2.3.	Mediocre Complaints	115
4.3.3.	Two-Step Machine Learning Approach	116

4.3.3.1.	The First Leg of Two-Step Machine Learning Approach	117
4.3.3.1.1.	Text Analytics	117
4.3.3.1.1.1.	Personality Traits	118
4.3.3.1.1.2.	Emotion	121
4.3.3.1.1.3.	Sentiment Polarity	123
4.3.3.2.	The Second Leg of Two-Step Machine Learning Approach	125
4.3.3.2.1.	Multinomial Logistic Regression	125
4.3.3.2.1.1.	Results for Content-Based Categories	131
4.3.3.2.1.1.1.	Full Sample Results for Content-Based Categories	131
4.3.3.2.1.1.2.	Holdout Sample Results for Content-Based Categories	138
4.3.3.2.1.2.	Results for Interaction-Based Clusters	143
4.3.3.2.1.2.1.	Full Sample Results for Interaction-Based Clusters	143
4.3.3.2.1.2.2.	Holdout Sample Results for Interaction-Based Clusters	148
4.3.4.	Model Validation and Testing	152
4.4.	DISCUSSION OF THE RESULTS AND MANAGERIAL IMPLICATIONS	155
4.5.	LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH	161
REFERENCES		163

ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
B2C	Business to Consumer
BBS	Bulletin Board System
CCB	Consumer Complaint Behavior
CRM	Customer Relationship
eWOM	Electronic Word of Mouth
FCM	Fuzzy C-Means
HMMs	Hidden Markov Models
IE	Information Extraction
IM	Impression Management
IR	Information Retrieval
IRC	Internet Relay Chat
LDA	Linear Discriminant Analysis
ML	Machine Learning
MUD	Multi-User Domain
NER	Named Entity Recognition
NGO	Non-governmental Organization
ROI	Return on Investments
S&P	Standard & Poor's
SNS	Social Networking Site

SVM Support Vector Machine

UGC User-Generated Content

WOM Word of Mouth



LIST OF TABLES

Table 1: Changes in the Customer Complaining Behavior Understanding	9
Table 2: Personal Factors Influencing Complaining Behavior	16
Table 3: Non-Personal Factors Influencing Complaining Behavior	25
Table 4: Differences between Traditional CRM and Social CRM	58
Table 5: Common Machine Learning and Text Mining Algorithms	71
Table 6: Market Cap Ranking of the Selected B2C Service Providers	94
Table 7: Companies' Customer Care Accounts, Related Hashtags and Keywords	96
Table 8: Descriptive Statistics	98
Table 9: Examples For n-grams	102
Table 10: Key Phrases for Content-Based Classification	103
Table 11: Examples of Complaint Tweets that Contain Legal Threat	104
Table 12: Examples for Complaint Tweets that Contain Call for Boycott	106
Table 13: Examples for Complaint Tweets that Contain Brand Switch Warning	107
Table 14: Cluster Distribution	110
Table 15: Centroids	113
Table 16: A Sample of Results from IBM Watson's Personality Insights	120
Table 17: A Sample of Results from ParallelDots Emotion Analysis AI API	122
Table 18: A Sample of Results from ParallelDots Sentiment API	124
Table 19: Model Fitting Information for Full Sample Results of Content-Based Categories	132
Table 20: Full Observation Parameter Estimates for Content-Based Categories	137
Table 21: Model Fitting Information for Holdout Sample Results of Content-Based Categories	138

Table 22: Holdout Sample Parameter Estimates for Content-Based Categories	142
Table 23: Model Fitting Information for Full Sample Results of Interaction-Based Clusters	143
Table 24: Full Observation Parameter Estimates for Content-Based Categories	147
Table 25: Model Fitting Information for Holdout Sample Results of Interaction-Based Clusters	148
Table 26: Parameter Estimates for Interaction-Based Clusters	151
Table 27: Precision, Recall and Accuracy Results for The Validation Set	153
Table 28: Precision, Recall and Accuracy Results for The Test Set	154

LIST OF FIGURES

Figure 1: Day and Landon's Taxonomy of Customer Complaint Behavior	11
Figure 2: Singh's taxonomy of customer complaint behavior	12
Figure 3: Tronvoll's model for dynamic complaining behavior	13
Figure 4: The Chronological History of Social Media	43
Figure 5: Social CRM Process	54
Figure 6: The Social CRM House	55
Figure 7: Relationships Among Text Mining Techniques	69
Figure 8: Research Model	90
Figure 9: Cluster Quality	110
Figure 10: Clusters	112
Figure 11: Comparison of Spiteful Complaints Cluster with Other Clusters	114
Figure 12: Comparison of Pervasive Complaints Cluster with Other Clusters	115
Figure 13: Comparison of Mediocre Complaints Cluster with Other Clusters	116

INTRODUCTION

As internet-based applications have become an essential part of our daily lives, by 2019, the number of individuals who use social media actively reached 3.484 billion with an average of 8.9 accounts per person and an average daily use of 2 hours 16 minutes (Global Digital Report 2019, 2019). Internet-based applications and social media platforms are mainly preferred for disseminating information and online communication, and hence they are standing at the heart of the information flow and social interaction (Raacke and Bonds-Raacke, 2008). They re-shape the way of information acquisition, opinion sharing, and communication with the ubiquity they provide (Kapoor et al., 2018). Due to the ease of use, speed, and scope it provides, social media can alter society's public discourse and set trends in a wide range of topics varying from politics to entertainment (Asur and Huberman, 2010). From the marketing perspective, these platforms offer fruitful, organic and free consumer data varying from consumption experience and word of mouth (WOM) to complaints.

In addition to social media's ease of use and speed, consumers' inclination to engage with a third-party platform than communicating directly with the company (Vermeulen & Seegers, 2009) gave rise to sharing consumption experiences online. Consumers are now sharing their opinions regarding the consumption experience with brands, complaining, and recommending their favorite brands more than ever. Hence, social media holds the ability to enhance product acceptability and creating group opinions (Shareef et al., 2018). The upsurge in electronic WOM is both a blessing and a curse for companies. An effectively handled online complaint may result in winning the complainant back and attract new customers through shaping reputation (Einwiller and Steilen, 2015). On the other hand, even though web care is a promising tool for brand communication, it can dangerously backfire on the company when not handled effectively (Lee and Song, 2010). Mainly, complaint handling and responsiveness to complaints constitute a crucial part of the web care services. 83% of the consumers that post complaints on social media platforms declared that they liked hearing back from the company (Martiz Research, 2011). However, despite the proven impact and criticality of the online complaints on customer satisfaction, the extant literature points out that

companies are moderately responsive to online complaints, and often attempt to redirect the complaint to a private channel that cannot be tracked by other Internet users (Einwiller and Steilen, 2015). These facts bring out the first gap that constitutes the basis for this thesis.

Gap 1: Even though online complaints are incredibly significant due to their magnitude; companies are moderately responsive to online complaints and need an effective way of managing them for ensuring both customer satisfaction and brand reputation.

Following the importance of responsiveness to online complaints, the second vital aspect of online complaint handling is timeliness in responding. Einwiller and Steilen (2015) put forth that most companies submit their responses within 9 hours and 26 minutes. When the online world's pace is considered, 9 hours can be a long time for responding to customers, especially when they are waiting for an explanation regarding an unfavorable experience. A legendary case highlighting the prominence of response time is Dave Carroll's "United Breaks Guitar" music video in which Carroll complains about his experience with United Airlines. The airlines tried to reach Carroll in less than 24 hours after streamed the music video on YouTube, but the video has already reached 20,000 views at that time and hit 5 million views within a month (Deighton and Kornfeld, 2010). Therefore, replying the complaint in the shortest possible time is vital to prevent the complaint from reaching a wider audience and initiating an outrage (Pfeffer et al., 2014). However, when the volume of online complaints and companies' budget constraints are considered, it may not always be easy to reply every single complaint in a short time. Besides, the criticality of the complaint or complainant's expectation regarding the response time may vary. The finest solution, in this regard, is the prioritization of the complaints. In this way, companies can balance customer satisfaction with response time and the required budget for the complaint handling process. In addition, prioritization of complaints requires a flexible framework that can be customized by companies with different motivations. For the companies that are keen on brand image, complaints with increasing popularity (in terms of the number of interactions) may be considered as the most harmful ones. Also, for companies struggling with financial problems, the complaints that include "legal threat" emphasis may hold more importance and have priority. Hence, priorities may vary from company to company and from time to

time, resulting in the need for a flexible prioritization tool for ensuring the validity of prioritization. However, to the best of our knowledge, neither the prioritization of online complaints is investigated, nor a flexible framework is provided. Accordingly, the second gap that constitutes the basis for this thesis is presented below.

Gap 2: For balancing complainants' satisfaction with the response time, and the budget required for managing online complaints, it is crucial to prioritize online complaints.

For building a sufficient complaint handling procedure to manage online complaints, it is vital to understand customers and related outcomes that can impact company reputation and profitability. Concordantly, for managing complaints effectively and preventing any possible outrage or an online firestorm, robust interpretation of two factors plays a central role: (1) customer's motivation for complaining; (2) complaint's magnitude. Since the extant literature points out numerous personal factors that impact complaining intention, motivation for complaining may be different for each complainant. Demographic, psychological, emotional, cultural, social, and experiential factors can be listed as the primary personal factors proven to impact complaining intention (Tronvoll, 2007b). Extant literature points out personality traits and emotions as the most dominant personal variables among the listed personal factors. A vast number of studies validated the relationship between personality traits and complaining intention (e.g., Bolfing, 1989; Fornell and Westbrook, 1979; Malafi et al., 1993; Slama and Celuch, 1994; Liu et al., 1997; Markus and Kitayama, 1990), and the relationship between emotions and complaining intention (e.g., Bolfing, 1989; Folkes et al., 1987; Tronvoll, 2008). Even though these studies validate the relationship between personality traits, emotion, and complaining, there are two aspects to be considered. First, the vast majority of these studies employ surveying method, including close-ended questions that generate structured numerical data. Even though surveys can provide a wealth of information, they hold some drawbacks. The honesty of subjects who answer the questionnaire is a doubt for researchers, mainly for sensitive questions (Fricker and Schonlau, 2002). Since both personality and emotions are sensitive in nature, and honesty is a doubt that cannot be averted by the researchers, the surveying method can be considered to be more prone to bias when compared with the text analytics method. Analyzing autogenous customer data with text analytics may increase the strength of the relationship between these

constructs and enhance the relationship's predictive and explanatory power. Second, there is a need for online complaint typology to predict the magnitude of customer complaints. Due to the differences between offline and online environments, existing complaint typologies cannot fully explain online complaints' nature. However, for estimating the magnitude of an online complaint, it is vital to analyze the complaint by considering the complainant's characteristics. For instance, the complainant may have a tendency to provoke other people continuously, and such situations may result in a widespread call for a boycott or an online firestorm. In such cases, more intense and effective action is required to be taken in the shortest time possible. Therefore, accurate estimation and a more in-depth understanding may be obtained by analyzing unstructured, autogenous, textual data via text analytics and investigating these factors' concurrent impact on the complaining intention. Accordingly, the third and fourth gaps that constitute the basis for this thesis are presented below.

Gap 3: Although the extant literature validates the relationship between personality traits, emotions, and complaining intention through the surveying method, the method is prone to respondents' risk of honesty.

Gap 4: Due to the differences between online and offline environments, traditional complaint typologies cannot be fully integrated into online complaints. Hence, for obtaining a more accurate estimation of online complaint's magnitude, it is vital to identify online complaint types.

In this context, this thesis aims to both identify content-based and interaction-based online consumer complaint types and predict complaint types according to the complaint magnitude rooted in complainants' personality traits, emotion, Twitter usage activity, as well as complaint's sentiment polarity, and feedback frequency. Following the aim of the thesis, consumer complaints about B2C service providing companies listed in the S&P's 500 Index are obtained from Twitter for the time period between April 2017 and December 2018, referring to 240 thousand complaint tweets, around 220 thousand consumer profiles and around 24 million user tweets to be analyzed. Diverse text analytics software tools are employed to obtain complaint tweets' sentiment, consumers' personality traits, and emotions. Besides, the features regarding user profiles (the number of followers, the number of followings, etc.) and the features regarding complaint

tweets (the length of the conversation, the number of likes, the number of retweets, the number of replies from other users, etc.) are collected. This information, together with complaint sentiments, personality traits, and emotions of consumers, constructs independent variables of the thesis and used for predicting complaint type.

Initiated by the objectives of this thesis and the gaps in the literature, the following research questions are addressed in this dissertation. It should be noted that the stated questions are going to be revisited in the following chapters in detail.

- What are the characteristics of the content-based complaint categories (legal threat, call for boycott, switching intention)?
- What is the interaction-based online complaint types regarding complainants' personality traits, emotions, and Twitter usage activity, as well as the complaint's sentiment and feedback frequency?
- What are the characteristics of the interaction-based complaint clusters?
- How can the complaint type be predicted via machine learning and text analytics?
- How can online complaint typology be used for prioritizing customer complaints?

The introduction section provides readers an insight into the context of the thesis. The problem definition is presented along with the aim of the study and the expected contributions, and research questions are addressed.

Chapter 1 is devoted to complaining behavior and complaint handling. The chapter starts with defining the customer complaint concept, presents a detailed review of extant models on customer complaint behavior (CCB), and explains personal and non-personal factors that impact complaining behavior. The second half of the chapter focuses on online complaining behavior, addresses the importance and types of online complaints, and provides information about online complaint channels. Also, the organizational complaint management process and the importance of effective complaint management are discussed.

Chapter 2 is organized around social media and social customer relationship management (CRM) concepts. The chapter starts with the definition of social media, and

continues with the historical development of social media, indicates its differences from the traditional media, and explains the social media tools including blogs, microblogs, networking sites, content sharing platforms, social news, and bookmarking sites, rating and review sites, forums, and virtual worlds in detail. Then, the social CRM concept is defined, its differences with traditional CRM is explained, and its importance for companies is highlighted.

Chapter 3 is devoted to the machine learning and its application in the marketing domain with an emphasis on social media analytics. The development of machine learning techniques is explained, and types of machine learning techniques are described. Usage of machine learning in text mining is described, and related algorithms are defined. Finally, a detailed review of studies that employed machine learning and text mining in social media analytics is provided.

The last chapter, *Chapter 4*, is mainly devoted to the methodology and findings of this dissertation. Accordingly, the aim of this thesis, research model and design, data collection method are discussed. Following these, procedures undertaken for data analysis are explained, and findings are presented. The contribution and limitations of the thesis, along with the recommendations for future research, are provided.

CHAPTER ONE

COMPLAINING BEHAVIOR AND COMPLAINT HANDLING

1.1. THE CONCEPT OF CUSTOMER COMPLAINT

Despite all organizational precautions, companies may face problems in their relationship with customers (Homburg and Fürst, 2005). Such problems or dissatisfactions with the purchased products or services grounds consumer complaining behavior, which requires company attention due to the risk of becoming detrimental (Istanbulluoglu, 2017). Consequently, organizational post-complaint behavior and complaint handling process play a crucial role for firms. Whereas a fair recovery for undesired incidents is capable of escalating customer satisfaction and loyalty (Smith and Bolton, 1998), a poor recovery may intensify negative evaluation of customers (Bitner et al., 1990). Since eliminating undesired incidents is impossible, companies need to understand customer complaint behavior to establish an efficient complaint management process that ensures customer satisfaction.

One of the earliest definitions of customer complaint is put forth by Landon (1980), and describes the concept as "...an expression of dissatisfaction by individual consumers (or on a consumer's behalf) to a responsible party in either the distribution channel or a complaint handling agency." Oliver (1980) highlighted that complaints reflect dissatisfaction caused by negative disconfirmation of pre-purchase expectations. Jacoby and Jaccard (1981) define complaining behavior as "...an action taken by an individual which involves communicating something negative regarding a product or service to the firm which is manufacturing or marketing that product or service, or to some third-party organization's entity." Singh and Howell (1985) expand the definition by emphasizing non-behavioral responses and describes customer complaint as "...a set of all behavioral and non-behavioral responses which involve communicating something negative regarding a purchase episode and is triggered by perceived dissatisfaction with that episode."

As can be understood from the provided definitions, the traditional marketing approach views complaining as a post-purchase behavior triggered by customers' disappointment (Oliver, 1997; Singh, 1990). Nevertheless, complaints are not always triggered by dissatisfaction. Hence, dissatisfaction is a necessary but not sufficient reason for complaining (Day, 1984; Singh and Pandya, 1991). Since complaining is a more complex behavior than a reaction to post-purchase disappointment, Tronvoll (2007b) suggests that the following aspects should be considered while defining complaint behavior: (1) Since service is co-created through interactions between employees and customers, negative incidents and consequential complaints should be considered as a part of the overall service-interaction process; (2) The value-in-use evaluation may trigger complaints during or after the consumption; (3) By many research on the subject (e.g., Bearden and Teel, 1983; Blodgett and Granbois, 1992), dissatisfaction is proven to be a minor cause of complaining behavior. Accordingly, Tronvoll (2007b) defines complaining as a process initiated by a service experience that lies outside of the customer's acceptance zone during the service interaction or the evaluation of value-in-use of the consumed service. One of the most contemporary definitions is made by Stephens (2000), and following the previous one, the definition states that "...complaining is a post-purchase process that may or may not occur when customers are disappointed."

Tronvoll (2008) argues that the change in epistemological perspective in complaint research requires a conversion in the lexicon of complaint behavior and puts forth the comparison between approaches indicated in Table 1. The transition from goods-dominant logic to service-dominant logic led the static relationship between customer and provider to evolve into a dynamic one, which changed traditional researchers' focus from the post-purchase and outcome-related behavior to intra- and post-relational behavior. Consequently, the basis of complaints changed from product failure to failure or missing competence from the service provider, which, in turn, changed understanding on motivation for complaining from customer dissatisfaction to negative impression towards the provider. As an outcome of such change, passively listening to customers gave its place to establishing a dialogue with customers and seeking feedback to make adjustments during service co-creation.

Table 1: Changes in the Customer Complaining Behavior Understanding

	Previous Understanding	New Understanding
Perspective	Goods-dominant logic	Service-dominant logic
Customer - provider interaction	Static/transactional	Dynamic/relationship
Complaint phase	Post-purchase	Intra and post-interaction
Basis of complaint	Product failure (based on the expected attributes)	Failure or missing competence (based on the perceived value proposition)
Result of failure	Dissatisfaction	Negative impression
Nature of communication	Listening to the customer, Noting specific complaint	Dialogue with the customer, Seeking feedback
Aims of service recovery	Correcting specific product failure, Obtaining a satisfied customer	Learning and adjustment during the co-creation process, Strengthening the relationship

Source: Tronvoll (2008)

1.2. MODELS ON CUSTOMER COMPLAINT BEHAVIOR

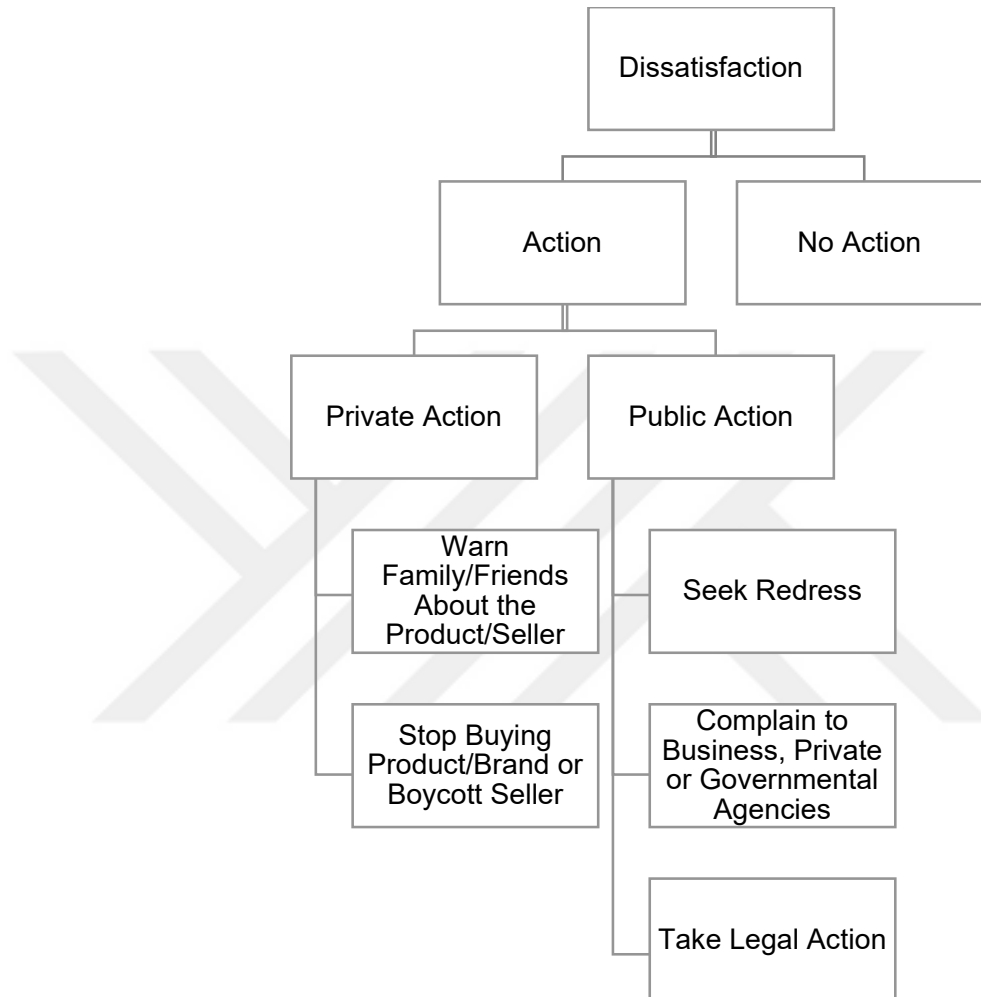
Several behavioral models have been proposed for outlining customers' responses to a negative service experience. Among them, the most referred to and extensively accepted complaint behavior models belong to Hirschman (1970), Day and Landon (1977), Singh (1988), and Tronvoll (2012).

In his book, *"Exit, Voice, and Loyalty: Response to Decline in Firms, Organizations, and States,"* Hirschman (1970) investigates how customer loyalty impacts exit and voice behavior. According to Hirschman, customers have two forms of active response options when they face failures; the first option is economic (exit), and

the second one is political (voice). Within this context, exit behavior denotes escaping from a deplorable condition, whereas voice behavior involves the intention to change the condition (Graham and Keeley, 1992). Hirschman advocates that as the degree of loyalty increases, the likelihood of voice upsurges accordingly. On the other hand, loyalty prevents deterioration by limiting customers' tendency to be the first to exit. However, it is indicated that although loyalty delays exit behavior, even the most loyal customers can exit depending upon the bargaining power vis-a-vis the organization. To sum up his views on complaining behavior, Hirschman (1970) advocates that loyalty postpones exit behavior and activates voice.

Another taxonomy of customer complaint behavior is put forth by Day and Landon (1977) and has achieved wide acceptance in the literature. As indicated in Figure 1, according to Day and Landon's taxonomy, there exist three options available for dissatisfied customers: taking no action, taking private action, or taking public action. The primary decision for customers is whether to take any action or not. At this first level, customers may prefer to refrain from taking action and forget about the incident or decide to take public or private action. Private actions include word-of-mouth activities such as warning family and friends about the product/retailer, stop purchasing product/brand, and boycotting the seller. Public actions include seeking redress directly from the seller, complaining to seller/manufacturer, complaining to media or voluntary organizations, and taking legal action to obtain redress. Day and Landon (1977) ground public/private action dichotomy on the nature and importance of the purchased product and the effort required and perceived outcome of the taken action. According to this view, expensive and complex products encourage more action to be taken publicly.

Figure 1: Day and Landon's Taxonomy of Customer Complaint Behavior

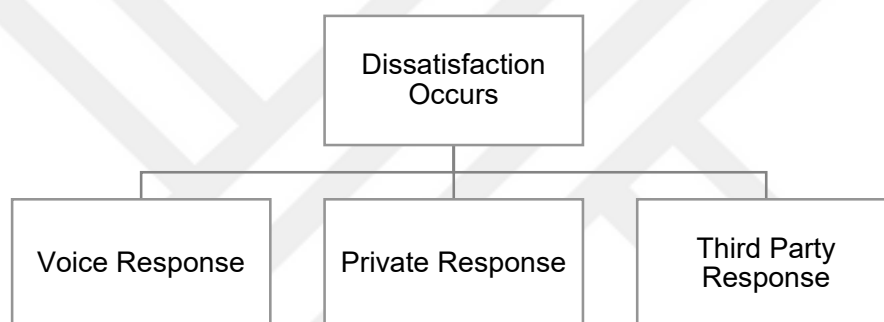


Source: Day and Landon (1977)

Singh (1988) developed a three-dimensional taxonomy that classifies customer complaint behavior responses according to the complaint's addressee. As indicated in Figure 2, Singh's model employs external-not external (to the social circle of the customer) and involved-not involved (in the dissatisfactory transaction) criteria for classifying customer responses following a dissatisfactory transaction. Voice behavior includes behavioral and non-behavioral responses directed at external parties to the

customer's social circle but involved in the dissatisfaction transaction (e.g., seeking redress from retailer or manufacturer). On the other hand, third-party responses are directed at external parties to the customer's social circle and not directly involved in a transaction (e.g., taking legal action, consumer agencies, media). Finally, private responses are directed at parties that are not external to the customer's social circle and not directly involved in the transaction (e.g., word-of-mouth communication with family and friends).

Figure 2: Singh's Taxonomy of Customer Complaint Behavior



Source: Singh (1988)

By considering the shift from goods-dominant logic to service-dominant logic, a more contemporary and dynamic view on customer complaints is put forth by Tronvoll (2012). Unlike the previous models that are grounded on the customer dissatisfaction resulting from the discrepancies between pre-purchase expectations and purchase experience, taxonomy put forth by Tronvoll grounds unfavorable experiences that may be: (1) a negative critical incident; (2) various negative incidents, league of which become critical; (3) negative evaluation of value-in-context during service delivery; or (4) an undesirable evaluation in post-purchase phase. According to Tronvoll's model, customers have three action options with a different threshold; (1) not complaining, (2) communicating the unfavorable experience, (3) taking complaint actions. The model, depicted in Figure 3, holds that complaining behavior is dependent on; (1) resource-

configuration and context including customer's resources (e.g., competence, time, finance, etc.), resources provided by the company (e.g., complaint channels, information about how to complain, etc.), (2) the context containing entire conditions surrounding the customer during complaining process (e.g., physical environment, market situations, etc.). If a change occurs in the context or the resource-configuration, customer's thresholds regarding communicative or action complaint responses may vary accordingly. It is suggested that in cases where the unfavorable experience does not surpass the complaining threshold, the customer will not take any kind of complaining action. In contrast, in cases where the unfavorable experience exceeds the threshold, the customer will engage in "communication complaint response" and/or "action complaint response."

Figure 3: Tronvoll's Model for Dynamic Complaining Behavior



Source: Tronvoll (2012)

1.3. FACTORS THAT INFLUENCE COMPLAINING BEHAVIOR

Extensive research has focused on the antecedents of customer complaint behavior, and the vast majority has highlighted the cognitive approach, which is grounded on the disconfirmation theory (Tse and Wilton, 1988). However, it is argued that over-emphasizing cognitive aspects may fail to take sufficient account of other aspects (Tronvoll, 2011). According to Edvardsson et al. (2005), the service experience

is a multifaceted process that creates cognitive, emotional, and behavioral responses. Besides, environmental and company-related issues impact the complaining behavior of customers (Lam and Zhang, 1999). At the broadest level, antecedents of complaining behavior can be divided into personal and non-personal factors.

1.3.1. Personal Factors

Personal factors influencing complaining behavior are divided into seven groups by Tronvoll (2008); demographical, psychological, personality-related, emotional, cultural, social, and experiential factors. Items belonging to each of these factors and studies investigating the impact of these items on complaining behavior are provided in Table 2.

Studies show that young and middle-aged consumers show more complaining tendencies than the older consumers (e.g., Andreasen and Best, 1977; Bearden, 1983; Hogarth and English, 2002; Moyer, 1985; Kolodinsky, 1993; Singh, 1990). Westbrook and Newman (1978) explain this inverse relationship between complaining intention and age based on more experience and skill in pre-purchase research. In terms of gender, women are more inclined to complain than men (e.g., Hustad and Pessemier, 1973; Kolodinsky, 1993; Solnick and Hemenway, 1992). Moreover, the literature highlights that consumers with higher income (e.g., Andreasen, 1988; Kraft, 1977; Mason and Himes, 1973) and education level (e.g., Crosier et al., 1999; Hogarth and English, 2002; Kolodinsky and Aleong, 1990) complain more. Other demographics that impact complaining intention are the place of residence (e.g., Crosier and Erdogan, 2001; Crosier et al., 1999), the number of households (e.g., Mason and Himes, 1973; Pfaff and Blivice, 1977), life cycle stage (e.g., Kolodinsky, 1993) and socio-economic level (e.g., Kraft, 1977; Spalding and Marcus, 1981; Tronvoll 2007b).

Psychological factors that impact customer complaining behavior include personal values (Rogers and Williams, 1990), personality (e.g., Fornell and Westbrook, 1979; Davidow and Dacin, 1997), attitudes towards complaining (e.g., Day, 1984; Bearden and Oliver, 1985), attitudes towards business (e.g., Jacoby and Jarrard, 1984), and personal confidence (e.g., Richins, 1983a). Among these constructs, the most important and widely investigated ones are personality and attitudes towards

complaining. Bolting (1989) found that customers' self-confidence and assertiveness impact complaining behavior and help them choose between negative WOM and complaining to management. Fornell and Westbrook (1979) indicated that assertiveness and aggressiveness impact complaint behavior through affecting how customers struggle with frustration, whereas other studies (e.g., Liu et al., 1997; Markus and Kitayama, 1990) pointed self-esteem as a psychological factor that positively impacts complaining behavior. Another stream of research (e.g., Malafi et al., 1993; Slama and Celuch, 1994) explained complaining as a way to respond to peer-pressure. Finally, various studies (e.g., Bearden and Oliver, 1985; Day, 1984) investigated the role of attitudes toward complaining in the consumer complaining behavior framework.

Attribution to blame and emotion can be considered as emotional constructs that impact complaining behavior. The extant body of knowledge points out a positive direct relationship between blaming attribution and complaint behavior (Folkes et al., 1987; Richins, 1983a; Stephens and Gwinner, 1998). The previous research on complaining behavior also reveals that emotions are a good predictor of complaining intention. Accordingly, Bolting (1989) highlighted that the emotional cost of failure has a positive impact on complaining, Folkes (1987) pointed out that anger mediates the relationship between attributions and complaining, and Tronvoll (2008) revealed that negative emotions predict complaining.

Other personal aspects impacting complaining behavior can be listed as cultural, social, and experiential factors. Various dimensions of culture are shown to be impacting consumers' complaining behavior following failure incidents (e.g., Liu et al., 2001; Liu and McClure, 2001). Consumers' social and political involvement (Robinson et al., 1982; Warland et al., 1975), employment status (Ash, 1978; Crosier et al., 1999; Liefeld et al., 1975), and the social climate (Jacoby and Jarrard, 1981) were found to impact complaining behavior. Finally, customer experience is a predictor of complaining intention (Day, 1984; Jacoby and Jarrard, 1981; Singh, 1990).

Table 2: Personal Factors Influencing Complaining Behavior

Factor	Items	Studies
Demographical	Age	Andreasen & Best (1977), Bearden (1983), Hogarth & English (2002), Moyer (1985), Kolodinsky (1993), Singh (1990b)
	Gender	Hustad & Pessemier (1973), Kolodinsky (1993), Solnick & Hemenway (1992)
	Income	Andreasen (1988), Kraft (1977), Mason & Himes (1973)
	Education	Crosier et al. (1999), Hogarth and English (2002), Kolodinsky and Aleong (1990)
	Place of Residence	Crosier & Erdogan (2001), Crosier et al. (1999)
	Number of Households	Mason & Himes (1973) Pfaff & Blivice (1977)
	Life Cycle Stage	Kolodinsky (1993)
	Socio-economic Level	Kraft (1977), Spalding & Marcus (1981), Tronvoll (2007a)
Psychological	Personal Values	Rogers & Williams (1990)
	Personality Traits	Bolfing (1989), Fornell & Westbrook (1979), Malafi et al. (1993), Slama & Celuch (1994), Liu et al. (1997), Markus & Kitayama (1990)
	Attitudes Towards Complaining	Bearden & Oliver (1985), Day (1984), Richins (1987), Singh & Wilkes (1996), Volkov et al. (2002)
	Attitudes Towards Business	Jacoby & Jarrard (1981)
	Personal Confidence	Richins (1983b)
Emotional	Attribution to Blame	Folkes et al. (1987), Richins (1983b), Stephens & Gwinner (1998)
	Emotions	Bolfing (1989), Folkes et al. (1987), Tronvoll (2008)

Cultural	Dimensions of Culture	Liu et al. (2001), Liu & McClure (2001)
Social	Social & Political Involvement	Robinson et al. (1982), Warland et al. (1975)
	Employment Status	Ash (1978), Crosier et al. (1999), Liefeld et al. (1975)
	Social Climate	Jacoby & Jarrard (1981)
Experiential	Customer Experience	Day (1984), Jacoby & Jarrard (1981), Singh (1990a)

Source: Adapted from Tronvoll (2008)

Following the aim and scope of this thesis, personality traits and emotions are further explained in the following sections. Additionally, sentiment polarity that reflects the personal tone used during articulation of the complaint is considered and further explained as a personal factor influencing complaining behavior. Since the above-mentioned studies mainly rely on the data gathered via surveys containing closed-ended questions, consumers' information dissemination tone that reflects how they feel about the incident could not be attained. However, when the social media environment is considered, it can be stated that it is critical to employ sentiment polarity for both better understanding their emotions regarding the incident and developing a broader understanding of the phenomena.

1.3.1.1. Personality Traits

Personality traits can be defined as the primary psychological constructs that form individuals' personalities (Harris and Mowen, 2001) and are mainly characterized by enduring, cross-situational consistencies in individuals' behavioral and response patterns (McCrae, 2009). In a similar vein, Lazarus and Folkman (1984) define personality as the fundamental psychological constructs and "makeup" of individuals that shape their actions and responses to the environment. Since understanding consumers is vital in building effective marketing strategies, consumers' personality traits have been widely examined in the marketing domain. The main research areas in which the impact of

personality traits is widely examined include loyalty (i.e., Lombart and Louis, 2012), customer satisfaction (i.e., Ekinci and Dawes, 2009), dissatisfaction (i.e., Mooradian and Olver, 1997), WOM behavior (i.e., Adamopoulos et al., 2018), post-purchase behavior (i.e., Mooradian and Olver, 1997), and repurchase intentions (i.e., Gountas and Gountas, 2007).

The first effort for creating a shared taxonomy was put by Allport and Odbert (1936), and through collecting 18,000 adjectives used for labeling individuals' behaviors that distract them from others, Allport and Odbert established four main groups; personality traits, temporary states, and moods, evaluative judgments of personal conduct and reputation, and physical characteristics and talents (as cited in John and Srivastava, 1999). A few years later, Cattell (1943) selected 4,500 traits from the list created by Allport and Odbert and reduced them to 35 variables by employing semantic and clustering procedures. By building upon Cattell's study, Fiske (1949) came up with a model consisting of 22 variables, commonly referred to as big five personality traits. In the year 1961, Tupes and Christal (1961) revisited Cattell's (1957) bipolar variables (Goldberg, 1992), and the widely used NEO Five-Factor Inventory that is being used for measuring personality traits is advanced by Costa and McCrae (1992). Even though a consensus on the actual number of cardinal personality traits has not been reached yet (Allport, 1961; Block, 1995), the five-factor model (FFM) is widely recognized by psychologists (Tupes and Christal, 1992; Matzler et al., 2005). The five-factor model organizes personality traits in hierarchical order and lists the five common traits as; *agreeableness*; *openness to experience* (intellect, imagination, or culture); *extroversion* (surgency), *neuroticism*; and *conscientiousness* (will to achieve) (Connor-Smith and Flachsbart, 2007).

Extroversion refers to being sociable, cheerful, and energetic (Mooradian and Swan, 2006) and is distinguished by affiliation, positive affectivity, power, and ambition (Matzler et al., 2005). Extroverts are mainly characterized by having positive emotions and being friendly and eager (Agarwal et al., 2016), and show a tendency to reflect their emotions to other individuals in response to stressors (David and Suls, 1999). Adjectives describing extroverts include "sociable," "forceful," "energetic," "adventurous," "enthusiastic," and "outgoing" (John and Srivastava, 1999). Extroversion facilitates the

pursuit of pleasurable experience, and individuals who score low on extraversion tend to be reserved, retiring, and cautious (Roccas et al., 2002).

Agreeableness denotes being straightforward, helpful, and trustful (David and Suls, 1999), and agreeable individuals can be characterized by being pleasant, likable, warm, and considerate by others (Graziano and Tobin, 2009). This trait explains an individual's quality of interpersonal orientation. Adjectives defining agreeable individuals include "forgiving," "not demanding," "warm," "not stubborn," "not show-off," and "sympathetic" (John and Srivastava, 1999). Agreeable individuals tend to be good in nature, cooperative, and care about motivational goals of conformism and traditional values, whereas those scoring low on agreeableness tend to be doubtful, short-tempered, and stubborn (Roccas et al., 2002).

Conscientiousness can be defined as an individual's degree of organization, perseverance, and motivation in goal-directed behavior, and individuals who score high in conscientiousness are mainly organized, reliable, and hardworking (Costa and McCrae, 1985). In other words, conscientious people are characterized by being dependable, reliable, achievement-oriented (Digman, 1990), efficient, responsible, and thorough (McCrae and John, 1992), whereas those who score low in conscientiousness can be defined as careless, apathetic, and undependable (Digman, 1990). Adjectives defining conscientious individuals include "efficient," "organized," "not careless," "thorough," "not lazy," and "not impulsive" (John and Srivastava, 1999).

Openness to experience denotes a willingness to divert new and progressive ideas (Agarwal et al., 2016). Openness indicates the ability and inclination to discover sensory and aesthetic information via perception, imagination, and artistic endeavor (Shiner and DeYoung, 2013). Individuals who score high in openness have creative minds, aesthetic sensitivity, intellectual curiosity, flexible thinking, a wide variety of interests, and various preferences (Digman, 1990; McCrae and John, 1992). Adjectives defining conscientious individuals include "curious," "imaginative," "artistic," "wide interests," "excitable," and "unconventional" (John and Srivastava, 1999).

Neuroticism can be defined as individuals' tendency to experience negative affect and subsequent emotional upheaval (Aggarwal et al., 2016). People who score high in neuroticism are likely to have psychological sufferings, unrealistic thoughts, extreme

desires or urges, and maladaptive coping responses (Costa and McCrae, 1985). Neurotic individuals can be characterized by being anxious, self-pitying, unstable, tense, and worrying (McCrae and John, 1992). Adjectives defining conscientious individuals include "tense," "irritable," "not contented," "shy," "moody," and "not self-confident" (John and Srivastava, 1999).

The extant body of knowledge reveals that personality traits play a crucial role in explaining complaining behavior following failures. Mooradian and Oliver (1997) found that consumers who score high in neuroticism are less likely to complain following a failure, but they are also found to be score low on repurchase intention following a negative experience. Complaint behavior has also been linked to agreeableness, and research showed that people who score high in agreeableness hold less propensity to complain (Kowalski and Cantrell, 1995; Kowalski, 1996). Another research conducted by Keng et al. (1995) revealed that complaining consumers are more assertive, confident, and strong, whereas non-complainers tend to be more conservative. Similarly, extant research validates that complainers are more assertive, self-confident, and aggressive (Richins, 1987; Phau and Sari, 2004), whereas non-complainers are submissive, quiet, and unconfident in nature (Bolfing, 1989; Fornell and Westbrook, 1979). Consumers who score high in openness to experience are also found to be more inclined to complain (Ekinci et al., 2016). Even though extroversion is the most examined trait in the five-factor model, a consensus about the direction of its impact on complaining has not been reached yet. One stream of research advocates that consumers who score high in extroversion have more tendency to complain. Harris and Mowen (2001) reported that consumers who score high in extroversion and conscientiousness hold a higher propensity to complain than those who score low on these traits. Donoghue and Klerk (2006) found that aggressive, extrovert, and altruistic individuals are more inclined to complain in a similar vein.

On the other hand, another stream of research advocates that extroversion has a negative impact on complaining propensity and justifies it with extroverts' social adaptability. Likewise, depending upon the positive correlation between positive affect and extroversion (Costa and McCrae, 1980; Watson and Clark, 1992), and the inverse relationship between positive affectivity and complaining behavior, it is advocated that there exists an inverse relationship between extroversion and complaining propensity.

Despite the mentioned disaccords in the direction of the relationship between personality traits and the tendency to complain, it is an agreed-upon fact that personality plays a vital role in determining consumers' propensity to complain.

1.3.1.2. Emotions

Emotions can be defined as a "mental state of readiness that arises from cognitive appraisals of events or thoughts; has a phenomenological tone; is accompanied by physiological processes; is often expressed physically; and may result in specific actions to affirm or cope with the emotion, depending on its nature and meaning for the person having it." (Bagozzi et al., 1999). In service failure literature, emotions can be described as consequences of a stressful cognitive assessment (Lazarus, 1991). Understanding consumer emotions is significant in various areas of marketing, since they impact information processing and goal setting, mediate consumers' reactions to persuasive appeals, exhibit goal-setting behavior, measures consumer welfare and impact of marketing stimuli (Bagozzi et al., 1999). Since service failures are negatively-valenced experiences in the nature (Bougie et al., 2003), the majority of the extant literature on the subject focuses on consumers' negative emotions (Williams and Aaker, 2002). However, it is revealed that an efficient recovery process may reduce the negative emotions (DeWitt et al., 2008) and make consumers recall their experience as a highly satisfying one (Bitner et al., 1990). Therefore, consumer emotions are widely examined in failure and recovery encounters. McColl-Kennedy and Smith (2006) list areas that are relevant to the study of customer emotions in service failure and recovery encounters as evaluation of the incident, identification of the main constituents of cognitive assessments and their impact on marketing stimuli, evaluation of post-purchase behaviors, identification of responses to emotions, detection of the impact of emotions on information processing, identification of emotion variability across cultures, determination of antecedents of emotional states, and examination of inter-relationships among emotions.

Since anger is one of the most frequently experienced emotions in the post-failure process, considerable emphasis has been made on studying anger. The way consumers express their anger is critical due to individuals' tendency to attack verbally and/or non-verbally (Fitness, 2000; Deffenbacher et al., 2002). Consumers' anger following failure

situations may result in complaining, negative WOM, boycotting, and exiting (DeWitt and Brady, 2003; Keaveney, 1995; Stephens and Gwinner, 1998), and such responses hold vital importance for companies since they result in reduced consumer loyalty, lower ROI (return on investment), and amplified consumer attrition (McColl-Kennedy and Smith, 2006). Another primary emotion that has similar valence to anger is fear. However, it should be noted that even if these two emotions have a similar valence, they hold distinct patterns of cognitive appraisal, especially in attributing responsibility (Tiedens and Linton, 2001). For instance, if consumers attribute the cause of a failure to be external and uncontrollable, they may respond with anger and disgust (Averill, 1983), whereas if they attribute the cause failure to be unstable, they may feel afraid and sad (Su et al., 2018). Extant literature also reveals that consumers who are dissatisfied with the service, experience some degree of disgust along with the previously mentioned negative emotions (Komunda and Osarenkhoe, 2012). Shi et al. (2011) revealed that consumers who feel anger, disgust, and surprise tend to engage in complaining behavior and prefer not to complain when they feel depressed or anxious.

Whereas anger, fear, sadness, and disgust are labeled as negative emotions, love, surprise, and joy are categorized as positive emotions (Lazarus, 1991; Shi et al., 2011). Extant research reveals that positive emotions are strongly linked with satisfactory experiences. Plutchik (1980) exposed that "delight" characterizes a mixture of surprise and joy, and expectedly, Westbrook and Oliver (1991) reported that consumers who score high in surprise and joy categorize their consumption experience as a satisfactory one. Mano and Oliver (1993) revealed a positive relationship between surprise and satisfaction in a similar vein. The inter-relationships between positive and negative emotions have also attracted the attention of researchers in the field. Plutchick (1980) found that a mixture of surprise and anger emotions is tapped by the word "outrage," whereas the combination of surprise and sadness is tapped by "disappointment." It should also be noted that the nature of affective reaction following an emotion may alter its positivity/negativity. Scherer (1984) denoted that when a positive affective reaction follows a surprise, we call it a "positive surprise," whereas when an adverse affective reaction follows it, we call it "negative surprise." For instance, an unexpected flight upgrade offered by airlines may be perceived as a "positive surprise" by the passengers, whereas the cancellation of the flight may be perceived as a "negative surprise."

Therefore, it may be stated that positive and negative emotions may co-exist due to individuals' cognitive structures and inter-relationships among emotions.

1.3.1.3. Sentiment Polarity

At the broadest level, sentiment may be described as the tone in which an individual articulates his/her subjective judgment. In the CCB domain, sentiment may be defined as "the semantic orientation or polarity of a word indicates the direction the word deviates from the norm for its semantic group or lexical field" (Hatzivassiloglou and McKeown, 1997). The semantic orientation (or tone) can indicate positivity or negativity. A positive semantic orientation denotes praise, whereas a negative semantic orientation signifies criticism (Turney and Littman, 2003). Consumers who experience a service failure may express their dissatisfaction through negative WOM, and expressions in such negative tones may offend service providers, profits, and company image (Bitner et al., 2000). Moreover, the expression of intense negative sentiment may trigger further negative comments and result in an online firestorm. Therefore, how consumers express their thoughts about the consumption experience and their evaluation's tone is essential.

Sentiment analysis aims to classify consumer reviews/comments based on consumers' emotions and accordingly identifying expressed sentiments in terms of being positive (favorable) or negative (unfavorable) (Agarwal et al., 2008). Reviews or comments that do not indicate any opinion may be categorized as "neutral" in terms of sentiment. In the marketing domain, sentiment analysis is used to create a market and/or business intelligence, discover opportunities and threats, and comprehend public sentiment articulated in various forms of textual communications (Kumar and Sebastian, 2012). By analyzing consumer opinions, companies can notice the satisfaction level or intensity of complaints. Following a service failure, analyzing the intensity of complaints may benefit companies in two ways; to prioritize the most critical complaints and to offer a recovery that can reverse the negative consumer judgment into a neutral or positive one. Accordingly, consumer opinions' sentiment is vital in analyzing negative experiences and providing an effective complaint handling process to complainants.

1.3.2. Non-personal Factors

Non-personal factors that impact customers' complaining behavior are divided into three groups by Tronvoll (2008); situational, service-related, and market-related factors. Items belonging to each of these factors and studies that investigated the impact of these items on complaining behavior are provided in Table 3.

Situational factors that influence complaining behavior following failure incidents can be investigated in two sub-groups; economic factors and seriousness of failure. Economic factors are mainly centered on the cost of failure for a customer. Bolting (1989), Landon (1977), and Singh and Wilkes (1996) advocated that customers compare the cost of failure with the benefit of complaining while deciding on the actions to be taken, whereas Francken and van Raaij (1985) supported that they make a comparison between cost and quality of the service. The second situational factor, the seriousness of the failure, is mainly investigated through problem seriousness, perceived alternatives, failure type and magnitude, and criticality concepts. According to Richins and Verhage (1985), the seriousness of the problem impacts customers' complaining behavior. Fornell and Didow (1980) advocate that customer response to a failure is a choice between the exit and complaining, and when the exit option is unavailable, complaining is the only way of reacting. Smith et al. (1999) found that type and magnitude of the failure play a vital role in the post-failure process and highlighted that customers prefer recovery offers that match the experienced type of failure in the appropriate amounts for the magnitude of the failure. Finally, Ostrom and Iacobucci (1995) and Webster and Sundaram (1998) revealed that failure's criticality substantially impacts customers' post-failure reactions.

Service-related complaint behavior antecedents can be divided into two groups; those related to the consumed service itself and those related to the service provider. Perceived importance of the service (Blodgett and Granbois, 1992), complexity and price (Day and Landon, 1977; Landon, 1977), and product/service type (Day, 1984) are service-related determinants of complaint behavior. On the other hand, store type (Day and Landon, 1977), reputation (Bearden and Mason, 1984; Granbois et al., 1977), responsiveness (Bolting 1989; Jacoby and Jarrad, 1981; Richins, 1983), the likelihood of success (Blodgett et al., 1993; Blodgett et al., 1995; Granbois et al., 1977; Richins,

1987), employees' responsiveness, friendliness, reliability and promptness (Zeithaml et al., 1988), and company size (Oster, 1980) can be listed as service-provider related determinants of the post-failure attitude of customers.

Degree of competition and industry structure are market-related constructs that impact customers' post-failure complaining behavior. According to Andreasan (1985), in a loose monopoly, quality-conscious customers are the ones that are most sensitive to failures and more likely to vocalize their thoughts and exit. In a similar vein, Hirschman (1970) advocated that competitive structure determines individuals' responses to failures of firms, organizations, and states. Tronvoll (2007a) revealed that in monopolistic markets, complainers mainly belong to lower socio-economic groups. Kolodinsky (1995) and Singh (1991) investigated the impact of industry structure on complaint behavior. Kolodinsky (1995) found that personal characteristics are critical in explaining complaining behavior for both auto repair and medical services; supply-side characteristics, on the other hand, hold an important place for explaining it only for auto repair services. Accordingly, the research conducted by Singh (1991) revealed that individuals' responses to service failures vary for different service industries.

Table 3: Non-Personal Factors Influencing Complaining Behavior

Construct	Factor	Items	Studies
Situational	Economic	Cost/Benefit of Complaining	Bolfing (1989), Landon (1977), Singh & Wilkes (1996)
		Cost/Quality	Francken & van Raaij (1985)
	Seriousness of Failure	Problem Seriousness	Richins & Verhage (1985)
		Perceived Alternatives	Fornell & Didow (1980)
		Failure Type	Smith et al. (1999)
		Failure Magnitude	Smith et al. (1999)
		Criticality	Ostrom & Iacobucci (1995), Webster & Sundaram (1998)

Service-Related	Service	Perceived Importance	Blodgett & Granbois (1992)
		Complexity/Price	Day & Landon (1977b), Landon (1977)
		Product/Service Type	Day (1984)
	Service Provider	Type of Store	Day & Landon (1977b)
		Reputation	Bearden & Mason (1984), Granbois et al. (1977)
		Responsiveness	Bolting (1989), Jacoby & Jarrard (1981), Richins (1983)
		Likelihood of Success	Blodgett et al. (1993), Blodgett et al. (1995), Granbois et al. (1977), Richins (1987)
		Employees' responsiveness, friendliness, reliability, and promptness	Zeithaml et al. (1988)
		Company Size	Oster (1980)
Market-Related	Market	Degree of Competition	Andreasen (1985), Hirschman (1970), Tronvoll (2007a)
	Structure	Industry Structure	Kolodinsky (1995), Singh (1991)

Source: Adapted from Tronvoll (2008)

1.4. ONLINE COMPLAINING BEHAVIOR

Online complaining behavior refers to customers' public opinion sharing activities on online platforms (Tripp and Gregoire, 2011). Since prospective buyers trust independent evaluators' opinions more than their circle of connections (Gelb and Sundaram, 2002), and consumers are more apt to engage with a third-party platform than complaining directly (Vermeulen & Seegers, 2009), online complaining and opinion sharing are at the center of attention for both prospective buyers and complainers.

Even though opinion-sharing and information/opinion-gathering are not novel concepts for marketing, the nature and context of online complaining differ from the traditional form, known as "word-of-mouth" (WOM). Traditional WOM refers to the process in which consumers share their "attitudes, opinions, or reactions about businesses, products, or services" with other individuals (Jansen et al., 2009). The introduction of Web 2.0 technologies provided new platforms for authorities who voluntarily share information, opinions, and WOM in an electronic form, which is called electronic word-of-mouth (eWOM). As e-WOM gained more popularity with the rise of Web 2.0 technologies, "market mavens" who enjoy sharing information and opinions with the people they know (Feick and Price, 1987) gave way to self-designated authorities to do it without personal contact. The anonymity and lack of personal contact in online platforms result in conversations in which both parties assume authority and identities that they cannot claim otherwise, and negative opinions are more apt to be communicated (Gelb and Sundaram, 2002).

The critical difference between online and offline complaining behavior is the fact that offline complainers are more likely to have the objective to redress, whereas the opposite is valid for online complainers (Andreassen & Streukens, 2013). Hennig-Thurau et al. (2004) listed the motives for writing negative eWOM messages as; social benefits, concern for other people, economic incentives, and personality traits of extraversion and self-enhancement. Other reasons highlighted in the literature are getting a solution to the existing problem (Thøgersen et al., 2009), venting negative emotions for reducing anxiety (Richins, 1984), prevent others from experiencing similar problems (Parra-Lopez et al., 2011), encouraging company to improve its business (Verhangen et al., 2013).

1.4.1. The Importance of Online Complaints

Even though interaction, collective experience, and feedback are considered essential characteristics of social media (Zerfass et al., 2014), studies reveal that only one-third of the online complaints receive a response from the related organizations (Maritz Research, 2011). Some companies, including American Airlines, US Airways, and Southwest Airlines, declared that they would not be accepting e-complaints due to

the cost of assigning numerous employees to complaint handling (Adams, 2000). However, since customers are susceptible to negative information (Lutz, 1975) and such cases can decrease sales more than twice as strong as positive WOM promoting sales (Arndt, 1967), customer complaints should be handled carefully.

Unlike the traditional complaining that occurs via telephone, mail, and other traditional ways of communication, online complaints voiced in companies and organizations' social media pages are available for public view, and these comments are accessible for all Internet users (Einwiller and Steilen, 2015). Hence, providing satisfactory responses to those complaints is vital for ensuring perception about the company and reputation. Moreover, if not appropriately managed, such cases may turn into online firestorms with other social media users' participation via enormous quantities of negative WOM messages (Pfeffer et al., 2014). A further negative incident that may also follow such unsatisfactory resolutions includes brand switching (Stauss and Seidel, 2004) and negative WOM (Ennew and Schoefer, 2003). The literature points out that customers who are not satisfied with the company's post-complaining actions share their experience with nine to ten people on average (Maurer and Schaich, 2011), and in some cases, this number increases up to twenty people (Tschohl, 2007).

On the other hand, making it easy for customers to complain is a critical issue (Hart, 1990), and providing multiple communication channels to customers plays a vital role (Ahmad, 2002) in understanding customers' experience with the company, and for having a chance to recover such negative incidents. The ability to identify and address complaints promptly provides substantial opportunities for building trust, strengthening relationships (Au et al., 2009), and increasing customer satisfaction (Mattila, 2001). Moreover, a satisfactory complaint handling process is strongly associated with customer trust and commitment (Kelley and Davis, 1994). Another critical point is that such cases, in which customers directly describe their experience with the product/service, provide companies an effective way of understanding the service encounter and company-customer relationships without a need for a third party (Mattila and Mount, 2003).

1.4.2. Online Complaint Channels

Recent developments in Internet technology and online communication have changed the pace of customer complaining. Online channels differ from traditional ones such as calling, writing a letter, or visiting a store to enable customers to access companies immediately (Strauss and Pesce, 1998). By building upon Day and Landon's (1977) customer complaining behavior typology, Mattila and Wirtz (2004) present two types of channels that stem from redress seeking: (1) interactive channels (e.g., face-to-face, telephone); (2) remote channels (e.g., letters, electronic messages). Even though the classification presented by Mattila and Wirtz (2004) provides the base for classification of complaint channels and involves both online and traditional channels, further updates are required due to the fast-changing nature of Internet technologies and increase in the usage of social media as a vital part of our daily lives.

There has been scarce research providing a typology of online complaint channels, and Bailey (2004) stands alone in categorizing this. According to Bailey (2004), online complaint channels can be divided into three categories: (1) company-specific hate sites developed by dissatisfied customers; (2) general complaint sites; (3) sites of related government agencies or non-profit organizations (NGOs). Even though it is the pioneering typology in the field, some critical issues remain to be addressed about this categorization. The established categories are determined mainly by considering the developers of the sites. The first category of sites is established by haters, the second one by impartial third-party companies to create a platform for articulating complaints, and the third one by governmental organizations or NGOs. This approach results in failing to capture one of the most powerful platforms that is social media. However, networking sites such as Facebook and Twitter are in the midst of our digital age (Greenberg, 2010) and are widely utilized by individuals for voicing their thoughts on purchased products and services. Therefore, the typology established by Bailey (2004) may be revisited by dividing online channels into two groups as (1) online complaint sites (including company-specific hate sites, general complaint sites, and sites of related government agencies or NGOs), (2) social media platforms (including all social media platforms in which customers can voice their dissatisfaction).

1.4.2.1. Online Complaint Sites

According to the online complaint channels typology put forth by Bailey (2004), online complaint channels can be divided into three main categories: (1) company-specific hate sites; (2) general complaint sites; (3) sites of related government agencies or NGOs. Company-specific hate sites are those that are established by dissatisfied and angry customers. On these sites, customers publicly share their complaints with other individuals, and instead of seeking redress, such sites act as a channel for informing and warning others not to purchase a product/service from the target company (Lee and Cude, 2011). Examples of these types of complaint channels include <http://www.paypalsucks.com> (PayPal), <http://www.stop-microsoft.org> (Microsoft), <http://www.hateverizon.org> (Verizon), <http://www.walmart-blows.com> (Walmart), <http://www.verizonpathetic.com> (Verizon), and <http://www.unitedpackagesmashers.com> (United Parcel Service) (Bailey, 2004; Lee and Cude, 2011). Wolrich (2005) conducted interviews with the founders of these sites and some of the target companies, and results show that, in most cases, such sites are founded due to either poor complaint handling or poor customer relationship management. For instance, Eric Gray, the founder of [paypalsucks.com](http://www.paypalsucks.com), explains the reason for launching the site as "I called PayPal to inquire about a straightforward question. I do not even recall what the problem was now, just a simple question. But the person I spoke with could not answer it. After repeated requests to get a supervisor online, they hung up on me. I called back, got a different person, and, believe it or not, they hung up on me, too. I felt like I was in the twilight zone." In a similar vein, John Westpah, the founder of [amexsux.com](http://www.amexsux.com), noted the reason as "Although I started the site because of a small complaint, now--two years and 20,000-plus postings later--I realize how much more widespread and serious American Express' wrongdoing is." From the target company side, two types of responses were recorded; understanding the customer and threatening the customer for closing the site. PayPal representative said that they support customers' right to express their opinion. On the other hand, some companies, including American Express and Microsoft, refused to interview Wolrich, founders of the hate sites claim that company representatives and lawyers threatened them.

The second category is general complaint sites, in which customers rate companies, and provide open-ended feedback messages so that their posts can be sent to target companies (Lee and Cude, 2011). These sites mainly have separate pages for each product or service provider, including all consumer feedback about the target company. All feedbacks and company responses (if available) are accessible to Internet users who visit those complaint sites. Posted complaints include all personal data withheld and are publicly available by the data they were posted (Bailey, 2004). Consumer Direct (a UK based complaint site), <http://www.complaints.com> (a USA-based complaint site), <http://www.sikayetvar.com> (a Turkey-based site) can be counted as examples of this category. The service offered by these general complaint sites are: (1) Helping customers make better-informed buying decisions; (2) Helping customers to reach better customer service from target companies; (3) Helping companies to provide better service (Bailey, 2004).

The third category of complaint sites contains platforms supported by various governmental agencies or NGOs (e.g., Consumer Center Malta in the EU, and econsumer.gov) to protect consumer rights (Bailey, 2004). Some of these sites aim to help customers reach their complaints to companies, whereas others (mainly government agencies) aim to find patterns that warrant consumer protection actions (Lee and Cude, 2011). [Econsumer.gov](http://econsumer.gov) is one of the best examples of this category, and this international platform features 35 consumer protection agencies worldwide intending to help customers in reporting complaints, provide online shopping tips against fraud, and tips on complaint resolution (Bailey, 2004). Examples of member agencies include DG for Consumer Protection and Market Surveillance (from Turkey), Federal Trade Commission (from the USA), and Competition Bureau (from Canada).

1.4.2.2. Social Media Platforms

Empowered by social media and mobile technologies developments, most customers prefer social media platforms for voicing their complaints in real-time (Gunarathne et al., 2017). Accordingly, the complaining process has transformed from a private matter between the company and customer to a public process tracked by all Internet users (Ward and Ostrom, 2006). The power shifted to customers in such a

setting due to companies' inability to control published content about their brand (Bruce and Solomon, 2013). Since these platforms enable quick and effective dissemination of information (Doerr et al., 2013), Twitter and Facebook are leading social platforms chosen by customers to voice their complaints.

In Twitter, which is a widely used microblogging platform, consumers can post about their purchase and post-purchase impressions about a target company with a vast number of connected individuals in real-time (Hennig-Thurau et al., 2015). Since high-speed connection with a vast number of Internet users makes Twitter a robust platform through which communicative dynamics between customers and companies can be analyzed in case of a failure situation, many companies, including JetBlue and Marriott International, put a great emphasis on complaints articulated on these channels (Abney et al., 2017). For instance, the senior director of public relations at Marriott International stated that "We would hear from people (on Twitter) and looked at it as a tremendous opportunity for customer recovery...It began to grow. It is something we felt was an extension of what we originally started." (Murray, 2011). As indicated, Twitter is a widely used and credited platform for customers, and also a vital platform for companies through the lens of which they can obtain critical customer insights. However, it should be noted that tracking customer insights from Twitter is not sufficient for ensuring customer satisfaction. Research shows that being able to view customer feedback in real-time comes at a cost, and 72% of the consumers expect companies to respond to their Twitter messages within an hour (Melancon et al., 2018). Therefore, tracking customer insights, managing online complaints, and providing proper responses can be counted as necessary conditions for reaching efficient Twitter usage for companies.

Facebook, a membership-based Internet community, enables members to share profile information, communicate with other members via public or private messages, share posts and photos (Pinto and Mansfield, 2012). As a result of these features, Facebook is an essential and widely used medium for self-expression. Therefore, an offline corporate wrongdoing or service failure can easily migrate to Facebook via angry posts or wall conversations (Champoux et al., 2012). The antecedents of the inclination to Facebook complaints are perceived unfairness and perceived likelihood of service recovery (Balaji et al., 2016) and redress seeking (Ma et al., 2015). Moreover, the

reasons behind choosing Facebook over the other platforms are customers' motives to vent, take revenge, and self-enhance because the platform is relatively an emotionally-charged one (Presi et al., 2014). Contrary to customers' expectations, most of the Facebook complaints are being followed by double deviation (Balaji et al., 2016; Tripp and Gregoire, 2011). Companies' failure to manage customer complaints properly makes customers experience double deviation and makes them feel wronged (Bitner et al., 1990).

1.5. ORGANIZATIONAL COMPLAINT MANAGEMENT

Organizational complaint management, or complaint handling, refers to companies' strategies for resolving failures to manage customers' post-purchase dissatisfaction (Tax et al., 1998). Hart et al. (1990) broadened this definition by making an emphasis on potential company benefits that may be obtained through the handling process and indicates that companies learn from failures to (re)establish company's reliability and reputation in the eyes of customers. Even though it is evident that the way companies handle the situation determines how complainants will remember the incident (Bitner et al., 1990), companies still need to enhance their organizational complaint handling processes. Research reveals that approximately half of the complainants are dissatisfied with the companies' complaint handling processes (e.g., Estelami, 2000; Grainer, 2003), and this brings us to the result that companies are not fully informed on how to handle failures successfully (Tax et al., 1998). However, for avoiding double deviation for customers and amplified negative evaluation (Bitner et al., 1990), providing an appropriate response to complainants is crucial.

One of the first efforts in explaining dimensions of organizational responses to customer complaints was put forth by Boshoff (1999), and the RECOVSAT scale that measures satisfaction with the service recovery was developed. The study results revealed six constructs (communication, empowerment, feedback, atonement, explanation, tangibles) to be employed for measuring satisfaction with transaction-specific recovery. Boshoff (1999) describes these six dimensions as:

- *Communication*: The way in which a company and its employees address customer complaints.

- *Empowerment*: Employees' access to resources and the decisions they are permitted to take.
- *Feedback*: The information provided about the reason for the incident and what is being done for resolving it.
- *Atonement*: Compensation provided by the company to restore inputs' ratio to outputs for reimbursing the inconvenience caused by the failure.
- *Explanation*: Explaining to the complainant about the reason for the encountered problem clearly.
- *Tangibles*: Appearance of the employees, the equipment they use, and the physical environment in which they handle complaints.

The other first-movers in explaining dimensions of organizational complaint management were Smith et al. (1999). The research revealed that failure context and recovery context impact perceived justice in the post-failure process, which, in turn, affects satisfaction with the service encounter. Smith et al. (1999) revealed four constructs (compensation, response speed, apology, initiation) could be employed in analyzing the recovery context. These four dimensions are described as:

- *Compensation*: Company's response to the inequity caused by the failure (discount, refund, coupons, etc.)
- *Response Speed*: Speed of responding to the complainant.
- *Apology*: Communicating politeness, concern, effort, and empathy to the complainant.
- *Initiation*: The beginning of the recovery process.

Building upon researches on complaint recovery frameworks by Boshoff (1999) and Smith et al. (1999), Davidow (2000; 2003) established an integrated framework for explaining organizational responses to complaints and their impact on post-complaint customer responses. According to Davidow (2000), dimensions of organizational responses (timeliness, facilitation, redress, apology, credibility, and attentiveness of the organization) impact customers' post-complaint satisfaction, which, in turn, impacts

WOM likelihood, WOM valence, and repurchase intention. Dawidow (2003) describes these six dimensions of organizational response as:

- *Timeliness*: The perceived speed of responding to complainant's feedback.
- *Facilitation*: The policies, procedures, and structure according to which the company handles the process.
- *Redress*: Benefits or outcomes the complainant obtains from the company as a response to the complaint.
- *Apology*: Acknowledgement by the company of the complainant's distress.
- *Credibility*: The company's willingness regarding presenting an explanation for the failure, clarifying what went wrong, and explaining what the company will do to ensure such an incident will not be experienced again.
- *Attentiveness*: Interpersonal communication and interaction between the company and the complainant that expresses respect, empathy, and willingness to learn about the complainant's concern.

Estalami (2000) investigated the impact of competition and complaint management procedures on satisfaction with complaint resolution. According to the model, competitive intensity and consumer loyalty constitute competitive determinants, whereas promptness, employee behavior, and compensation constitute procedural determinants of the complaint handling process. Complaint management procedures of Estalami (2000) commensurate in some aspects with the constructs articulated by Davidow (2003). Estalami's promptness construct is analogous with Davidow's timeliness, employee behavior with attentiveness, and compensation with redress.

Homburg and Fürst (2005) provided another framework for understanding organizational complaint management by building upon the behavioral theory of the firm. According to this framework, the quality of the process, behavioral aspects, and quality of outcome guidelines impact customers' justice evaluation that, in turn, impacts customers' satisfaction evaluations and loyalty. Moreover, it is indicated that the type of business and industry and supportiveness of the internal environment intervenes in the relationship between quality guidelines and customer justice evaluations. Homburg

and Fürst (2005) define the quality of process guidelines as “the degree to which a formal organizational procedure for registering and processing customer complaints exists and is consistent with complainants’ needs.” This construct captures timeliness, facilitation, and credibility dimensions of previously mentioned frameworks. Quality of guidelines is described as “the degree to which an explicit organizational policy for employees’ behavior toward complainants exists and is consistent with complainants’ needs,” and captures the attentiveness dimension. Finally, quality of outcome is defined as “the degree to which a formal organizational policy for providing compensation to complainants exists and fits customers’ needs” and captures empowerment, redress, and atonement dimensions of previous frameworks.

Finally, Gelbrich and Roschk (2011) provided a meta-analysis on organizational complaint handling and customer responses following this process. Their meta-analytic framework divides organizational responses into three categories; compensation, favorable employee behavior, and organizational procedures. Within this framework, Davidow’s (2003) redress and apology dimensions are referred to as “compensation” construct, attentiveness and credibility are referred to as “favorable employee behavior,” and facilitation and timeliness are referred to as “organizational procedures.” The model suggests that these organizational responses impact the justice perception of complainants that determines post-complaint satisfaction. According to the framework, loyalty and positive WOM are outcomes that are dependent on post-complaint satisfaction.

To sum up, all organizational responses to complaints (e.g., providing redress, favorable employee behavior, timeliness, proper procedures) shape complainants’ justice perceptions in terms of distributive, interactional, and procedural justice. These perceptions, in turn, impact customers’ post-complaint satisfaction. Satisfaction with the post-complaint process determines customers’ further behavioral intentions (e.g., loyalty, positive WOM).

1.5.1. The Importance of Effective Organizational Complaint Management

The post-purchase processed following customer dissatisfactions include complaining, negative WOM communication, a decrease in brand loyalty and

repurchase intention (Lapidus and Pinkerton, 1995). Even though a complaint situation following a dissatisfying experience may result in negative affect towards the company, a well-handled and satisfying post-complaint process may transfer the customer's disposition into a satisfactory encounter (Bitner et al., 1990; Gilly and Gelb, 1982).

Conducting an effective complaint handling process positively impacts customer satisfaction, commitment, and trust (Morgan and Hunt, 1994). The extant research on the subject validates the positive impact of a well-handled complaint management process on the company's customer evaluation. Customers seek an opportunity to articulate their dissatisfying experiences to the target company and accept apologies positively (Richins, 1979). Efficient complaint management is proven to create second-order satisfaction (Cambra-Fierro et al., 2015; Etzel and Silverman, 1981; Westbrook, 1987), increase loyalty towards the company (Etzel and Silverman, 1981; Westbrook, 1987), and such second-order satisfaction is found to yield higher levels of patronage (Gilly, 1987), enhance customer engagement (Cambra-Fierro et al., 2015), increase the chance of positive WOM from these customers (Blodgett and Anderson, 2000; Orsingher et al., 2010). Since keeping existing customers costs less than finding attracting new ones, the economic return of a well-handled process is simple to calculate (Fornell and Wernerfelt, 1987). Effective complaint management that takes place online can bear even more opportunities than those in traditional channels. The complainant's consequent positive reaction may impact other Internet users who observe the interaction positively, and thus, they may provide the opportunity to win back the complainant and new ones by shaping reputation (Einwiller and Steilen, 2015).

On the other hand, an inefficient complaint handling process results in further damage to the company's relationship with the complainant. Insufficient complaint handling practices harm the relationship and cause customer dissatisfaction (Gelbrich and Roschk, 2010). It can promote negative WOM intention causing potential customers to keep away from doing business with the company (Lapidus and Pinkerton, 1995), resulting in reduced trust and loyalty, reduced uptake of products and services, and customer attrition (Cook, 2012). Literature also suggests that an insufficient complaint handling procedure ultimately results in losing customers and thereby losing market share, while obtaining new customers through various marketing promotions costs a substantial amount of money and reduces profitability (Homburg and Fürst,

2005). Cook (2012) lists other results of a flawed complaint handling process as: (1) between 5% and 15% lower revenues, and consequent reductions in profit; (2) reputational risk resulting in criticisms by other customers, NGOs, media, consumer associations, and trade organizations; (3) risk of non-compliance with relevant regulations and sanctions that result in legal punishments; (4) employee morale can be affected negatively, or staff capacity can be wasted with repeated complaints; (5) inability to reaching root causes of problems that can create further dissatisfactions.



CHAPTER TWO

SOCIAL MEDIA AND SOCIAL CUSTOMER RELATIONSHIP MANAGEMENT (CRM)

2.1. DEFINITION OF SOCIAL MEDIA

Social media is lying at the heart of our daily lives, and hence is an everlasting hot topic both for the practitioners and academia. As of 2019, 4.4 billion people use social media actively across several devices, global daily social media usage is 2 hours and 18 minutes, and in one social media minute, 41.6 million messages are being sent over messengers, 347,222 users are scrolling Instagram, 87,500 people are tweeting, and 4.5 million videos are being watched on YouTube (Viens, 2019). Even though the extensive global usage and the prominence of social media platforms is an agreed-upon fact, due to the divergence regarding the key elements to be included, the consensus on the definition of “social media” has not been reached yet (Ahlqvist et al., 2010). Both academicians and practitioners seem to be uncertain about what should be included under the definition and how social media varies from Web 2.0 and user-generated content (UGC) (Kaplan and Haenlein, 2010). Hence, before defining social media, Web 2.0, and user-generated content should be explained.

Web 2.0 can be described as tools and services that allow people to publish and communicate in ways that were previously only available to organizations, developers, or enthusiasts (Cooke and Buckley, 2008). The term “Web 2.0” was phrased by Tim O’Reilly and Media Live International (Levy, 2009). Musser and O’Reilly (2006) define this revolution as a combination of economic, technological, and social trends that constitute the new Internet generation, which is more mature and attaches importance to user involvement, and states its motto as “building applications that harness network effects to get better the more people use them.” While applications such as personal web pages and content publishing were extant in the era of Web 1.0, Web 2.0 replaced them with blogs, wikis, and collaborative projects (Kaplan and Haenlein, 2010). Subsequently, what makes Web 2.0 different from its predecessor is its participatory and collaborative nature that enables individuals to create and modify content and applications (Cooke and Buckley, 2008).

User-generated content (UGC) can be defined as online content “created outside of professional routines and practices” (OECD, 2007). The critical fact about UGC is that internet users play the most crucial role by being the producers and consumers of content concurrently (Mendes-Filho and Tan, 2009). Even though earlier definitions of the concept highlight the control shift from company-generated to user-generated nature of online content (Hennig-Thurau et al., 2010), more recent definitions shed light on several key features (Malthouse et al., 2016). Online content should fulfill three essential requirements to be considered as user-generated content: (1) should be made available through a publicly accessible media or a social networking site; (2) should reflect a certain amount of creative effort; (3) should be created “outside of professional routines and practices” (Christodoulides et al., 2012; Kaplan and Haenlein, 2010). Even though the concept of user-generated content was available before Web 2.0, noteworthy changes in technological drivers (i.e., variety of daily-used products with internet access), economic drivers (i.e., increased availability and decreased cost of tools for creating such content), and social drivers (i.e., increased need for online engagement) make today’s user-generated content understanding essentially different from what was observed previously (Kaplan and Haenlein, 2010).

As indicated previously, an agreement on the exact definition of “social media” has not been reached yet. The vast majority define social media as a set of software. Kaplan and Haenlein (2010), Berthon et al. (2012) describe it as a set of Internet-based applications built on Web 2.0 foundations, enabling the formation and exchange of UGC such as text, pictures, videos, and networks. In a similar vein, Parr (2008), Alt and Reinhold (2016), and Quesenberry (2018) explain the term as a set of computer and Internet-mediated technologies that support the production and sharing of various types of expressions and experiences in an efficient way. Another set of researchers describe social media as a contact between creators and users of information. Accordingly, Lewis (2010) and Kietzmann et al. (2011) include both channels and content in the definition, through which producers and users of information connect and interact. Heggde and Shainesh (2018) brings these two views together by defining social media as “the online media, enabled by Web 2.0 technologies that allow multidirectional conversations and real-time interactions, and encompass a gamut of platforms such as

social networking sites (SNS), blogs, discussions forums, video sharing, content sharing, social bookmarks, podcasts, and wikis.”

2.2. HISTORICAL DEVELOPMENT OF SOCIAL MEDIA

Even though the term “social media” is mainly described through digital mediums of communication, its history dates back to the 1950s. According to Borders (2010), phone phreaking, which refers to unauthorized access to telephone systems through homemade electronic devices, enabled phreaks to sneak into unused corporate mailboxes for hosting the first blogs and podcasts. In the 1960s, E-mail was invented as a way of transferring messages between computers. However, it was unavailable to the public until 1991 (Edosomwan et al., 2011). In the late 1960s, ARPANET and CompuServe were developed for providing a network of time-sharing computers (Borders, 2010).

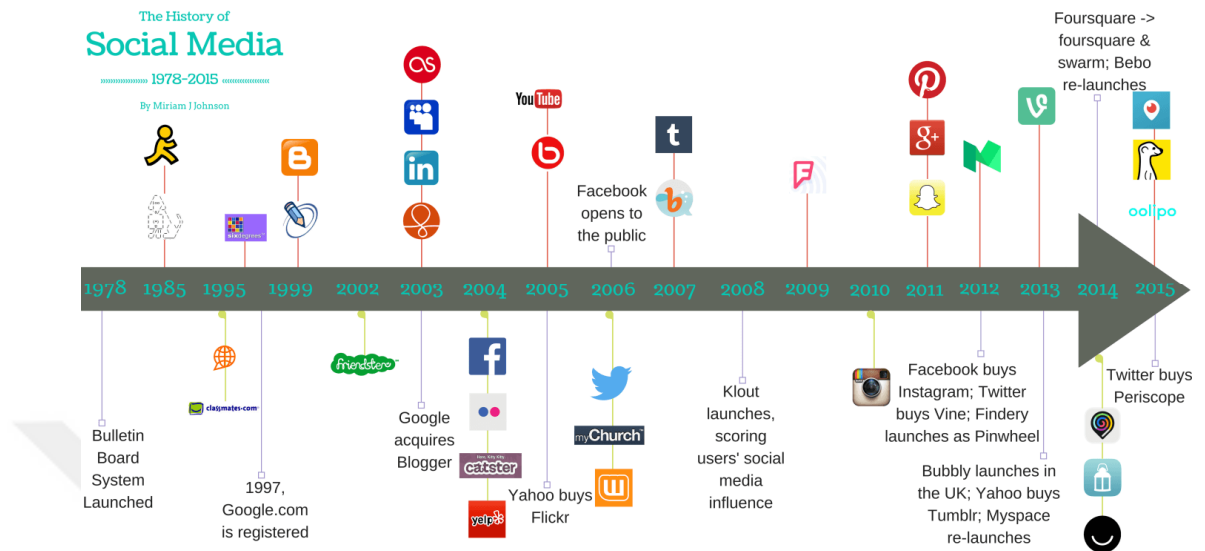
The 1970s and 1980s were the transition era from relatively primitive computer technologies to a more enhanced one. In 1978, Multi-User Domains (MUD) used for real-time game playing and chatting, and Bulletin Board Systems (BBS), which can be assumed as the father of the World Wide Web, used for uploading and downloading content, reading news, and exchanging messages were invented (Edosomwan et al., 2011). In 1980, the Usenet system that operated without a central server was released to post articles and news (Ritholz, 2010). In the late 1980s, the real predecessors of social media came to the scene. Listserv, a mail listing software that allows the sender of a mail to send one message to several people, was released in 1986 (Edosomwan et al., 2011). In 1988, Internet Relay Chat (IRC), which is a platform for real-time messaging (chatting), was launched (Ritholz, 2010).

In the year 1991, the World Wide Web invention opened a new page in history and changed the way we communicate by forming a novel basis for network communication (Van Dijk, 2013). The idea of creating online social sites for enhancing the interaction between individuals accelerated in the 1990s. Concordantly, the first social network site, SixDegrees, which allows users to create profiles, and list friends, was launched in 1996 (Boyd and Ellison, 2007).

The early 2000s were the blooming era of social networking sites with various new players enabling people to share their common interests in music, movies, friendship, and education (Edosomwan et al., 2011). In 2002 Friendster launched as an online dating site designed for helping friends-of-friends meet, and reached 300,000 uses within a year (O'Shea, 2003). In 2003, Wikipedia, the world-famous online encyclopedia, and LinkedIn, the online platform for bringing professionals together, were launched. In February 2004, a groundbreaking social networking site named Facebook joined the game by extending its membership requirements from Harvard-students-only to everyone-that-is-13-and-older (Edosomwan et al., 2011). In 2005, the era of user-generated content was crowned with the launch of Youtube and Reddit. In the subsequent year, one of the most famous microblogging sites, Twitter, was launched. Tumblr, a social platform for sharing every kind of post, including text, photos, music, and videos, was launched in 2007. In 2009, Whatsapp's launch changed the way of daily communication and opened a brand-new era in our daily lives.

In analogy to the 2000s, the early 2010s were also fertile regarding the launch of new social media platforms. In 2010, Instagram and Flickr were launched. Instagram positioned its features of like button and hashtag for creating a technical idiom that is compatible with Facebook's and Twitter's (Van Dijk, 2013). Flickr is also a photo-sharing platform, whereas it differentiates from Instagram by allowing users to track more detailed photography details about photo specs. In the subsequent year, 2011, another photo and video sharing platform called Snapchat was released and brought the concept of sharing photos that become invisible after a certain period of time determined by the account owner. Vine, a 7-seconds-long video sharing platform, was released in 2013. In 2015, Foursquare, a location and information sharing platform, and Periscope, a live video streaming platform, was launched. Lastly, in 2016, TikTok, a platform for sharing short comedy videos, was released. The chronological history of social media platforms is depicted in Figure 4.

Figure 4: The Chronological History of Social Media



Source: Allen (2017)

2.3. DIFFERENCES BETWEEN SOCIAL MEDIA AND TRADITIONAL MEDIA

Traditional media represents passive communication by adopting a one-to-many communication approach and contains the radio, television, and print as the main mediums of communication (Heggde and Shainesh, 2018). On the other hand, social media differs from traditional media by leaving more space for interaction and active participation of the users. The distinction between traditional and social media is multifaceted, and hence, goes beyond the variation in employed mediums of communication. Apuke (2016) described the main differences based on the properties of traditional and social media as:

Expediency: When compared with traditional media, social media mediums are accessible at a lower cost.

Usability: Content creation for the traditional media mediums requires some special skills in addition to training, whereas social media requires only modest reinterpretation of extant skills.

Feedback: Feedback from the audience is instant in social media while delayed in traditional media.

Permanence: Social media is more flexible in altering published content through comments or editing, whereas traditional media content cannot be altered.

Interactivity: Communication in traditional media is passive, while in social media, users have the opportunity of interacting easily and creating their content.

Dispersal: Unlike traditional media, social media is less centralized, and therefore more responsive to individual choices.

Content Production: Social media provides an opportunity to become a publisher to everyone, whereas traditional media is reserved for content creation and broadcasting professionals.

In addition to the above-stated differences, communication varies significantly between social and traditional media. Social media's interactivity enables two-way communication and content control, whereas the one-way nature of traditional media communication provides limited navigation options and a more static communication environment (Quesenberry, 2018). The limited navigation in traditional media represents a one-to-many communication approach, which is also known as passive communication, and it is based on transmitting a given message through mediums (Heggde and Shainesh, 2018). Keim and Noji (2011) define the divergence by making an emphasis on the interactive nature of social media and states that social media contains multidirectional information flow, low information control, high adaptability, collective intelligence, and individual empowerment, while traditional media contains unidirectional information flow, high information control, low adaptability, proprietary intelligence, and organizational empowerment.

2.4. SOCIAL MEDIA PLATFORMS

The definition of social media entails technologies enabled by Web 2.0 and concordantly contains a wide variety of platforms such as podcasts, networking sites,

video sharing, social bookmarks (Heggde and Shainesh, 2018). For gaining an in-depth understanding of the characteristics of these diverse platforms, a further distinction is required. Therefore, in this section, characteristics of blogs, microblogs, social networking sites, content sharing platforms, news, and bookmarking platforms, rating and review sites, forums, and virtual worlds are explained.

2.4.1. Blogs

Blogs, abbreviated from the term Weblogs, are content management websites that encompass the chronological order of posts in a diary-like structure and contain multimedia content on a wide range of subjects (Quesenberry, 2018). These journals enable interactive communication through text, image, video, and audio and are arranged in reverse chronological order (Huang et al., 2007). Blogs can be considered the social media equivalent form of personal web pages usually managed by a single person and enable interaction through comments (Kaplan and Haenlein, 2010).

The first blog dates back to the year 1994, in which Justin Hall, a college student, started writing about video games and gaming; in 1997, the concept of “weblog” which is a combination of the words web and log were officially born, and in 1999 LiveJournal and Blogger were released (Zarrella, 2009). 2003 was a vital year for bloggers since open-source blogging platform WordPress and TypePad were released, The Oxford English Dictionary included the “WebLog” term, and Google purchased Blogger (Quesenberry, 2018). The release of Tumblr, a blogging platform allowing multimedia content sharing, in 2007, was followed by the launch of Medium in the year 2012 (Quesenberry, 2018).

Blogs can be divided into four main categories; individual blogs, corporate blogs, collaborative blogs, and media blogs. *Individual blogs* contain personal thoughts on a subject in the form of an online diary; *corporate blogs* are owned and edited by representatives of legal entities; *collaborative blogs* are written on a specific topic and updated regularly by a collection of users; and *media blogs* concentrate on various types of content and are named after the content they produce (i.e., video blogs are named vlogs) (Thevenot, 2007). Each of these blog types may have its further sub-branches according to either the expertise of the blogger (expert vs. non-expert) or the

sphere of the blog (public sphere including newspapers and broadcasters vs. private sphere including lifestyle and travel blogs) (Youngs, 2009).

2.4.2. Microblogs

Microblogs, or micro sharing sites, is a form of blogging in which content size and length is limited (Tuten and Solomon, 2018). Unlike traditional weblogs, microblogs limit the posted content to smaller bits such as short sentences, sentence fragments, images, videos, or links (Quesenberry, 2018). Ebner (2009) lists characteristic features of these platforms as; (1) limited characters for each entry, (2) providing the ability of blogging via mobile devices, web interface or desktop, (3) a social network consisting of followers and followed accounts, (4) possibility of sending private messages to followers, (5) being able to answer those private messages and initiating discussions. Additionally, restriction on the employed characters favors the use of abbreviations (e.g., LOL instead of laugh out loud), emoticons (e.g., emoji and smiley), and references (e.g., hashtags) in microblogs (Alt and Reinhold, 2016).

The first microblogging platform that drew remarkable attention was TXTmob, which was used in 2004 by political convention protesters to communicate through short text messages, and shortly after TXTmob, Jack Dorsey launched an SMS broadcast system called Twitter and confessed that the idea was inspired by TXTmob (Zarrella, 2009). As of 2007, more than 100 microblogging platforms, including Plurk, Sina Weibo, and identi.ca, were in operation, but Twitter is by far the most well-known today (Tuten & Solomon, 2018). One of the main reasons behind the popularity of Twitter is the fact of being used as an instant reporting platform and source of news during crises, and acting as a vital source of information for individuals (Quesenberry, 2018).

The network-based subscription and replying option make microblogging a novel hybrid communication method that can be considered a blend of blogging, instant messaging, and status updates (Passant et al., 2008). Consequently, social media platforms of various other types may be misclassified as microblogs. For preventing such misclassifications, it should be noted that platforms such as Facebook and LinkedIn are social networking sites that contain a microblogging feature of status updates (Quesenberry, 2018).

2.4.3. Social Networking Sites (SNS)

Social networking sites (SNS) are online hosts, which enable its members to create and maintain personal user profiles that provide the ability to identify other users, and participate in consumption, creation, and/or communication of content (Kaplan and Haenlein, 2010; Tuten and Solomon, 2018). SNS enables members to create public or private profiles within a bounded system, list the connections (followers, friends), and view and manage their connections (Boyd and Ellison, 2007). Personal profiles of members may include various types of information and content, including text, video, audio, and photo (Kaplan and Haenlein, 2010). Members maintain their relationship with connections through sharing their availability, photo, mood, friend list, and status, and this content may be shared through diverse ways, including direct messages, posts, or instant messaging (Tuten and Solomon, 2018).

The history of SNS dates back to the 1980s' bulletin board systems (BBS) that enabled users to send private messages and share public posts. Zarella (2009) explains the development of SNS in the modern sense as: "The late 1980s and early 1990s witnessed the growth of the desktop applications such as CompuServe, Prodigy, and AOL that allowed users to connect to the Internet, and create profiles, post events, chat, and send messages. As the World Wide Web gained popularity, social networking moved to web-based applications. Concordantly, in 1995, Classmates.com and Match.com were created. The modern era of social networks began in 2002, with the launch of Friendster. In 2003, MySpace was released, and within the same year, Mark Zuckerberg created Facemash, a website letting students compare two photos to select the more attractive one. Later, Zuckerberg created 'The Facebook,' a social network that began as an exclusive site for college students." As indicated, the early 2000s were the blooming era of SNS in the modern sense. In addition to MySpace and Facemash, LinkedIn is a business-oriented networking application enabling members to share their professional experience and form connections with other professionals was released in 2003 (Quesenberry, 2018).

Social networking sites may be categorized into three groups according to the aim of usage: (1) Individuals (i.e., Facebook, Google+), (2) Business (i.e., LinkedIn,

Xing), and (3) Science (i.e., ResearchGate) (Alt and Reinhold, 2016). Even though a foundational distinction between SNS types can be articulated, further research is still being conducted for revealing how individuals use these platforms, for what purposes are SNS being used, and how these platforms shape the way of thinking, working, and communicating (Heggde and Shainesh, 2018).

2.4.4. Content Sharing Platforms

Content sharing platforms enable users to share various media types, including audio, photos, videos, presentations, text, and documents (Heggde and Shainesh, 2018; Kaplan and Haenlein, 2010). These platforms, similar to other social media platforms, host content but also feature the media types mentioned above types rather than text or a mix of media (Tuten and Solomon, 2018). Another feature that distinguishes content sharing sites from other social media platforms is that users are not required to create a personal page, and if they do, the profile only contains necessary information or statistics about the user (Kaplan and Haenlein, 2010). What makes these media sharing sites “networked” is the option of following content posted by specific users (Tuten and Solomon, 2018).

The first content-sharing site, IFILM.net, was released in 1997 as an online assortment of short videos streamed through various browser plug-ins such as Windows Media Player, QuickTime, and RealPlayer, and later in 1999, photo-sharing platforms such as Ofoto and Shutterfly were launched (Zarrella, 2009). With the spread of Web 2.0, high-speed Internet, digital cameras, and other recording mediums, content sharing platforms gained remarkable momentum. The blooming process was initiated with the release of Youtube, a video sharing platform with more than a billion users today, followed by the launch of photo-sharing application Instagram in 2010, and continued with the release of Snapchat in 2011 (Quesenberry, 2018). Other remarkable media sharing platforms can be listed as; Flickr, Vimeo, SoundCloud, Slideshare, and Scribd.

Major content sharing platforms can be grouped according to shared media as (1) Photo sharing platforms (i.e., Instagram, Flickr, Snapchat); (2) Video sharing platforms (i.e., Youtube and Vimeo), (3) Audio sharing platforms (i.e., Soundcloud and

Audiofarm), (4) Document sharing platforms (i.e., Scribd, SlideShare) (Tuten and Solomon, 2018; Heggde and Shainesh, 2018). Although each of these platforms holds its features, the defining characteristic of connectedness with other users and content sharing leads them to be a social media part.

2.4.5. Social News and Bookmarking

Social news sites enable users to submit and vote content extant on the Web to isolate the most interesting ones (Zarella, 2009), whereas social bookmarking platforms enable users to save links to access them whenever they like (Tuten and Solomon, 2018). Social news and bookmarking platforms are both enabling classification of content for later consumption that covers reading, sharing, and promoting (Heggde Shainesh, 2018). The main idea behind these collaborative projects is that many individuals' joint effort leads to a more enhanced outcome than any of them could individually achieve (Kaplan and Haenlein, 2010). These two platforms are very akin, with social news being more concentrated on content distribution and social bookmarking being focused on the content organization (Tuten and Solomon, 2018).

Social news and bookmarking platforms play a significant role for both individuals and companies for two main reasons; trend analysis and content curation. Collaborative tagging that establishes links to resources is widely used in trend analysis and determination of content popularity (Quesenberry, 2018). Such analysis allows companies and individuals to track main trends and chronological movements and enable users with similar taste or interest to connect. On the other hand, content curation enables the collection of information relevant to a specific topic or area of interest (Quesenberry, 2018). This process helps people determine which sites to visit among many options and act as a valuable and refined source of information for following developments in the area of interest.

The first bookmarking platform offering public and private storage of favorite links, ITList, was launched in 1996, and the release of Slashdot in 1997 followed it, and by Fark, a platform enabling users to upload links 1999 (Zarella, 2009). Delicious, the first bookmarking site in the modern sense, was launched in 2003 (Quesenberry, 2018). Following Fark, Digg, a social news platform integrating news and publishers' streams

through voting, was founded in 2004. Furthermore, in 2006, Reddit and BuzzFeed, the most prominent segment players, were launched (Quesenberry, 2018).

2.4.6. Rating and Review Sites

Reviews can be defined as reports indicating individuals' opinions regarding the product, service, or performance, whereas ratings refer to measuring how good/bad something is (Quesenberry, 2018). Users of rating and review sites can post questions, comments, ratings, experience, or comment on other users' posts to share insights and knowledge on a specific product or service (Heggde and Shainesh, 2018). Since they provide a natural environment for articulating the real feedback (Zarella, 2009), such platforms are considered the key for obtaining consumer insights, achieving better performance, and attaining a good reputation (Nguyen and Coundounaris, 2015).

The concept of reviewing and rating products or services called Amazon Reviews was first released by Amazon.com in 1995 by enabling individuals to anonymously review and rate products available on the site on a scale of one to five, and Citysearch.com followed the trend by providing users the ability to review various city attractions (Zarella, 2009). Until the year 2000, Amazon.com and Citysearch.com remained as the leading players in reviewing and rating. In 2000, online booking and review company TripAdvisor enabled its users to add their comments and ratings regarding their experiences, and user-generated reviews surpassed professional ones by reaching "over 500 million reviews covering 7 million accommodations, restaurants, and attractions all around the world in 28 languages" (Quesenberry, 2018). In 2004, Yelp, a platform for rating and reviewing local businesses, was released and has grown to 188 million visitors per month and 142 million reviews (Zarella, 2009; Quesenberry, 2018).

2.4.7. Forums

Forums are online discussion platforms allowing users to create content representing a current discussion and hold them on related topics via published messages (Alt and Reinhold, 2016; Quesenberry, 2018). These platforms can be

considered modern and interactive bulletin boards focused on discussion, and users contribute to these communities by answering previous posts and providing useful information (Zarrella, 2009). Forms differ from blogs in that forums are created by numerous users' participation (Alt and Reinhold, 2016), and from chat rooms in that conversations occur more like email conversations (Zarrella, 2009). Users create profiles and contribute to conversations via posting questions, responses, news, and photos that result in a threaded discussion (Tuten and Solomun, 2018). These threads may be anything from short statements to longer opinions or how-tos (Quesenberry, 2018).

The first platform that can be described as a forum, Usenet, was established in 1979 as a joint project of Duke University and the University of North Carolina (Tuten and Solomun, 2018). Usenet, which still exists, started as a mailing list that allowed users to post articles and responses, and much of the slang and the concept of spam were first appeared there (Zarrella, 2009). In contrast to other social media platforms, there are thousands of forums all over the world. However, the most popular forums can be listed as Reddit, 4chan, and 2Channel. Besides their members, these platforms attract companies and researchers' attention due to being a fertile source of analyzing individuals' thoughts and behaviors. Moreover, according to Hagel and Armstrong (1997), the most vital profit-making potential on the Internet can be found on such platforms that contain vital insights about consumer behavior.

2.4.8. Virtual Worlds

Virtual worlds may be described as platforms that imitate a three-dimensional setting in which users are enabled to take place in the form of their personalized avatars and interact with other avatars (Kaplan and Haenlein, 2010). These platforms have three main aspects that differentiate them from other social media platforms: (1) Users can interact in real-time; (2) Virtual worlds enable users to express themselves with fully customized avatars; (3) Users can explore the virtual world in three dimensions (Kaplan and Haenlein, 2009).

Virtual worlds can be divided into two as; game worlds and social worlds. Virtual game worlds consist of social multi-player games that are competitive and goal-oriented

and involve engagement rules that yield both attention and active involvement (Tuten and Solomun, 2019). Even though the rules of these virtual games usually restraint the degree of self-presentation and –disclosure, when users spend the majority of their time with the games, their avatar starts resembling their real-life personality (Kaplan and Haenlein, 2010). All in all, these platforms can be summarized as digital worlds, in which one can create one's avatar and collect digital possessions (Heggde and Shainesh, 2018). The most famous examples of virtual game worlds include World of Warcraft, Azeroth, EverQuest, and Norrath. The second type of virtual world, social worlds, offer an unlimited range of self-representation by enabling its users to behave more freely and create a virtual life similar to their real-life (Kaplan and Haenlein, 2010). Such communities offer their users a virtual life that can satisfy real human emotions and desires and make users progressively move to an atomistic lifestyle with a self-serving mindset from the earlier nuclear lifestyles (Heggde and Shainesh, 2018). The most prominent examples of social worlds are Second Life and Sims.

Even though the history of virtual worlds dates back to the late 1990s and early 2000s, the concept dates back to 1992, the year in which a groundbreaking novel was published. The cyberpunk novel *Snow Crash*, written by Neal Stephenson, inspired many web technologies, including Google Earth and the usage of avatars (Zarella, 2009). One of the book's most attention-grabbing elements was the Metaverse, a platform upon which people can build their realities. Drawing upon Metaverse, Ron Britvich created *Alpha World* in 1995, Finnish *Habbo* was founded in 2000, and Philip Rosedale launched *Second Life* in 2003 (Kaplan and Haenlein, 2009). The trend was followed by the launch of other virtual worlds such as World of Warcraft and Sims Online.

2.5. SOCIAL CRM CONCEPT

Relationships between brands and their consumers are formed via either experiences or one-to-one interactions that mainly occur in the form of a cycle, which is a part of numerous stimuli and responses (Heggde and Shainesh, 2018). Company-consumer interactions that constitute the relationship may occur at the point of experience, in social media, or both. The remarkable upsurge in social media usage

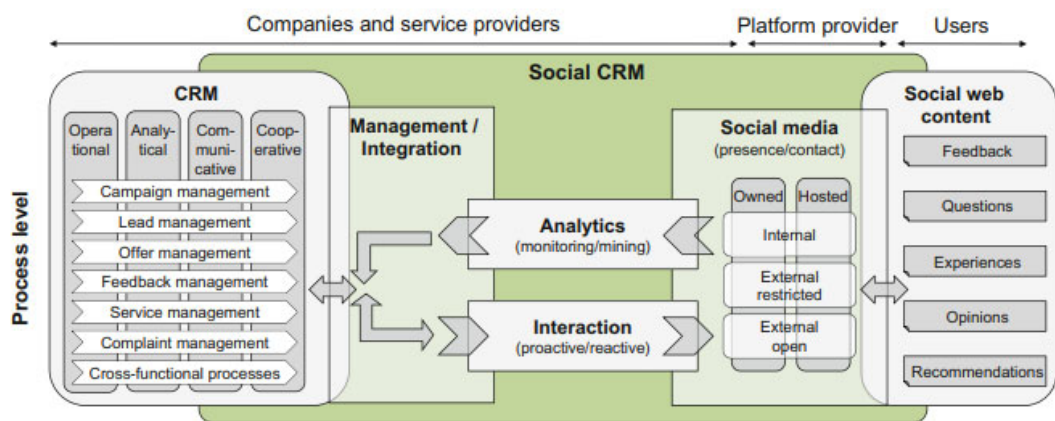
led to the diversion of consumers' attention to these platforms and caused the emergence of the "social customer" concept that made practitioners and researchers reconsider the way of managing customer relationships (Greenberg, 2010). Following the pervasive social media usage, the traditional CRM concept broadened its perspective through recognizing novel technological capabilities and social shifts brought by social media applications (Trainor, 2012). Even though traditional CRM solutions remain aggregating and analyzing consumer data for optimizing business processes, social CRM solutions contribute to these processes by engaging with social consumers, elevating their experience with a brand, and providing flexible solutions (Band and Petouhoff, 2009).

In the broadest sense, social CRM, collaborative CRM, or CRM 2.0, refers to the interaction between brands and their customers on social media platforms and the linkage of these platforms to CRM processes and systems (Alt and Reinhold, 2016). Greenberg (2010) incorporates the change brought by the technological developments to the traditional CRM concept and defines the social CRM concept as "*a philosophy and a business strategy, supported by a system and a technology, designed to engage the customer in a collaborative interaction that provides mutually beneficial value in a trusted and transparent business environment.*" This new and rapidly evolving concept of CRM reshapes the way of building and maintaining customer relationships within the contemporary business and IT landscape (Trainor, 2012). Within this context, social CRM serves both brands and consumers through being responsive to interactive society and fostering consumer involvement via online interaction (Faase et al., 2011).

The generic social CRM view contains four focal actors and five key elements as the main building blocks. *Users* share content on the social media platforms hosted by *platform providers*, and both individuals and *companies* have presences on these platforms in which interaction with other companies and *service providers* occurs (Alt and Reinhold, 2016). Concordantly, social CRM's focal actors consist of users, platform providers, companies, and service providers. Additionally, presence in *social media platforms* for communicating with consumers, *analysis* of consumer-oriented activities and social web content for tracking activities, appropriate *interactive techniques* for ensuring proactive communication with consumers, *CRM integration* with internal systems for providing functionalities, and finally, *management and integration* of social

media and internal activities concurrently for moderating multiple presences are vital for conducting social CRM activities successfully (Alt and Reinhold, 2016). Therefore, social media, analysis techniques, interactive techniques, CRM integration, and management of multiple presences are considered as crucial elements of social CRM. Figure 5 designates the mentioned actors and elements of social CRM.

Figure 5: Social CRM Process

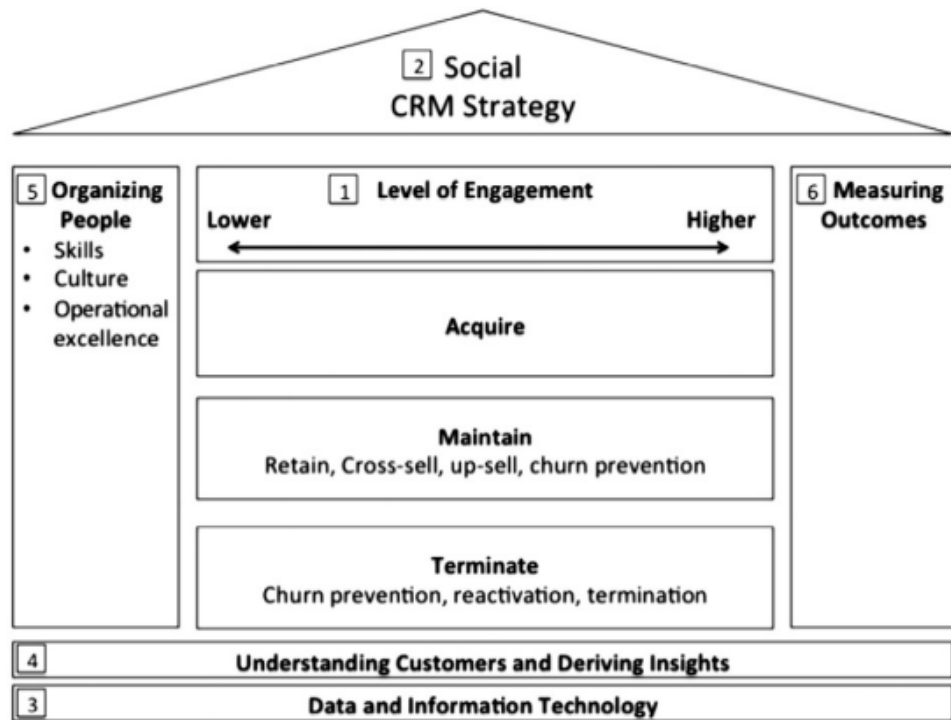


Source: Alt and Reinhold (2016)

Another view on the constituents and dimensions of the social CRM concept was proposed by Malthouse et al. (2013). According to the “Social CRM House” of Malthouse et al. (2013), the social CRM concept contains *CRM* and *social media* as the two main dimensions, and the CRM dimension is composed of three constituents of the traditional CRM framework: relationship acquisition, retention, and termination. This view postulates data and information technology as the foundation and organization of employees and measurement of outcomes as the pillars' social CRM house. Within this context, social media is viewed as the dimension that impacts the core areas (acquisition, retention, termination) and supporting business areas through the level of social media engagement. Brand’s social CRM strategy in terms of value creation and

content marketing is considered the house's roof. The social CRM house is shown in Figure 6.

Figure 6: The Social CRM House



Source: Malthouse et al. (2013)

There exist divergent views on the elements of and requirements for an outstanding social CRM strategy. The 4Cs model presented by Kietzmann et al. (2011) postulates that for achieving a successful social CRM strategy, brands should (re)*cognize* and comprehend the social media landscape and its elements, ensure that social media strategies are *congruent* with company goals, *curate* social media content, and interactions, and *chase* for information varying from consumer opinions to competitors' strategies. Acker et al. (2011), on the other hand, suggests the "MASTER" approach for building efficient social CRM capabilities. According to this approach,

“Monitoring” social consumer communities, *“Assessing and analyzing”* the results of the monitoring process, *“Strategizing and structuring”* social CRM proposition, *“Testing”* the strategy, *“Embedding”* the proposition and strategy into social media activities and campaigns, and *“Reviewing”* embedded activities frequently are vital for a successful social CRM strategy.

2.6. DIFFERENCES BETWEEN TRADITIONAL CRM AND SOCIAL CRM

In today’s fast-moving online world, consumer expectations are higher when compared to the offline world, immediate responses play a vital role in ensuring satisfaction, and satisfying customers is not sufficient to prevent churn (Grover, 2011). To meet the requirements of the new age and consumer expectations, moving the traditional CRM approach one step further is vital, and social CRM meets them by enabling companies to develop bilateral relationships by facilitating dialogue and aiding consumers to create content (Shimp, 2009). Hence, by definition, traditional CRM and social CRM concepts differ both in their nature and application.

Social CRM extends the traditional approach by enhancing social functions, processes, and capabilities that target consumer-brand interaction and the interaction among consumers (Greenberg, 2010). The main difference is that traditional CRM assumes consumers as passive participants solely responding to company actions through buying behavior, but social CRM considers them active and enables them to engage with the company (Malthouse et al., 2013). Greenberg (2009) lists further differences between these two concepts as: (1) Traditional CRM is tactical and operational, whereas social CRM is strategic; (2) Traditional CRM considers company-consumer interaction as a company-dominant relationship, but social CRM views it as a collaborative effort; (3) Traditional CRM only focuses on company-consumer relationship, social CRM also includes among company, business partners, and between-consumer interactions; (4) Traditional CRM aims shaping consumer opinions about a product or service, whereas, starting from the product/service development phase, social CRM views consumers as a partner; (5) In traditional CRM, the focus is on products and services that satisfy consumer, but in social CRM focus is on environment and experiences that yield consumer engagement; (6) Traditional CRM

obtains consumer insights through the single view of consumer data that constitutes insight by integrating consumers' data across all channels of the company, whereas in social CRM insights are obtained by integrating consumer data, consumers' personal profiles on social media, and consumers' participation during this acquisition; (7) In traditional CRM, marketing is focused on processes that deliver targeted and highly specific messages to consumers, whereas in social CRM marketing activities mainly aim building relationships with consumers; (8) Traditional CRM can be described as utilitarian, functional, and operational, whereas social CRM also emphasizes design and style, (9) Traditional CRM is bi-directional, social CRM is uni-directional.

Another notable comparison between traditional and social CRM was put forth by Trainor (2012). According to this comparison, main differences between these concepts can be listed as: (1) Traditional CRM considers customer relationship as a one-to-one relationship, whereas social CRM views it as a network consisting of many-to-many relationships; (2) In traditional CRM customer interaction is one-way, but social CRM holds an interactive perspective; (3) Traditional CRM employs technology for increasing employee efficiency, and effectiveness, whereas social CRM uses it for supporting the connection between company and its environment; (4) In traditional CRM organizational learning is obtained through customer information that is used as an input to decision support systems, but social CRM considers company-brand interaction as the locus of learning; (5) Traditional CRM holds that targeting right consumers is vital for delivering value, whereas social CRM advocates dual creation of value; (6) In traditional CRM information is delivered from consumer to firm, but in social CRM company and consumers co-create information; (7) Traditional CRM captures consumer information and shares it across the company, whereas in social CRM consumers are actively take place in information updates. Distinctions indicated by Trainor (2012) are summarized and further categorized in Table 4.

Table 4: Differences between Traditional CRM and Social CRM

	Traditional CRM	Social CRM
Customer Relationship	The one-to-one relationship between customer and organization.	A network of many-to-many relationships between (1) customers and organization and (2) customers and peers.
Customer Communication	One-way monologue.	Interactive dialog.
Technology Use	The method often deployed to increase employee efficiency and effectiveness; internally focused and usually marketing/sales oriented.	The method deployed to support outside-in processes that connect the firm to the external environment; customer-oriented
Organizational Learning	Customer information is viewed as input to decision support systems and internal responses.	Knowledge co-created with customers; locus of innovation/learning is the interaction between customer and firm
Value Creation	Company-centric focus: targeting and managing the “right” customers and defining value to be delivered.	Emphasis on dual creation of value: customer co-created experiences, collaborative service, and support, innovation co-creation.
Data Sources	The information is delivered from customer to firm.	Information created through interactions with and between customers
Information Process	Customer information is captured and made accessible for use across the organization.	Customer information from legacy systems and customer-centric social applications integrated to create a complete view of customers and networks of customers. Customers are active participants in information updates via interaction-enabling technology.

Source: Trainor (2012)

Social media platforms enable companies to engage with customers online and access previously unavailable customer insights that lead to improving the customer-brand relationship and CRM (Greenberg, 2010). Within this context, social media acts as a bridge between traditional CRM and social CRM. Therefore, despite the several distances between them, traditional and CRM concepts should be considered as complementary.

2.7. THE IMPORTANCE OF SOCIAL CRM FOR COMPANIES

Web 2.0 technologies and the social customer phenomenon force brands to extend the traditional CRM approach and move beyond optimizing consumer-company relationship by including the instantaneous interactions among consumers in the equation. Notwithstanding, companies' behavioral intention to adopt social CRM activities differs in accordance with the perceived usefulness (Askool and Nakata, 2011; Harrigan et al., 2015), perceived ease of use (Askool and Nakata, 2011; Harrigan et al., 2015), perceived reliability (Askool and Nakata, 2011), social impact (Venkatesh et al., 2003; 2012), hedonic motivation (Venkatesh et al., 2012), habit (Venkatesh et al., 2012), and facilitating conditions (Venkatesh et al., 2003; 2012)

Social CRM solutions contribute to main business functions consisting of market research, marketing, sales, support, and product development through its capabilities that are supported by web 2.0 technologies. According to Bernoff and Li (2008), listening and monitoring consumers' interactions with each other facilitates market research. Moreover, the same research suggests that participating in those conversations helps marketing activities, enabling "fanboys" to market your products or services contributes to sales, making consumers support each other facilitates support activities and enabling consumers to work collaboratively to improve extant products or services assists product development. These capabilities constitute social CRM objectives, which can be listed respectively as; listening, talking, energizing, supporting, and embracing.

Grover (2011) listed the benefits of proper social CRM implementations as amplified customer loyalty, more effective marketing, enhanced customer service, and cost reduction. A useful social CRM application enhances consumer-brand interaction

and relationships, which, in turn, amplifies customer loyalty and enhances customer service. Moreover, social CRM provides companies a detailed portrait of customer profiles that can predict purchase preferences, attract customer attention, and create more focused marketing campaigns. Finally, automated data mining and analysis allow companies to set accurate goals, enhance forecasting, measure return on investments more accurately, and acknowledge underperforming strategies and products, hence decreasing the company's costs.

Since the ultimate objective of the social CRM approach is to build long-term relationships with consumers based on high customer engagement (Greenberg, 2010), high engagement-based proceeds can be considered outcomes of practical social CRM implementations. In this context, marketing activities supported with proper social CRM implementation are likely to reach both extant and potential consumers concurrently, and therefore, are capable of enhancing consumer acquisition and retention (Malthouse et al., 2013). As higher customer engagement yields more consumer financial value, customer engagement has a substantial impact on financial performance and is valid for all phases of the previously mentioned customer life cycle, including acquisition, retention, and value development (Woodcock et al., 2011). Another critical point is that a higher level of engagement makes termination of the relationship more challenging to the customers (Malthouse et al., 2013).

One of the most remarkable outcomes of implementing effective social CRM is obtaining precise and meaningful customer insights. These insights include information regarding the nature of conversations about the brand, consumer debate and its environment about the brand, emotional engagement of consumers, and related news about the company (Shokohyar et al., 2017). Moreover, consumer insights facilitate customer-centric innovation, acknowledge the company, suppliers, and intermediaries about customer behavior, attitude, and mood (Woodcock et al., 2011). Consumer insights include data on consumer behavior, consumers' sentiment towards the brand, profile information, and experience maps (Greenberg, 2010), and these products -, brand- and/or firm-related insights reveal consumer needs (Alt and Reinhold, 2012). In turn, these needs serve companies in enhancing the relationship with consumers and increasing customer lifetime value (Kumar et al., 2010).

Another vital advantage of implementing an effective social CRM strategy arises from integrating consumers into operational processes. Adopting a service-oriented social CRM approach reduces companies' service costs (Chesbrough, 2007), enables responding to consumers' service requests promptly, and gives rise to consumer-to-consumer support (Breidbach et al., 2014). Thereby, brands and customers benefit from the social CRM approach by influencing product developments and enhancements, obtaining an accelerated response to support requests, and interacting with other consumers.



CHAPTER THREE

MACHINE LEARNING CONCEPT AND APPLICATIONS IN MARKETING

3.1. DEFINITION OF MACHINE LEARNING

Arthur Samuel (1959), the eponym of “machine learning,” defines the technique as providing computers the capability of learning without being overly programmed. Machine Learning (ML) refers to a computational process in which the human learning process is imitated to obtain knowledge from the real world and improve performance at given tasks based on the previously obtained knowledge. Human beings learn from their experiences as a consequence of their ability to reason. However, computers learn through algorithms that automatically adjust their structure through repetition (experience) to improve their ability to achieve the desired task (El Naqa and Murphy, 2015). In which various sets of input data are provided with desired outcomes, this process is called training. Subsequently, the ML algorithm aligns itself optimally to obtain the anticipated outcome from new and unseen data. Michalski et al. (1985) summarize this task as: “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.” All in all, ML involves providing computers the capability to optimize given parameters using either example data or prior experience (Alpaydm, 2004).

A machine learning model may be predictive to make forecasts or descriptive to gain knowledge or both. However, regardless of the aim, a basic model contains two main components: a learning element that obtains data from the environment and a performance element that uses the obtained information for making decisions (Yao and Liu, 2013). The role of the learning element is vital for two reasons: (1) An efficient algorithm is needed for both solving the optimized problem and storing and processing the large set of data in training; (2) efficiency of the learning is often as important as the model’s predictive accuracy (Zhang, 2010). Hence, it may be stated that the fundamental scientific objective of an ML model is to explore possible learning mechanisms by using statistical theories for model building.

3.2. DEVELOPMENT OF MACHINE LEARNING TECHNIQUES

ML, in the current sense, originates from Donald Hebb's book titled "The Organization of Behavior," which was published in 1949. Hebbian learning rule revealed a model that mimics the relationship and communication between neurons, which has later been widely used in the artificial neural network domain (Hebb, 1949). Following Hebb, in 1952, Arthur Samuel developed a program for playing checkers that recorded and remembered all positions that have been played and employed this information to calculate chances of each side winning (Nilsson, 2009). The motivation behind the creation of "machine learning" was to eliminate the need for intense effort to program computers to learn from experience (Samuel, 1959). Samuel's attempt was followed by Rosenblatt's perceptron machine, which combined Hebb's model with Samuel's ML concept. Perceptron was built for image recognition; however, it failed to recognize many visual patterns, such as faces (Nilsson, 2009). In 1974, a new impulse was given to the domain by Werbos' discovery of multilayers, which employed multiple layers to boost processing power.

Further improvements continued with the development of decision trees by Quinlan (1986) and support vector machines by Cortes and Vapnik (1995). In 1990, Schapire brought a boosting concept to the scene by revealing the idea of bringing weak classifiers for creating a single strong one. This was followed by the development of Long Short-Term Memory (LSTM) by Schmidhuber and Hochreiter (1997), widely used in speech recognition training. When the development sequence is considered, it can be stated that although the ML concept dates back to the 1950s, it has gained remarkable momentum in the 1990s.

Considering the historical development of the ML and the research fields of key contributors, it may be noted that the concept was motivated by the research in diverse disciplines (Yao and Liu, 2013). Nilsson (1982) lists key disciplines that have contributed to the ML domain as:

- **Statistics:** Statistical methods employed for solving a problem regarding value estimation of an unknown function at a point given the values of a function at a set of sample points may be considered as examples of ML.
- **Brain Models:** Non-linear elements with weighted inputs are originated from basic models of brain neurons, and many noteworthy ML techniques employ such non-linear elements that often-called neural networks.
- **Adaptive Control Theory:** This theory deals with controlling a process that includes unknown parameters to be estimated during operation. During operation, these parameters alter frequently, and the control process should track these alterations.
- **Psychological Models:** Psychologists conducted numerous researches on the performance of humans in learning tasks. Decision trees, semantic networks, and goal-seeking behavior are some of these psychology-based vital concepts that inspired ML development.
- **Artificial Intelligence:** AI researchers discovered the role of similarities in learning and shaping future actions based on previously seen cases. The Discovery of rules for expert systems using decision tree methods, inductive logic programming, and explanation-based learning are examples of AI contributions to the domain.
- **Evolutionary Models:** Inspired by nature and biological evolution, genetic algorithms and genetic programming are developed to enhance computer programs' performance.

Coherently, these algorithms have been employed in various settings, including pattern recognition, spacecraft engineering, business, entertainment, computer vision, computational biology, and medical applications (El Naqa and Murphy, 2015). Zhang (2010) relates this widespread application and multi-faceted nature of ML applications to its abilities of;

- Producing precise outputs for a large input sample, appropriately constraining input and output function for approximating the association inherited in the examples.
- Extracting relations that are unrevealed among large data samples (data mining).
- Being used for the job enhancement of present machine designs.
- Being able to adapt to an altering environment without a need for constant redesign.

3.3. TYPES OF MACHINE LEARNING

The term “machine learning” is being used to connote general techniques to deduce patterns from large data sets or to the ability to make predictions based on what is learned from analyzing previous data (Zocca et al., 2017). Hence, this broad definition comprises various techniques. These techniques may be classified based on the employed learning process as supervised, unsupervised, semi-supervised, and reinforcement learning (Portugal et al., 2018).

3.3.1. Supervised Learning

The first learning type is supervised learning, in which algorithms require external assistance, and data is a sample of input and output patterns. Learning a function is enabled by training data with examples, including the inputs and the desired outputs (Mitchell, 2006). The input dataset is divided into two: (1) *train dataset*, which contains output variable to be predicted or classified, (2) *test dataset* for prediction or classification (Dey, 2016). Concordantly, each training example of input data is paired with a classification label for the learner to absorb similarities and differences (El Naqa and Murphy, 2015). In other words, the process may be described as using experience for gaining experience, in which the “experience” (training example) contains essential information that is absent in the unobserved test examples to which the learned expertise will be applied (Shalev-Schwartz and Ben-David, 2014).

According to the type of outputs, supervised learning may be divided into *classification learning* and *regression*. In *classification* learning, the task is assigning new inputs into one of the discrete categories, whereas in *regression*, output space consists of continuous variables (Camastra and Vinciarelli, 2015). To put it another way, the output of the function can either be a continuous value (regression) or can predict the class label of the input (classification) (Zhang, 2010).

3.3.2. Unsupervised Learning

The second type is unsupervised learning, in which no specific output value is provided to the algorithm, and the algorithm learns to discover the natural statistical uniformities and internal representations for encoding features to generate classes automatically (Yao and Lin, 2013). Unlike supervised learning, there exists no difference between training and test data in unsupervised learning, and the learner processes input data to reach a sort of summary or compressed version of the provided data (Shalev-Schwartz and Ben-David, 2014). Unsupervised learning treats input data as a set of random variables (Zhang, 2010) and deals with learning patterns in the input when a specific output value is not supplied (Russell, 2013). Hence, the aim is to describe how the data is clustered (Zhang, 2010).

In unsupervised learning, the aim is extracting a structure from the input data by finding a short description or summary of the data through specifying similarities between pairs of objects (Camastra and Vinciarelli, 2015). This idea underlies *clustering* algorithms, the objective of which is grouping similar objects together under the same cluster.

3.3.3. Semi-Supervised Learning

The third type of machine learning is semi-supervised learning, in which some part of the data is labeled, and others are unlabeled, and the labeled part is employed for the learning of the unlabeled part (El Naqa and Murphy, 2015). The objective of semi-supervised learning is to comprehend the impact of combining labeled data with unlabeled on the learning behavior and using readily available unlabeled data for

advancing supervised learning when the labeled data are scarce or employing labeled data expensive (Zhu and Goldberg, 2009). Since the algorithm is supplied with supervision information in addition to the unlabeled data, semi-supervised learning may be considered as a midway between supervised and unsupervised learning (Chapelle et al., 2006).

In semi-supervised learning, there exists two learning paradigms; transductive learning and inductive learning. Transductive learning aims to make predictions on the unlabeled training data without generalizing output for future test data (Zhu, 2011). On the other hand, the object in inductive learning is learning a parameterized classifier that is generalizable to unseen cases (Belkin et al., 2006).

3.3.4. Reinforcement Learning

The last type of learning is reinforcement learning, in which learning actions affect future inputs, and the learner is not stated which actions to take but is required to discover which actions produce the most fruitful reward by trying them out (Sutton and Barto, 1998). In other words, reinforcement learning can be described as learning to control a system in a way to maximize a numerical value that represents a long-term objective (Littman and Szepesvari, 1996).

The most critical challenge in reinforcement learning is that for obtaining a substantial reward, the learning algorithm should select actions that have been tried formerly and found to be useful, but for revealing such actions, the algorithm needs to try actions that have not been selected before. Since the algorithm is required to exploit what it already knows for obtaining a reward, but it also should make exploration for making more fruitful action selections in the future, the algorithm must try a bunch of actions and gradually favor the ones that yield better results (Sutton and Barto, 1998).

3.4. MACHINE LEARNING AND TEXT MINING

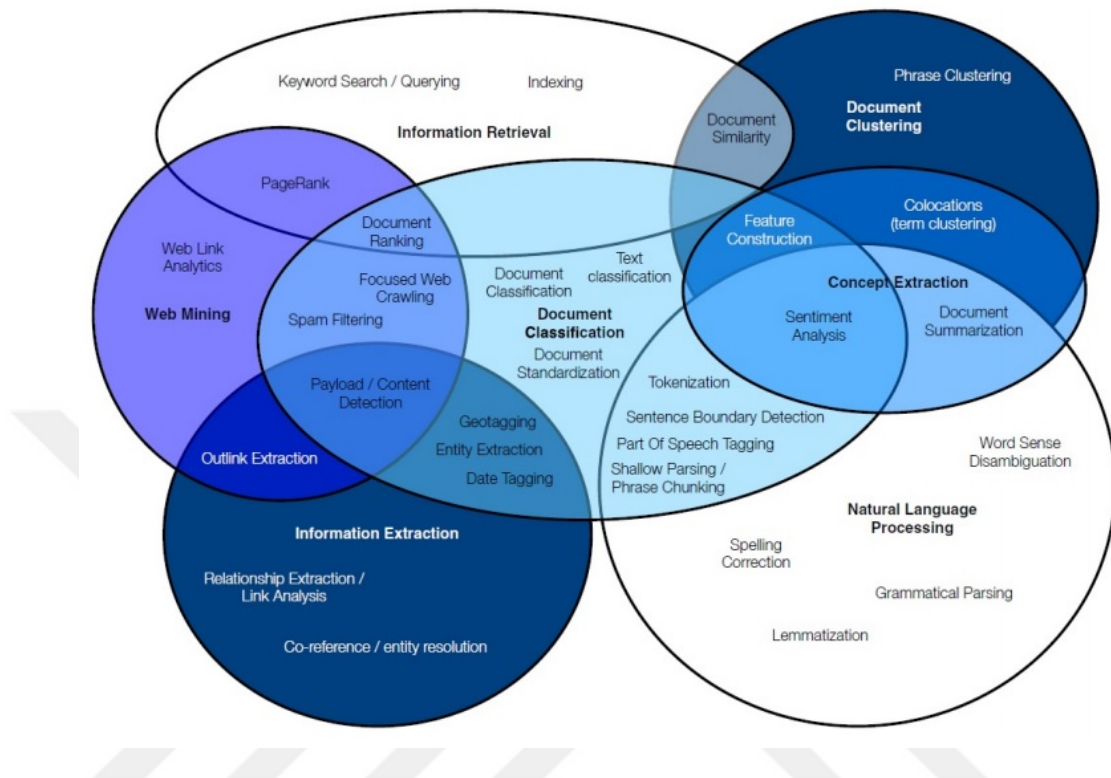
Extracting non-trivial patterns and information from a textual data source through various statistical algorithms is referred to as *text mining*, *text analytics*, or

machine learning from text (Aggarwal, 2018). Given the definition, text mining may seem to be quite similar to data mining; however, there exist vital differences between these concepts to be highlighted. Data mining tools extract information that is structured, previously unknown, and hidden in the input set (Witten and Frank, 2000). On the other hand, text mining tools extract information that is unstructured or semi-structured (i.e., emails, blogs, articles, HTML files) and explicit (Moreno and Redondo, 2016). Based on its properties, text mining can be viewed as a focused side-shoot of data mining and machine learning (Aggarwal, 2018).

The steps followed in a typical text mining application is summarized by Moreno and Redondo (2016) as follows: Before the analysis, a text-mining tool retrieves collected documents and checks the format and character sets. Then the text goes through the text analysis phase, each component of which performs a specific process on the text. This process's output, consisting of various phrases, provides structured or semi-structured information to be further analyzed. Considering the intersection point of machine learning and text mining approaches, the process may be summarized as follows (Miner et al., 2012): A set of input features should be defined, and documents should be classified before applying machine learning algorithms. After the model is trained, unclassified documents may be submitted to the classification model, in which patterns of the categories in the training section is employed.

Various text mining techniques are employed to investigate patterns in a given text and their mining phase (Gupta and Lehal, 2009). The following figure by Weiss et al. (2010) depicts the relationships between and intersections of diverse text mining techniques. Information extraction, information retrieval, clustering, information extraction, and text summarization are explained in the following section.

Figure 7: Relationships Among Text Mining Techniques



Source: Weiss et al. (2010)

3.4.1. Information Extraction

Information extraction (IE) refers to extracting meaningful information from a crowded text (Talib et al., 2016). IE systems enable the detection of critical phrases and relationships, which plays a vital role in converting unstructured text into a structured representation, and hence, it should be considered a pre-processing step in text mining (Aggarwal, 2018). One of the most popular forms of IE is named entity recognition (NER), which discovers different types of entities from text such as names, locations, organizations, temporal expressions, quantities, and monetary values, and then establishes relationships among those entities (Aggarwal, 2018; Moreno and Redondo, 2016; Talib et al., 2016). Khan et al. (2010) list the steps taken for information extraction as; tokenization (partitioning the text into tokens), removing stop words (elimination of stop words such as “the,” “and,” “a/an” etc.), and stemming word (converting different

word forms into a uniformed style). IE results in a corpus stored in a database for further processing, and precision and recall results of the extracted data are employed to check the relevance of the results (Talib et al., 2016).

3.4.2. Information Retrieval

Information Retrieval (IR) refers to the extraction of relevant patterns related to a given set of words (Talib et al., 2016). In other words, IR systems find related documents that provide answers to given questions (Khan et al., 2010). Search engines use IR systems to extract relevant data on the Web, employ query-based algorithms for tracking trends to obtain more accurate results, and provide users with the more appropriate information according to their needs (Talib et al., 2016). To achieve this, statistical measures are employed to automatically process text data and compare it to the given questions (Khan et al., 2010).

3.4.3. Clustering

Clustering refers to the process of grouping related texts by creating a vector of topics for each document and assigning the text to a particular topic cluster (Moreno and Redondo, 2016). Since explicit examples of desired categories of documents are not provided previously, clusters are not known a priori, and hence, clustering is an unsupervised process (Aggarwal, 2018). Clustering differs from IR by being unsupervised. IR systems imply supervised learning since the previous knowledge is used for assigning a given text to a given category, whereas clustering implies unsupervised learning due to a lack of previously defined categories (Moreno and Redondo, 2016). Some of the widely used clustering techniques are hierarchical, density, distribution, centroid, and *K*-means (Talib et al., 2016).

3.4.4. Summarization

Summarization can be described as the process of creating a summary of a text for reckoning, whether it meets the needs of the user and worth further reading (Moreno

and Redondo, 2016). Key application areas of summarization are; news articles, search engine results, review summarization, emails, scientific articles, and improving other automated tasks (Aggarwal, 2018). Summarization methods can be divided into two groups; (1) *shallow analysis*, which is constrained by the syntactic level of representation and used for extracting essential parts in a document, (2) *deeper analysis*, which adopts semantics level of representation of the document (Moreno and Redondo, 2016). It should also be noted that the summarization process of single-document and multi-document are dissimilar since the summarization of multi-document ascends in a setting where documents are closely related, and the process needs to be more inclusive of the specific needs of the user (Aggarwal, 2018).

3.5. MACHINE LEARNING AND TEXT MINING ALGORITHMS

As indicated previously, text mining is a field that is drawing upon knowledge and methods from several external disciplines such as artificial intelligence, statistics, machine learning, and computer science. Hence, various machine learning algorithms are employed in text mining for classification, clustering, and information extraction. The earliest use of machine learning algorithms for text mining purposes was recorded by Li et al. (1991) for neural networks, Apte et al. (1994) for decision trees, and Cortes and Vapnik (1995) for Support Vector Machine (Miner et al., 2012).

Table 5: Common Machine Learning and Text Mining Algorithms

<i>Algorithm</i>	<i>Type of Learning</i>	<i>Common Practice Area in Text Mining</i>
<i>Support Vector Machine (SVM)</i>	Supervised	Document classification
<i>Discriminant Analysis</i>	Supervised	Document classification
<i>Naïve Bayes</i>	Supervised	Document classification
<i>Nearest Neighbor</i>	Supervised	Document classification

<i>Linear Regression</i>	Supervised	Document classification
<i>Ensemble models</i>	Supervised	Document classification
<i>Decision Trees</i>	Supervised	Document classification
<i>Neural Networks</i>	Supervised, Unsupervised	Document classification
<i>K-Means</i>	Unsupervised	Clustering
<i>Fuzzy C-Means</i>	Unsupervised	Clustering
<i>Hierarchical Models</i>	Unsupervised	Clustering
<i>Hidden Markov Models</i>	Unsupervised	Information extraction

Source: Developed by the author

3.5.1. Support Vector Machine (SVM)

Support Vector Machine (SVM), which is a blend of linear and instance-based learning, performs classification by creating a hyperplane that splits classes in a way to maximize the distance between them (Miner et al., 2012). Data points called “*support vectors*” are selected from each class, and for eliminating the risk of misclassification, a linear discriminant function that separates them as far and wide as possible is built (Witten and Frank, 2005). In other words, support vectors are the training examples nearest to the maximum margin hyperplane (the decision boundary), and there exists a minimum of one support vector for each class (Flach, 2012). The set of support vectors outlines the maximum margin hyperplane, and all remaining training examples become irrelevant-thus may be removed without shifting the hyperplane (Witten and Frank, 2005). The main idea in SVM is using data as an input and producing hyperplane as an output that ensures points lie on the precise class’ side (Aggarwal, 2018).

3.5.2. Linear Discriminant Analysis

Based on the linear discriminant of Fisher, Linear Discriminant Analysis (LDA) searches for the vectors that *best discriminate* among classes-instead of those describing the data (Martinez and Kak, 2001). LDA reflects high-dimensional data on a low-dimensional space, where the data is reformed for enhancing class separability (Ye et al., 2006). LDA contains three main matrices; within-class, between-class, and total scatter matrix. Within this context, LDA aims to maximize the between-class measures while minimizing within-class ones (Martinez and Kak, 2001). On the other hand, the total scatter matrix is multiple of the sample covariance matrix and is non-singular (Ye et al., 2006). This non-singularity is called as “*singularity problem*” and considered a vital drawback of LDA since in the majority of applications (i.e., information retrieval, text mining, and face recognition), scatter matrices are all singular due to the high dimensionality of data, and the dimension exceeds the number of data points (Ye et al., 2004).

3.5.3. Naïve Bayes

Bayes' theorem states that the occurrence probability of an event can be calculated by multiplying the intrinsic probability (calculated from the past data) with the probability of future occurrence (calculated from the present data) (Miner et al., 2012). By drawing upon this theorem, Naïve Bayes classifiers hold that each class's fraction in a data set includes the posterior probability that any new example will belong to either one class or another (Miner et al., 2012). Naïve Bayes classifier assumes independency among input data while ignoring possible dependencies, or in other words, correlations, and reduces multivariate problems to a group of a univariate problem (Alpaydin, 2004). Despite its naïve design, these classifiers are highly effective and accurate across various domains (Kelleher et al., 2015), mainly when the number of variables is high (Witten and Frank, 2005). Another reason behind the high accuracy is the fact that errors in the calculation of the prior probabilities do not essentially result in prediction errors (Kelleher et al., 2015).

3.5.4. Nearest Neighbor

Nearest neighbor classifiers, also known as *instance-based learners*, *memory-based learners*, or *lazy learners*, use the existing data to predict the class of new observations by finding the most similar observation among exemplars (Miner et al., 2012). These classifiers operate by memorizing the training set and leaving the classification to the end without training (Aggarwal, 2018). The idea is predicting the class of new examples based on the class of its closest neighbor in the training set (Shalev-Shwartz and Ben-David, 2014). In other words, existing data is employed to predict new observations by finding the most similar observation in the training set and predicting that the new one will have the same outcome (Miner et al., 2012). The nearest-neighbor classifiers contain low bias but high variance, indicating the risk of overfitting when the training set is noisy, limited, or unrepresentative (Flach, 2012).

3.5.5. Linear Regression

Linear regression algorithm provides an adequate description regarding the impact of an input on a given output (McClendon and Meghanathan, 2015). It is a widely employed tool for revealing the relationship between explanatory variables and numeric, continuous outcomes (Shalev-Schwartz and Ben-David, 2014). The output is calculated by adding up the inputs multiplied by some constants, or in other words, weights (Harrington, 2012). It should also be noted that linear regression may be employed for classification problems when numeric attributes exist. In such cases, regression is performed for each class, and output is set to be one of the training examples belongs to the class, and to be zero if not (Witten and Frank, 2005).

3.5.6. Ensemble Models

Ensemble models combine multiple classifiers' prediction for optimizing performance and overcoming drawbacks of single classifiers (Harrington, 2012). These models employ a base method as input and advance it by either applying it to various modifications of the data or various variants of the same model (Aggarwal, 2018). Each model with diverse strengths and weaknesses provides a different mathematical

perspective on the modeled relationship (Miner et al., 2012), and hence, contributes to the performance of prediction. Bagging and boosting are the most widely employed ensemble methods. In bagging, several datasets in the same size as the original set are built through a random selection from the problem domain (Harrington, 2012), and the employed models receive equal weights (Witten and Frank, 2005). Boosting, on the other hand, refers to the application of diverse classifiers consecutively to a dataset (Harrington, 2012), and more weight is given to the more successful models (Witten and Frank, 2005).

3.5.7. Decision Trees

Decision trees operate by dividing the training dataset hierarchically by imposing conditions regarding attributes, and each class's documents are placed in a single node (Aggarwal, 2018). Decision tree predictors forecast the label of an instance through moving on the path from the root node to a leaf, and at each node, the descendant is chosen by splitting the input space (Shalev-Schwartz and Ben-David, 2014). In univariate splits, the condition is imposed over a single attribute, whereas in multivariate split, it is imposed over several attributes (Aggarwal, 2018). The process, including testing the value of a descriptive feature and descending a level in the tree, is repeated until the process comes to a leaf node at which a prediction can be made (Kelleher et al., 2015). It should also be noted that numerous decision trees may be combined for creating “*random forests*” (Aggarwal, 2018).

3.5.8. Neural Networks

Artificial neural networks (ANN) simulate human neural networks, consisting of many neurons that are interconnected at contact points called synapses (Shalev-Schwartz and Ben-David, 2014). In biological networks, learning is performed as the individual nodes (neurons) obtain input from other neurons, make computations on these, and feed them into other neurons, and the computation is affected by the weights provided by the training data (Aggarwal, 2018). Neural networks may either be supervised or unsupervised. Supervised ANNs may be employed for both prediction

and classification. Supervised ANNs operate by approximating known values in the data set and comparing outputs with the known values. The error is used for modifying the link weights for the next iteration, and processing endures until the minimum error value is achieved (Miner et al., 2012). Unsupervised ANNs, on the other hand, are used for classifying documents through self-organizing maps (Miner et al., 2012). All in all, the main idea of ANNs is incrementally modifying the weights whenever wrong predictions are made until the best accuracy is achieved (Aggarwal, 2018).

3.5.9. K-Means Clustering

Lloyd's, or *K-means*, the algorithm operates by forming a particular number (K) of clusters that minimize distances between observations within a cluster while maximizing it between observations in different clusters (Miner et al., 2012). The algorithm chooses K random points as cluster centers, assign observations to the closest cluster, and calculate the mean of the observations in each cluster (Witten and Frank, 2015). Several iterations are being run during K -means clustering, and iterations continue until the cluster centers are stabilized (Aggarwal, 2018). Even though K -means clustering is being preferred due to its simplicity, it has two vital drawbacks that may be overcome via statistical mixtures: (1) Assuming that a point may only hold characteristics of a single cluster may be biased; (2) Determining the optimal number of clusters beforehand may be problematic if the aim is just to cluster (Rogers and Girolami, 2016).

3.5.10. Fuzzy C-Means

Fuzzy C-means (FCM) algorithm brings a novel solution to a previously indicated clustering problem, assigning a given data point to a cluster and defining its characteristics depending on a single cluster. The fuzzy C-means algorithm identifies a data point's shared similarity with each cluster by producing fuzzy c-partitions of a data set and describes the membership of each data point in all clusters by a membership function ranging between 0 and 1 (Bezdek et al., 1984). Hence, the fuzzy C-means algorithm is suitable for the datasets, including data points belonging to more than one

cluster. The algorithm first chooses c random cluster centers from the data set, and then, after several iterations, finds the genuine cluster center (Hung and Yang, 2001). Hence, it is vital to choose initial centers for finding the actual cluster centers at least possible iterations.

3.5.11. Hidden Markov Models

Hidden Markov models (HMMs) operate by assuming that the data is originally containing label-word pairs, but these labels are hidden, and thus, should be revealed through information extraction (Aggarwal, 2018). Hence, it may be stated that HMMs are stochastic processes with a “hidden” underlying stochastic process that can only be revealed through other iterative stochastic processes that generate the sequence of observed symbols (Rabiner and Juang, 1986). HMMs strive to model the most likely sequence of terms by maximizing the joint probability of terms that specify the frequency of several events occurring together (Miner et al., 2012). HMMs treat data as sequential entities rather than as independent ones by treating a data point, for instance, a sentence belonging to a summary, as a sequence of its preceding sentence included in the summary (Aggarwal, 2018). HMMs are employed for solving one of the following problems; (1) *the scoring problem* referring to the probability of HMM to produce the sequence), (2) *the alignment problem* referring to the optimal sequence that HMM can use for producing the sequence, (3) *the training problem* referring to the search for the best fitting structure and parameters of HMM (Eddy, 1996).

3.6. LITERATURE REVIEW ON MACHINE LEARNING AND TEXT MINING APPLICATIONS IN SOCIAL MEDIA ANALYTICS

The notable upsurge in the daily amount of social media utilization has driven social networks to become a vital part of consumers’ lives (Shareef et al., 2018) and re-shaped information sharing and communication tendencies (Kapoor et al., 2018). As an outcome of this broad utilization, the ascending number of active social media users, and the fertility of the domain, social media communication attracted the notice of both researchers and professionals within the field (Shareef et al., 2018). Therefore, both

commercial and non-profit organizations endeavor to comprehend trending issues, patterns, and threats in those broadly utilized social media platforms for inferring some insights (Stieglitz et al., 2018). From a marketing viewpoint, social media platforms driven by user-generated content offer prolific research paths for various settings such as consumer and seller behavior, entrepreneurship, political marketing, and customer relationship management (Greenwood and Gopal, 2015).

The remarkable expansion in social media usage resulted in the accrual of enormous data sets in several formats, including text, images, videos, sounds, and geotags, which are called Social Media Big Data (Stieglitz et al., 2018). With their capability to process high-volume and complex data, ML and text mining algorithms have become a prevailing approach in social media analytics and noticed as a promising area. The substantial influence of ML and text mining applications in social media analytics makes the domain more attractive, and researchers from diverse disciplines directed attention to the field.

The extant body of knowledge on text mining and machine learning applications in social media analytics can be grouped according to the source of analyzed user-generated content. This approach reveals five groups of research, namely as; review sites (Yelp, TripAdvisor, IMDb, etc.), microblogs (Twitter, Tumblr, Foursquare), social networking sites (Facebook, Whatsapp, WeChat, etc.), visual content communities (Youtube, Instagram, Flickr, etc.), and cross-media studies (i.e., both Facebook and Twitter data). The characteristics of each group are explained, and related studies are reviewed in the following sections.

3.6.1. Review Sites

The vast majority of the studies that employ machine learning and text analytics algorithms for analyzing social media content derive data from consumer review platforms such as TripAdvisor, Amazon reviews, Yelp, and IMDb. These studies' primary motivation can be stated as investigating the sentiment polarity of reviews for acquiring an in-depth understanding of the helpfulness and sentimental tendency of the reviews from several perspectives. Among review sites, the most examined platform is Amazon, and TripAdvisor and Yelp follow it.

A substantial portion of these studies (Agnihotri and Bhattacharya, 2016; Eslami and Ghasemaghaei, 2018; Heng et al., 2018; Rathan et al., 2018; Costa et al., 2019) investigated sentiment polarity of consumer messages on Amazon reviews by using diverse learning functions and statistical packages. Agnihotri and Bhattacharya (2016) examined sentiment of more than 2.000 Amazon reviews through LIWC (Linguistic Inquiry and Word Count) package to reveal the impact of readability and sentiment polarity on the helpfulness of the review. This study's results highlighted that content and sentiment polarity of the review are critical predictors of helpfulness, and the reviewer experience moderates them. Eslami and Ghasemaghaei (2018) investigated the influence of review polarity and review score inconsistency on product sales by examining over 15.000 Amazon reviews on musical instruments and digital music by employing the Comprehensive R Archive Network (CRAN). This research pointed out that sentiment polarity, review score, and score discrepancy are determinants of sales for high involvement products, whereas sentiment polarity and score are critical for the sales volume of low involvement products. Similarly, Costa et al. (2019) examined the sentiment polarity of 100.000 Amazon reviews by employing VADER (Valence Aware Dictionary and Sentiment Reasoner) library for investigating if a review on published book/electronics was incentivized. The generated decision tree model used various characteristics, including the review length, helpfulness rate, and the sentiment polarity achieved 75% accuracy in biased review prediction. Rathan et al. (2018) employed Support Vector Machine (SVM) for investigating 400.000 phone reviews on Amazon and developed a model that detects review sentiment with 80% accuracy. As an exception in this stream of research, Heng et al. (2018) adopted the topic modeling approach and employed linear discriminant analysis (LDA) for investigating the impact of review helpfulness on consumers' coffee buying decisions. Findings revealed that comments regarding Amazon service experience, subjective expression, flavor feature, and physical feature affect reviews' helpfulness.

As mentioned previously, comments posted on TripAdvisor constitute the second widely examined user-generated review content in the literature. Unlike the Amazon reviews that denote consumers' experience with purchased goods and Amazon's service, this stream of research focuses on evaluating consumed accommodation service. Calheiros et al. (2017) classified online hotel reviews by

employing topic modeling through Linear Discriminant Analysis (LDA) and revealed that hotel food yields common positive sentiments, whereas hospitality yields both ordinary and strong positive feelings. Chun Chang et al. (2017) analyzed ratings and sentiment polarity of over 600.000 reviews and found that business travelers give lower ratings and express more negative emotions than couples, and rates tend to be lowest in July and highest in December. Finally, for developing a precise hotel recommendation scheme, Nilashi et al. (2018) used multi-criteria ratings to investigate the performance of several algorithms in achieving the most precise recommendations. Results revealed that Fuzzy Rule-Based Ensemble and Expectation-Maximization Ensemble yield the most precise recommendations by reaching 94.5% accuracy.

The third stream of research is focused on the examination of consumer reviews posted on Yelp. For revealing online consumer engagement dimensions, Bilro et al. (2018) analyzed over 15.000 Yelp reviews via WordNet2.1 software. Findings of this study pointed out that cognitive processing of engagement and hedonic experience strongly affects consumers' review effort. Nave et al. (2018) investigated the sentiment polarity of 12.000 Yelp reviews via Semantria software and presented a data-driven decision support system (DSS) for managers by employing Correlated Topic Models (CTM). The presented DSS provided managers the opportunity of investigating reviews from diverse perspectives. Lastly, Jimenez-Marquez et al. (2019) provided a data management framework that focuses on the information contained in Yelp reviews on hotel experiences. The results revealed that binary classification yields the most outstanding results with 88% accuracy.

The last stream is focused on deriving insights from the sentiment polarity of film reviews posted on IMDb. Hu et al. (2018) examined whether reviews affect readers' desire to watch evaluated movies by predicting movie box-office performance from reviews. Results showed that reviews impact the box-office performance, and performance would endure satisfactory as the number of reviews endured high, even if the sentiment polarity is negative. Tripathy (2018) classified review sentiments by employing various supervised ML algorithms and revealed that as the "n" in n-gram accrues, the accuracy in sentiment polarity prediction declines.

3.6.2. Microblogs

The second-largest stream of research that employs machine learning and text analytics algorithms for social media analytics consists of studies that derive data from microblogging platforms such as Twitter and Sina Weibo. These studies' primary motivation can be stated as examining the sentiment polarity of microblogging posts, mainly by employing the SVM algorithm for data analysis. Among microblogging sites, the most examined platform is Twitter, and Sina Weibo follows it.

The majority of the studies focus on examining the sentiment polarity of consumer tweets. Hawkins et al. (2016) examined the sentiment polarity of more than 400.000 patient tweets by employing the Python library named "TextBlob" to reveal US hospitals' perceived quality. Surprisingly, the study results revealed that sentiment polarity is not allied with extant quality ratings. Kim et al. (2016) compared customer reviews and sales performance via examining the sentiment polarity of 220.000 tweets via a lexicon-based approach to analyze the sentiment and SVM for determining purchase intention. Findings highlighted that the volume of tweets and purchase intention reveals a significant gap between the market leader and the follower. By examining 42.000 posts of British MPs during the 2015 UK General Election campaign, Walker et al. (2017) anticipated the retweet count of political tweets via examining the sender, sentiment, content, and structure with 76,7% accuracy. D'Avanzo (2017) examined 3.000 tweets for analyzing Twitter users' sentiment on the issues that take place in Google trends, and by employing various sentiment analysis algorithms (i.e., naïve Bayes detection algorithm, a simple voter algorithm, an NLTK-based sentiment analysis algorithm, commercial sentiment analysis tool) credibility of the proposed approach is confirmed. Rogers et al. (2017) investigated the sentiment polarity of 350.000 tweets via 'Sentiment' software package using Naive Bayes and exposed that more famous brands suffer more from amplified negativity in tweets. Adamopoulos et al. (2018) examined the influence of personality trait resemblance on the effectiveness of WOM via LDA and revealed a positive relationship between the level of personality resemblance and the likelihood of a subsequent purchase from a recipient of WOM message. Lastly, by using the Sentistrength package, Park et al. (2018) combined utilization of geocodes, sentiment polarity, hotspot analysis, and GIS mapping in the visualization of emotions associated with on-site customer experiences, used and

identified three hot spots in Disneyland where tweets with positive sentiment were posted.

The remaining studies of this stream either employed more complex machine learning and text analytics methods (i.e., Ghiassi et al., 2013; Ikeda et al., 2013; Pournarakis et al., 2017; Bigne et al., 2019) or derived data from microblogging platforms other than Twitter (i.e., Wang et al., 2019). Ghiassi et al. (2013) analyzed over a million tweets to develop a sentiment classification using SVM. The second step revealed 181 terms for sentiment expressions and six brand-specific expressions by employing the developed Twitter-specific lexicon and DAN2 machine learning approach. Ikeda et al. (2013) analyzed 100.000 tweets intending to propose a hybrid text and community-based technique for assessing demographics and reached an 80% accuracy in estimating demographic facts by examining tweet history and follower-user relationships. In a similar vein, Pournarakis et al. (2017) aimed to develop an algorithm for generating insights on the tweet volume, discussion topics, sentiment towards a brand, and by employing SVM and LDA methods and succeeded to assess an overall average sentiment of a given topic with 83,2% accuracy. Through analyzing the number of retweets and replies, event tweets, tourist attraction tweets, and retweets of a hotel's Twitter account via Artificial Neural Network (ANN), Bigne et al. (2019) estimated occupancy rate with above 90% accuracy. Finally, Wang et al. (2019) analyzed around 700.000 Sina Weibo posts by employing Naïve Bayes to quantify the ideal level of marketing aggressiveness essential for maximum popularity. The results revealed that marketing aggressiveness is capable of predicting popularity with 96% accuracy.

3.6.3. Social Networking Sites

The third stream of research that employs machine learning and text analytics algorithms for social media analytics consists of studies that derive data from social networking sites. Entire studies in this stream preferred Facebook as the data source, and the majority aimed at predicting interaction and engagement rates. In contrast, another sub-group of research concentrated on examining the sentiment polarity of posts shared on Facebook brand pages.

The first sub-group that aims at predicting engagement and interaction rates contains three primary studies. In their research, Kwok and Yu (2013) and Moro et al. (2016) analyzed above 1.000 Facebook posts by employing an SVM classifier. Kwok and Yu (2013) found that photo and status updates yield higher interaction rates while more users are endorsing conversational messages, and Moro et al. (2016) revealed that status updates double up the interaction rate of other sorts of posts, while special offers and competitions are yielding more significant impact. Unlike Kwok and Yu (2013) and Moro et al. (2016), Lee et al. (2018) analyzed above 100.000 posts by employing a blend of numerous classifiers with ensemble learning and revealed that directly informative content yield lower interaction rates but a higher level of engagement might be attained through merging informative content with brand personality-related attributes.

The second sub-group of research concentrated on examining the sentiment polarity of posts shared on Facebook brand pages. İlhan et al. (2018) investigated the consequences of Attack Defense and Across in terms of sentiment polarity and page volume by examining 1.500 Facebook comments via the SVM algorithm. Results denoted that negative sentiment stated by competitor brand fans encourage fans of the other brand to defend the brand and such a situation results in a net increase in social media performance. Dhaoui et al. (2017) compared lexicon-based and machine learning algorithms' performance in detecting sentiment polarity by analyzing 850 consumer comments on 83 different Facebook brand pages. Findings signified that both approaches yield similar accuracy rates but differ significantly in classification ensembles, requiring a combined approach for enhanced performance. Finally, by analyzing Samsung mobile comments, Wong et al. (2017) proposed an ANN-based approach for investigating the influence of comments about product attributes and comment length; it is revealed that fewer information is conveyed through common attributes.

3.6.4. Visual Content Communities

The fourth stream of research employing machine learning and text analytics algorithms for social media analytics consists of studies that derive visual or verbal data

from visual content communities. For determining the attractiveness of touristic sites, Giglio et al. (2019) analyzed 26.000 Flickr photos via *K*-means, ImagelIdentify statistical package, and clustering and mapped the photographic activity touristic in cities. Klostermann et al. (2018) clustered 10.000 Instagram posts according to their sentiment polarity, and through integrating image, tags, and text, researchers allocated these posts according to the content and obtained differentiated views on brand perception. Finally, by building upon Value Theory, Jang et al. (2013) aimed at assessing the causality between Youtube user profiles, value structures, and attitudes. Examination of the sentiment polarity of 2.670 Youtube comments revealed that the constructed framework can extract reviewers' attitudes, and two-dimensional value structure (utility and hedonic) are determinants of attitude.

3.6.5. Cross-Media Studies

The fifth stream of research consists of studies that derive data from various social media platforms to investigate diverse phenomena. In other words, studies in this group utilized cross-media data sourced from multiple types of social media platforms simultaneously. When compared with the previous streams, it can be stated that studies employing data from various platforms are those investigating broader phenomenon such as media communication strategies of election candidates, factors influencing individuals' investment decisions, or consumers' consumption decision-making journey.

Ryoo and Bendle (2017) and Pineiro-Chousa et al. (2017) both conducted researches obtaining data through both a microblog (Twitter) and a social networking site (Facebook). By employing the LDA Topic Model, Ryoo and Bendle (2017) analyzed 85.000 Twitter and Facebook posts of presidential election candidates for revealing the topics that the candidates focus on overtime. On the other hand, by employing Logit and fsQCA, Pineiro-Chousa et al. (2017) analyzed over 90.000 investor posts for examining the influence of investors' Twitter and Facebook activities on the Chicago Board Options Exchange Market Volatility Index (VIX). Results revealed that combinations of diverse social network variables affect investors' decisions, causing a variation in the market risk. Similar to these researches, Vazquez et al. (2014) and Liu et al. (2016) both utilized the Natural Language Processing (NLP) technique for

examining social media data attained from various platforms. For estimating consumers' content preferences, Liu et al. (2016) applied HDLDA to their search queries and managed to estimate CTR in sponsored search advertising. Vazquez et al. (2014) aimed to convey eWOM texts to a single phase of the consumer decision journey and detect marketing mix elements by analyzing around 23.000 posts by employing a decision tree and achieved a 78% accuracy rate in identifying business indicators automatically.

Among the remaining studies of this group, Schniederjans et al. (2013) and Nguyen et al. (2018) employed the SVM method while Pelaez et al. (2019) applying the ISMA-OWA operator for investigating social media data. By analyzing UGC through SVM and LDA, Nguyen et al. (2018) aimed to automatically mine key sentences and posts of a Web document by integrating its social context. Results denoted that social context supports local information in capturing summary indicators in sentences. Schniederjans et al. (2013) examined the extent to which impression management (IM) direct-assertive strategies in social media impact a firm's financial performance. By employing SVM, researchers analyzed 150 publicly traded firms' social media data and found that ingratiation, intimidation, organizational promotion, and supplication strategies could estimate the financial performance with a 75,25% accuracy rate. Finally, Pelaez et al. (2019) established a model for obtaining an alternative ranking that contextualizes opinions articulates on social media platforms by employing 60.000 social media comments through the ISMA-OWA method.

CHAPTER FOUR

RESEARCH FRAMEWORK, METHODOLOGY, AND FINDINGS

4.1. RESEARCH FRAMEWORK

4.1.1. Significance and Aim of the Study

With the rise of Web 2.0 technologies, internet-based applications and smartphones have become an ever-present part of our lives and started to act as a critical platform for disseminating information. Social media platforms are standing at the center of this information flow and social interaction (Raacke & Bonds-Raacke, 2008), and consumers' preference for engaging with a third-party platform than communicating directly with the company (Vermeulen & Seegers, 2009) made these platforms a meeting point for those who want to share their consumption experience. Consumers' inclination to share their experiences online and the resultant upsurge in the volume of electronic WOM developed companies' understanding about consumers' wants, needs, complaints, and satisfaction levels and provided them the opportunity to both discover and recover dissatisfactions. However, such opportunities come at a cost. Even though an effectively handled online complaint may result in winning the complainant back and attract new customers through shaping reputation (Einwiller and Steilen, 2015), it can dangerously backfire on the company when not handled effectively (Lee and Song, 2010). Therefore, effective online complaint management in social media platforms is vital for companies.

The extensive social media usage for communicating complaints resulted in a high volume of social media big data. Handling these massive amounts of unstructured textual data in social media requires a vast amount of time and budget, which constitute a vital burden for companies. However, companies should handle these effectively and efficiently to avoid negative consequences of customer dissatisfaction resulting from negative experiences. In addition to these negative consequences such as double deviation and decreased repurchase intention, unhandled online complaints may trigger outrage and initiate an online firestorm that can cause further severe damage to the brand image and reputation. Therefore, it is significant for companies to find an effective way of handling online complaints.

The vital point for finding the balance between response time and cost of handling is dividing complaints into categories that can act as a base for further prioritization. By including personal characteristics of complainants, interaction rates of the complaint, sentiment polarity, content of the complaint, and emotion to the equation, companies may better analyze both the complainants and the riskiness of complaint regarding initiation of further online crises. Thus, this thesis aims to identify content-based and interaction-based online consumer complaint types and predict complaint types according to the complaint magnitude rooted in complainants' personality traits, emotion, Twitter usage activity, and complaint's sentiment polarity, and feedback frequency. Following its aim, this thesis addresses the following research questions:

RQ1: What are the characteristics of the content-based complaint categories (legal threat, call for boycott, switching intention)?

RQ2: What are the interaction-based types of online complaints regarding complainants' personality traits, emotions, and Twitter usage activity, as well as the complaint's sentiment and feedback frequency?

RQ3: What are the characteristics of the interaction-based complaint clusters?

RQ4: How can the complaint type be predicted via machine learning and text analytics?

RQ5: How can online complaint typology be used for prioritizing customer complaints?

This thesis intends to provide evidence for developing an efficient online complaint handling framework that uses social media big data. The contributions of this thesis can be listed as; (1) identifying types of online complaints that reveal diverse customer complaint behavior on Twitter; (2) proposing a set of content and profile features that can be employed for revealing complaint type; (3) revealing the relationship between content features, profile features, and online complaint type (4) proposing a novel approach for excelling online complaint through prioritization.

4.1.2. Research Model

As mentioned previously, this thesis' objective is to establish an online consumer complaint typology and predict complaint type following service failure in accordance with the complaint magnitude rooted in personality traits, emotion, sentiment, and Twitter usage activity. In accordance with its aim, this thesis employs a multi-level approach for identifying interaction-based and content-based categories and revealing their characteristics. Accordingly, the model utilizes a two-step machine learning approach. In the first step, content features, including sentiment polarity, emotion, and personality traits, are determined through text analytics. In the second step, the relationship between the clusters and the independent variables mentioned in the previous step is revealed through multinomial logistic regression, a machine learning algorithm that is mainly used to evaluate the probability of categorical membership.

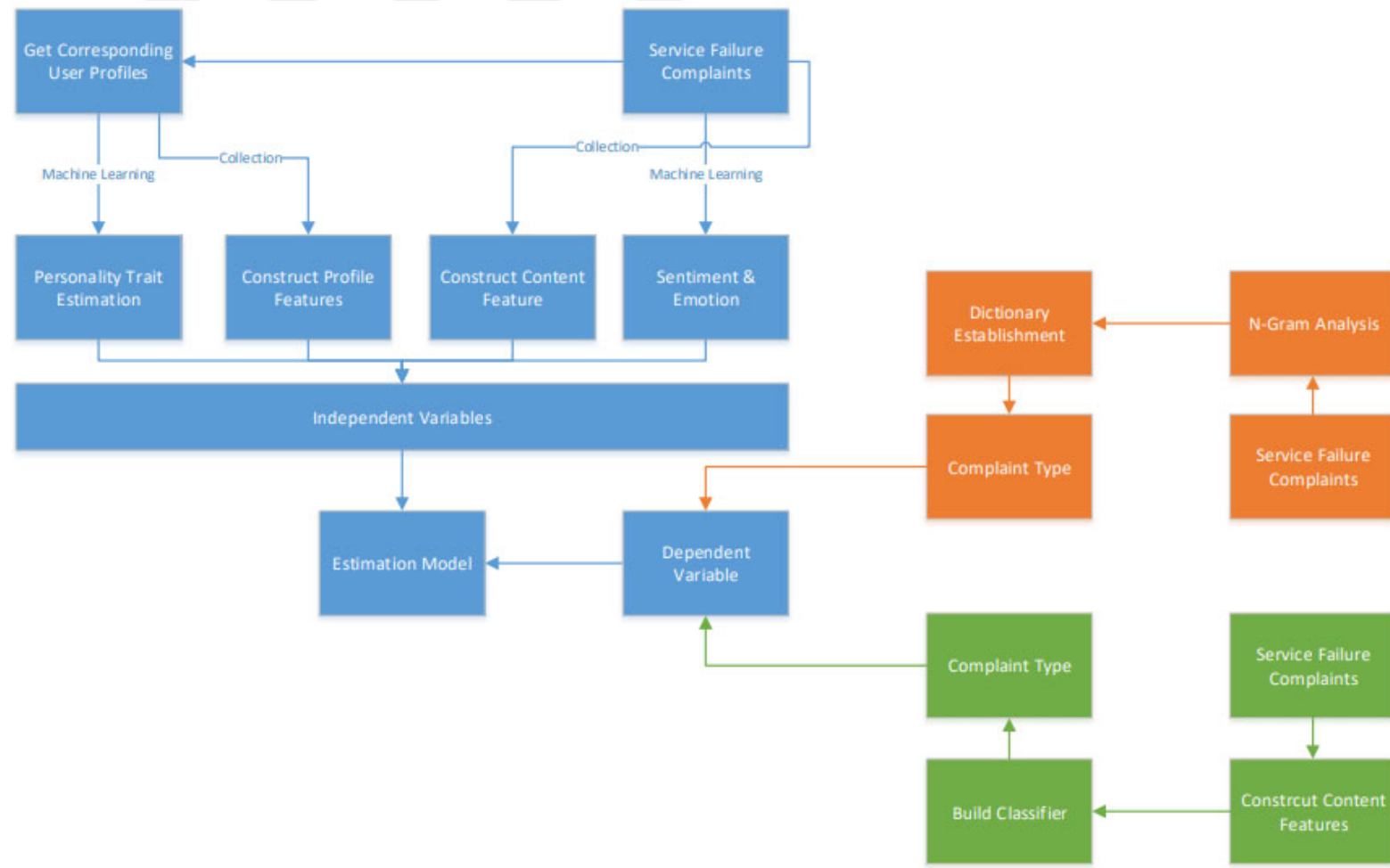
The analysis procedure begins with the gathering of public complaint tweets about listed twelve service providers via the employment of hashtag-based, keyword-based, and entity-oriented search queries, and obtaining those complaints' content features consisting of the number of mentions, favorites, replies, consecutive tweets, and used hashtags. In order to obtain the sentiment polarity of gathered complaints in terms of being positive, negative, and neutral, text analytics tools are being utilized. In the second step, profile features consisting of all extant tweets of complainants are collected and listed. From the gathered Twitter usage activity, complainants' personality traits and emotions are being obtained by employing text analytics tools. These two steps provide the independent variables to be used in the clustering process.

After collecting independent variables, complaints are grouped through (1) content-based classification, (2) interaction-based clustering. During content-based classification, complaints are being grouped in accordance with the behavioral intention mentioned by the complainant. Following the extant CCB models in the literature, complaints are classified into three categories; complaints that are (a) containing a legal threat, (b) calling for a boycott, and (c) containing brand switch warning. On the other hand, during interaction-based clustering, complaints are grouped in accordance with their interaction rates in terms of the number of favorites, retweets, consecutive tweets, and replies. At the end of this step, complaints are clustered according to their feedback

frequency. After completing the classification and clustering processes, the relationship between personality traits, emotions, sentiment, and complaint types is examined through multinomial logistic regression. The research model and the steps followed are depicted in Figure 8. The blue color section represents the collection of independent variables and the interface between independent variables and the dependent variable. The section highlighted in orange represents content-based classification, whereas the green section represents the interaction-based clustering phase.



Figure 8: Research Model



4.2. RESEARCH METHOD

4.2.1. Research Design

A research design depicts a study's framework by detailing all procedures essential for gathering the requisite data and its analysis for concluding (Malhotra and Malhotra, 2012). In this regard, a decent research design explicates the study's type, research setting, the extent of researcher interference, temporal features, and data analysis process (Sekaran, 2003). Furthermore, it is essential to explain dependent and independent variables, sampling procedure and frame, measurement, data collection procedures, and data analysis techniques for providing a comprehensive framework. Accordingly, this section is designated to explain the design of this research.

4.2.1.1. Independent and Dependent Variables

In accordance with this thesis' context, independent variables refer to the metrics used to classify and cluster online complaints. As mentioned previously, two diverse sets of features are employed for classification and clustering; content features and profile features. Therefore, it may be noted that content and profile features form independent variables of this research. Since these constructs are used for determining the type of online complaint, complaint type is the dependent variable of the research model.

Content features include emotion, sentiment polarity, and personality traits of complainants. *Sentiment* refers to the tone in which an individual articulates his/her subjective judgment. In the customer complaint domain, sentiment polarity can be defined as the semantic orientation of the words used in the articulation of complaints. *Emotions* reflect individuals' mental state of readiness arising from cognitive evaluations of events or thoughts (Bagozzi et al., 1999). In a service failure setting, emotions are consequences of customers' stressful cognitive assessment (Lazarus, 1991). Finally, *personality traits* are enduring, cross-situational consistencies in behavioral and response patterns (Harris and Mowen, 2001; McCrae, 2009). In this thesis, the Big Five taxonomy, which lists the five common traits of agreeableness, openness to experience, extroversion, neuroticism, and conscientiousness, is employed for revealing complainants' personality.

Profile features employed in this thesis include the count of favorites, retweets, and consecutive tweets. *Favorite count* reflects the number of times other users like a tweet of the user. A retweet represents a repost or forwarding of a tweet by another user. Accordingly, the *retweet count* represents the number of times another user forwards a tweet. Lastly, *consecutive tweet count* reflects the number of connected tweets during which the user articulates his/her complaint.

4.2.1.2. Selection of the Research Platform

The internet, web 2.0 technologies, and social media platforms provide consumers the opportunity of voicing their complaints quickly and conveniently. The ease of use, the ability to reach a wider audience, and communicating with the target companies conveniently, made these platforms vastly used outlets for articulating their negative experiences and complaints with service providers. Hence, these novel complaint channels are now being widely used to complain.

Facebook and Twitter are the major social media channels that are extensively preferred for communicating complaints (Einwiller and Steilen, 2015). However, what makes Twitter more alluring for this research is the congruity of its nature with this thesis' aim. Twitter enables tracking consumer opinions in a real-time setting, and 70 percent of the Fortune 500 companies are actively using a Twitter account (Slegg, 2013). Besides, Twitter is one of the most-visited sites for opinion sharing with 330 million monthly active global users, 40 percent of whom visit the platform on a daily basis (Lin, 2020). Moreover, around 500 million people visit Twitter without logging into an account and spend 3.39 minutes per session on average (Statista, 2019). Being among the top the social media sites, Twitter is an important platform to gather consumer insights. Therefore, Twitter provides a rich lens for researchers through which the dynamics of the communication between companies and complainants vis social media platforms can be analyzed (Abney et al., 2014).

The use of Twitter in the complaint handling context is apparent in the marketing practice. Several companies conduct complaint handling on their Twitter accounts, including Nike, Amazon, and UPS. Moreover, most of these companies moved their presence one step further by creating accounts for handling complaints, such as

@AmazonHelp, @UPSHelp, and @FedExHelp. Companies' strong presence on Twitter is not in vain. A survey conducted by Lithium Technologies revealed that 10 percent of Twitter users' tweet about their complaints, and 65 percent of those users indicated that they expect a response from companies in less than 2 hours (Vaughan, 2014). The same research put forth that 60 percent of the respondents cited negative consequences to the company if their complaint remains unanswered promptly, 29 percent reported that they would tell their friends and family about the experience, 26 percent would communicate their concern to the company via other sources of communication, 24 percent would have decreased repurchase intention, 21 percent would not recommend the company, and 15 percent would shame the company in social media. Due to the platform's mentioned intensive usage by both companies and consumers and the rich lens it provides, Twitter is employed as the research platform of this thesis.

4.2.1.3. Sampling Frame

To determine service-providing companies to be examined, the Standard and Poor's (S&P) 500 Index, which lists the market-capitalization-index of the 500 largest U.S. companies that are traded publicly, is utilized. The reason behind choosing S&P 500 among other indices is that it is regarded as the preeminent and the most commonly referenced gauge, providing the ability to track the largest-cap and, hence, the leading U.S. companies (Tracy, 2020).

As the first step, the S&P 500 Index is obtained from Thomson Reuters DataStream, and service providing companies are filtered and listed. As of the second step, among the listed service providing companies, B2B service providers such as advertising and marketing agencies were eliminated for ensuring congruity between data and the aim of the thesis. In the subsequent step, acquired B2C service providers were grouped by their economic sectors indicated by Thomson Reuters; (1) Financials, (2) Cyclical, (3) Healthcare, (4) Industrials, (5) Non-cyclical, (6) Telecommunication. Service providers operating under each group are ranked according to their market cap, and two companies with the largest market cap from each group are selected.

Consequently, twelve service providers were selected; JPMorgan Chase & Co., Bank of America Co., Amazon.com, Home Depot, UnitedHealth Group Inc., CVS Health Corp., United Parcel Service Inc., FedEx Corp., Walmart Inc., Walgreens Boots Alliance Inc., AT&T Inc., and Verizon Communications. The Market-cap ranking of the selected B2C service providers is listed in the following table.

Table 6: Market Cap Ranking of the Selected B2C Service Providers

Economic Sector	Market Cap	Service Provider
Financial	\$3.546.128.697.589,39	JPMorgan Chase & Co.
		Bank of America Corp.
Cyclicals	\$2.956.829.752.642,83	Amazon.com
		Home Depot
Healthcare	\$730.764.192.092,89	UnitedHealth Group Inc
		CVS Health Corp.
Industrials	\$570.557.191.860,09	United Parcel Service Inc.
		FedEx Corp.
Non-Cyclicals	\$534.593.721.858,74	Walmart Inc.
		Walgreens Boots Alliance Inc
Telecommunication	\$498.009.337.953,66	AT&T Inc.
		Verizon Communications

Source: S&P 500 Index

4.2.1.4. Data Collection

Even though various software tools exist for collecting real-time data from Twitter, there are two main options; purchasing Twitter API and undertaking in-house software development. Since the former is more restrictive than the latter and can be utilized for only a given period, the latter is chosen for data collection. Concordantly, historical Twitter data is obtained through a program that is coded with C Sharp in a way that enables collecting data regarding Tweet text, used hashtags, time and date of Tweet, the screen

name of the user, user ID number, Tweet ID number, number of replies, retweets, favorites, mentions, hashtags, and consecutive tweets.

For ensuring that all relevant public information is reached, data is collected via three diverse search methods; (1) through hashtags, (2) from customer care accounts, (3) through keyword search. In the first method, data collection through hashtags, previously collected complaint hashtags are searched (e.g., #amazonfail, #homedepotlies, #uhcfail), and public complaints that contain these hashtags were collected. The second method aimed at covering complaint tweets that do not directly cover company name and/or company-related complaint hashtags. For reaching these complaints, the “tweets and replies” section of the customer care accounts is utilized, and from the replies of customer accounts, related complaint tweets were reached. The third method, data collection through keyword search, aimed at reaching the complaints that do not contain any hashtag and mention customer care accounts. Accordingly, keyword combinations consisting of the target company’s name and previously determined keywords were operated. In accordance with the number of keywords, eight searches were initiated for each company (e.g., “UPS” and “bad service,” “UPS” and “poor service,” “UPS” and “fail,” “UPS” and “bad experience,” “UPS” and “bad customer experience,” “UPS” and “poor experience,” “UPS” and “bad customer service,” “UPS” and “poor customer service”).

Before the initiation of data collection, information regarding customer care accounts of companies extant in the research sample, the opening date of customer care accounts, company-specific hashtags, and keywords were gathered. Table 7 lists company names, the company’s customer care account’s opening date, company-specific hashtags, and search keywords. For ensuring consistency, data is obtained starting from the opening date of the newest customer care account that is the customer care account of Walmart. Accordingly, extant public data published between March 1, 2017, and December 31, 2018, were collected. The tweet gathering process resulted in 297,374 complaint tweets. After collecting complaint tweets, each complainant’s tweet history is gathered to obtain their personality traits. Compilation of complainant’s tweet history resulted in 2,183,368 user tweets.

Table 7: Companies' Customer Care Accounts, Related Hashtags and Keywords

Company Name	Customer Care Account	Opening Date of Customer Care Account	Company-Specific Hashtags	Keywords
JPMorgan Chase & Co.	@ChaseSupport	March 2011	NA	"bad service" "poor service" "bad experience" "bad customer experience" "poor experience" "bad customer service" "poor customer service" "fail"
Bank of America Corp.	@ BofA_Help	January 2009	NA	
Amazon.com	@AmazonHelp	October 2009	#amazonfail, #fuckamazon, #amazonpoorservice, #amazonstelas, #bewareofamazon, #amazoncheats, #amazonsucks, #notsoprime	
Home Depot	@ HDCares	December 2010	#homedepotsucks, #homedepotlies, #dontshophomedepot, #dontbuyfromhomedepot, #homedepotnightmare	
UnitedHealth Group Inc	@askUHC	March 2015	#uhcfail, #unitedhealthcarefail, #unitedhealthcarefraud	
CVS Health Corp.	@CVSHealth	March 2010	NA	
United Parcel Service, Inc.	@ UPSHelp	March 2011	#upssucks, #upslies, #neveruseups, #ihateUPS, #upscustomersupport, #upsfail, #upsruinedchristmas, #upsfailsagain, #neverups, #donttrustups,	

			#uselessparcelservices, #upsareliars	
FedEx Corp.	@ FedExHelp	November 2012	#fedexsucks, #fedexfail, #fedexlies, #fedexcustomerservicesucks, #badfedex, #fedexbadservice, #donttrustfedex	
Walmart Inc.	@walmarthelp	March 2017	#hatewalmart, #fuckwalmart, #walmartsucks, #walmartdoesntcare, #dontshopwalmart, #walmartfail	
Walgreens Boots Alliance Inc.	@Walgreens	June 2009	#walgreensfail	
AT&T Inc.	@ATTCares	August 2009	#ATTdoesntcare, #attthieves, #attlies, #attsucks, #attwirelessucks, #atandtsucks, #neveragainatt, #ihateatt, #attistheworst,	
Verizon Communications	@VerizonSupport	May 2009	#verizonfail, #verizonlies, #verizonsucks, #fuckverizon, #verizonoutage, #verizondoesntcareaboutcustomers, #sueverizon, #ihateverizon, #donttrustverizon, #verizonistheworst	

4.3. DATA ANALYSIS

4.3.1. Descriptive Statistics

Summary statistics for the variables of the thesis are presented in Table 8. It covers the mean, standard deviation, median values, and the number of observations for the variables that are employed in the multinomial logistic regression, content-based classification, and interaction-based clustering. In descriptive statistics, it can be observed that the median values for all variables are close to the mean values, indicating that the data are evenly distributed around the mean, signifying that the price series do not vastly disperse from the normal distribution. Furthermore, the standard deviation parameters for all independent variables are around the mean.

Table 8: Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Mention#	36471	,000	21	1,160	,817
Hashtag#	36471	,000	19	,320	,878
Reply#	36471	,000	1694	2,210	18,880
Retweet#	36471	,000	3955	1,406	34,604
Favorite#	36471	,000	20676	6,153	170,483
Sentiment	36471	,000	,999	,317	,307
Surprise	36471	,000	,331	,030	,038
Sadness	36471	,000	,426	,021	,039
Disgust	36471	,000	,292	,010	,027
Anger	36471	,000	,528	,009	,029
Joy	36471	,000	,744	,046	,087
Fear	36471	,000	,648	,022	,045
Agreeableness	36471	,000	,999	,358	,266
Conscientiousness	36471	,000	,998	,359	,263
Extraversion	36471	,000	,999	,492	,291
Neuroticism	36471	,000	,999	,433	,320
Openness	36471	,000	,999	,564	,272
Following	36471	,000	447029	1099,87	6262,187
Followers	36471	,000	108460488	24070,93	886842,283
Sticky2	36471	,000	25	,450	1,636
Valid N (listwise)	36471				

4.3.2. Classification and Clustering

For determining the type of the complaint, content-based classification and interaction-based clustering techniques are employed. Since the content-based categories are obtained through revising the extant CCB models in the literature, a set of keywords for each content-based complaint type (legal threat, call for boycott, brand switch warning) are determined by utilizing n-grams. Following the determination of keywords, rule-based classification is conducted for identifying content of the complaints. On the other hand, interaction-based clusters are identified via two-step clustering technique and the revealed interaction-based clusters are labelled according to their characteristics. This section provides the details regarding employed classification and clustering methods, obtained categories and clusters, and their characteristics.

4.3.2.1. Content-based Classification

Content-based classification is employed to extract content-related information from the complaint tweets and group these complaints according to their content. Following service failures, one of the most vital aspects is understanding and estimating how the dissatisfaction caused by the negative incident will reflect on the company. Since these reflections portray the cost of failure for the company, it is critical to ensure post-complaint satisfaction and to take actions to prevent complainants from carrying their dissatisfaction one step further (Davidow, 2000). The extant consumer complaint models mainly consider the offline environment; therefore, applying these models to online platforms requires some modifications. Thus, before the explanation of consumer reactions, the CCB model of Hirschman (1970), Day and Landon (1977), Singh (1988), and Tronvoll (2012) should be revisited for the online environment.

Hirschman (1970) advocates that customers either exit or voice their complaints, and their degree of loyalty towards the company impacts the relationship between the negative experience and complaint behavior. Day and Landon (1977) indicate that consumers may either; (1) take public action through seeking redress, complaining to third party agencies, and taking legal action, or (2) may take private action through boycotting the company and spreading negative WOM. Singh (1988) divides customer reactions into three categories as; (1) voice (articulation of complaint to the company for

obtaining a redress), (2) private response (boycotting and WOM), and (3) sharing the incident with third parties (court, consumer protection agencies, etc.). Finally, Tronvoll (2012) advocates that consumers may voice their complaints or take active/passive action.

Given the nature and dynamics of social media, two key aspects are valid for the entire mentioned CCB models; (1) the audience of the public response is wider than the offline environment, complaints remain until deleted by the complainant, and hence they are more accessible, (2) private responses (e.g., boycott, negative WOM) are no longer private due to accessibility of social media. When the accessibility of messages posted on social media is considered, it may be stated that the distinction between publicity and privacy of responses is blurred. Even though the complainant's account may be only accessible to his/her friends or followers, pure privacy cannot be advocated due to the complainant's extant followers/friends. Concordantly, the nature of social media diminishes the importance of the distinction between public and private reactions. However, a distinction based on previous models can be put forth by following the indicated further action. When the extant models and the nature of social media platforms are considered, three categories of reactions can be revealed: taking legal action, calling for a boycott, and brand switch warning (referred to as "exit" in previous models). Since publishing a complaint message on social media is "voicing" your experience in a way that enables constant "WOM," voicing the complaint and WOM behavior are not included in our categorization.

4.3.2.1.1. N-Grams

N-grams are a part of Natural Language Processing (NLP) and refer to "n" words' sequence. In other words, *n*-grams may be defined as a set of co-occurring words within a given frame. Since they are a combination of words, *n*-grams are not limited to complete phrases. However, *n*-grams are above words in the natural hierarchy since any language can be fragmented into its component *n*-grams in a similar way that a word can be broken into morphemes or letters. (Shaoul et al., 2013).

Given that *n*-grams are the sequence of *n* words, the *n*-gram case where *n* equals one is called "unigram." In a similar vein, the sequence of two words (*n*=2) is referred to

as “bigram,” and the sequence of three words ($n=3$) is called “trigram.” Given a sentence, a list of n -grams can be constructed by finding pairs of words. If X is the number of words in sentence S , the number of n -grams for sentence S would be $X-(N-1)$. For instance, given the sentence “Early bird catches the worm” consisting of five words, n -grams with the length of four words can be constructed by finding consecutive pairs of words. Then, n -grams for the example sentence may be listed as:

Unigrams ($n=1$): early, bird, catches, the, worm

Bigrams ($n=2$): early bird, bird catches, catches the, the worm

Trigrams ($n=3$): early bird catches, bird catches the, catches the worm

Fourgrams ($n=4$): early bird catches the, bird catches the worm

N -grams that provide all combinations of contiguous words/letters in the source text are being employed with various purposes. Word frequency estimation, speech recognition, automatic completion of sentences, tokenization of a text/sentence by word, text summarization, and automatic spell checks constitute some examples for n -gram usage. It should be noted that the determination of “ n ” is of vital importance in constructing n -grams. If the “ n ” value is too small and the n -grams are too short, significant differences may be hard to capture, whereas if the “ n ” value is too high and the n -grams are too long, a general pattern may be missed. The optimum value of n , which can also be regarded as the optimum length, depends on the application.

4.3.2.1.2. Dictionary Establishment

For determining the key phrases to be employed in the content-based categorization of complaints posted on Twitter, n -grams are created by using the Natural Language Toolkit (NLTK) suite of Python. Due to nature and short sentences used on Twitter, unigram, bigram, trigram, and fourgram phrases are generated. A list of entire unigrams, bigrams and trigrams that appear at least four times in the sample is generated. The n -gram generation process resulted in 124,863 n -grams (7,615 unigrams, 45,352 bigrams, 47,217 trigrams, and 24,679 fourgrams). Table 9 represents some examples regarding each n -gram (unigram, bigram, trigram, and fourgram) output.

Table 9: Examples For n-grams

	Term	Frequency
Unigram		
1	delivery	12146
2	last	1499
3	point	1235
4	help	3844
5	never	2930
Bigram		
1	custom, support	278
2	wait, for	1713
3	suppose, to	2189
4	unfortunate, that	555
5	no, one	1071
Trigram		
1	was, suppose, to	1311
2	your, custom, service	413
3	to, the, wrong	653
4	to, be, deliver	1348
5	can, you, help	370
Fourgram		
1	where, is, my, package	235
2	end, of, the, day	189
3	was, suppose, to, be	712
4	have, been, wait, for	121
5	to, an, access, point	110

Prior to determining key phrases, *n*-grams containing topic-specific terms and proper nouns were eliminated to ensure that no specific named-entities are included. The remaining *n*-grams are examined manually to properly allocate phrases to categories of legal threat, call for boycott, and brand switch warning. The final key phrase list contains twelve key phrases for “legal threat,” seven for “brand switch warning,” and twelve for “call for boycott” categories. The key phrase lists are provided in Table 10.

Table 10: Key Phrases for Content-Based Classification

Legal Threat	Call for Boycott	Brand Switch Warning
consumer protection	boycott	will be buying
state law	will never shop	will buy
lawyer	never again	switch to
unlawful	never use again	somewhere else
lawsuit	never shop again	will use
getting a lawyer	done with	will be using
court	stop using	from now on
illegal	stop buying	
against law	avoid using	
legal abuse	will no longer use	
sue	will no longer buy	
attorney	never buying again	

4.3.2.1.3. Results of Content-Based Classification

Following the establishment of dictionaries to classify complaints based on their contents, legal threat, call for boycott, and brand switch warning categories are formed. The identification of the content-based categories of obtained complaint tweets is conducted utilizing a specially utilized Python script. In total, 584 points in the dataset fell in one of our categories. Among these complaints, 39% contained a call for boycott, 35% brand switch warning, and 26% legal threat. The following headings provide details of legal threat, call for boycott, and brand switch warning categories.

4.3.2.1.3.1. Legal Threat

From a consumerist point of view, manufacturing companies and retailers are responsible for both encouraging consumers to communicate their dissatisfaction to them and recognizing that they are legally liable for protecting consumer rights (Cri , 2003; Rousseau, 2003). On the other hand, from the consumers' perspective, taking legal action can be considered a problem-focused coping method and a way of protecting their rights when the manufacturer or retailer does not redress failure. Therefore, when consumers have a negative experience with companies, they may take legal action towards the company or threaten companies by reminding them of legal regulations.

Along with complaining to the responsible company, public consumer protection agencies, and non-governmental voluntary organizations, taking legal action is a type of private action taken by consumers following a failure (Day and Landon, 1977; Phau & Sari, 2004). Reasons for contacting a legal representative following a failure include seeking assistance for obtaining redress when consumers' personal efforts fail, standing for consumer rights, noticing other individuals and/or organizations about the incident, and easing the anger/frustration against the company (Donoghue, 2008). Content-based classification revealed that only a minority of complainants emphasize a legal threat on Twitter. Results point out that only 26% (n=148) of the complaints that indicate a behavioral intention contains a legal threat. The majority (43.2%) of legal threat containing complaints are directed to companies operating in the industrials industry consisting of UPS and FedEx. It is followed respectively by cyclicals (Amazon and Home Depot) with 33%, non-cyclical (Walmart and Walgreens) with 16.4%, and healthcare (UHC and CVS) with 7.4%. Some examples for the complaint tweet that contain legal threat are provided in Table 11.

Table 11: Examples of Complaint Tweets that Contain Legal Threat

Complaint Tweet	Company	Key Phrase
@AmazonHelp Amazon has gone one step further in online shopping. Are you really going to deliver me a product that is unavailable for some time now? Should I sue you as you state 'Guaranteed Delivery' over there and do not provide option me to proceed.	Amazon	Sue
@CVSHealth since your insurance doesn't want to pay to replace my recalled blood pressure medication, who should my family contact about the lawsuit when I stroke out? #youdontcare	CVS	Lawsuit
@FedExHelp I swear to God if your damn delivery driver throws another package or box of mine I am going to sue your company for property damage! The driver didn't even put it on my porch! Left a 4-pound box at the bottom of my steps! Lazy much???	FedEx	Sue

4.3.2.1.3.2. Call for Boycott

The traditional CCB literature views boycott behavior as a private, “hidden,” indirect action taken by consumers that targets a retailer, brand, or product type (Day et al., 1981). However, in online environments, especially in social media platforms, articulation of boycott intention eliminates its feature of being “hidden” and turns into a call for a boycott that can turn into an online firestorm within the blink of an eye. Friedman (1985) defines such attempts by consumers as “an attempt by one or more parties to achieve certain objectives by urging individual consumers to refrain from making selected purchases in the marketplace.”

Boycotts are extreme consumer behavior cases and can significantly influence purchase decisions (Klein et al., 2004). Initiating boycotts with a high participation rate is a challenging achievement due to the social dilemma it creates. By boycotting a product, service, brand, retailer, or manufacturer, consumers choose between the personal benefit of consuming and the desire of refraining from consumption collectively so that all participants receive the shared benefits of an effective boycott (Sen et al., 2001). However, once it is achieved, companies face dramatic results of poor complaint management and consumer dissatisfaction varying from damaged brand image to decreased revenue. By being aware of the criticality of such cases for companies, the majority of the complaints that indicate a behavioral intention contains a call for a boycott. Results reveal that calls for boycott account for 39% (n=229) of the Twitter complaints that indicate a behavioral intention. The vast majority, referring to 76.8%, of complaints that contain a call for a boycott is directed to service providers operating in the industrials industry consisting of UPS and FedEx. It is followed respectively by cyclicals (Amazon and Home Depot) with 11%, non-cyclicals (Walmart and Walgreens) with 8.7%, and healthcare (UHC and CVS) with 3.5%. Some examples for the complaint tweet that contain legal threat are provided in Table 12.

Table 12: Examples for Complaint Tweets that Contain Call for Boycott

Complaint Tweet	Company	Key Phrase
What's the actual point of giving customers confirmation if you can't be bothered in delivering it?? This was supposed to be a happy time but I am so angry at this... That's me done with @AmazonUK , I'll go elsewhere now #notsomerrychristmas #is9daysnotenoughtime	Amazon	Done with
@FedEx Having my shipment not move for 3 days 100 miles from my home is why I try to avoid using your company. I'm sure it's damaged	FedEx	Avoid using
@FedEx Price gauging people sending food relief packages to Puerto Rico charging \$400 - \$600? SHAME ON YOU! Boycott FEDEX! Use ANOTHER carrier!	FedEx	Boycott

4.3.2.1.3.3. Brand Switch Warning

Switching refers to the probability that customers switch to other competitor brands due to their dissatisfaction with service providers' post-failure responses (Lin, 2010). Keaveney (1995) listed eight factors that cause consumers to switch to another service provider as; controllable core services containing the price, convenience, service attitude and environment, business moral, and uncontrollable factors containing competition and non-volunteer switching. Since consumers' switching intention denotes that the service provider cannot meet consumers' desired parameters, whether the triggering reasons for such intention are company driven or are a consequence of consumers' internal psychological assessment, these should be scrutinized and managed by companies (Riaz and Khan, 2016).

As intention precedes action, brand switch warnings articulated on social media are critical for companies due to a widespread audience. Following a boycott call, complaints that contain brand switch warnings constitute the second-largest group consisting of complaints that indicate an intention. Results show that brand switch warning accounts for 35% (n=207) of the Twitter complaints that indicate a behavioral intention. The majority, referring to 45.9%, of complaints that contain a call for a boycott

is directed to service providers operating in the industrials industry consisting of UPS and FedEx. It is followed respectively by cyclicals (Amazon and Home Depot) with 42.5%, non-cyclicals (Walmart and Walgreens) with 7.3%, and healthcare (UHC and CVS) with 4.3%. Some examples for the complaint tweet that contain legal threat are provided in Table 13.

Table 13: Examples for Complaint Tweets that Contain Brand Switch Warning

Complaint Tweet	Company	Key Phrase
I need to find somewhere else to shop because @amazon has really been pissing me off lately...	Amazon	Somewhere else
Hey @FedEx how about that no-show pickup today in Wolfeboro! How am I supposed to run my business w this kind of irresponsibility? Luckily the folks at your local drop-off came and picked up our packages to go out tonight. Might be time to switch to @UPS!	FedEx	Switch to
@Walmart I'LL ONLY BE GOING BE GOING TO TARGET FROM NOW ON TURN ON THE AC IN YOUR STORES YOU SHOULD BE ASHAMED OF YOURSELVES. TURN ON THE AC!	Walmart	From now on

4.3.2.2. Interaction-based Clustering

Social media users are inclined to use various cognitive heuristic clues; in other words, mental shortcuts, for processing information effectively (Lahuerta-Otero et al., 2018). Social media users mainly prefer using these heuristic cues over conducting systematic information processing to handle uncertainty and information overload in online environments (Metzger et al., 2010). Since social media users prefer to spend less cognitive effort unless necessary, they tend to process messages heuristically (Metzger et al., 2010; Zhang et al., 2014) by using various message and content-related elements. Accordingly, message-related elements (i.e., likes, favorites, retweets) that constitute social media posts' characteristics act as heuristic cues for social media users (Zhang and Watts, 2008) and are being used for judging the validity of content without further processing (Chaiken, 1987). In a social media setting, various attributes ranging from the author's popularity within the network to the interaction level act as vital cues that project

credibility and trustworthiness (Metzger et al., 2010; Kwak et al., 2010; Lahuerta-Otero et al., 2018).

In social media, some posts get more interaction and reach higher diffusion by becoming more popular than the other posts, whereas some posts attract minimal or no attention and cannot diffuse. Diffusion of social media complaints plays a crucial role for both other users and targeted companies. From the consumers' standpoint, a higher interaction rate indicates support and verification of other people and increases the probability of obtaining more attention from the targeted service provider. On the other hand, from the service providers' standpoint, complaints with higher interaction rates signal an escalated probability of a consumer boycott, an online firestorm, or an agreed-upon dissatisfaction that can lead to a sharp decrease in income. Moreover, such diffusions cause vital damage to a brand's image and require further marketing decisions for rebuilding the relationships with dissatisfied consumers and reshaping the brand image. Therefore, understanding interaction-based types of online complaints is crucial for companies. Accordingly, the count of retweets, consecutive tweets, favorites, and replies are used in interaction-based clustering.

How many times a tweet is replied, favorited, and retweeted reflects the attention it drew. Another critical factor in determining the interaction-based impact of a complaint is the count of consecutive tweets. Count of consecutive tweets reflects the number of complaint tweets posted by the complainant. Since longer tweets are more informative than the shorted ones, they are found to attract more attention on social media platforms (Chen et al., 2012; Zhang and Peng, 2015). Enge (2014) validated the relationship between tweet length and interaction by exploring that longer tweets get more retweets and likes. Hence, by employing a two-step clustering method, complaints are clustered in accordance with their count of replies, favorites, retweets, and consecutive tweets.

4.3.2.2.1. Two-Step Cluster Analysis

Two-step cluster analysis is a statistical procedure mainly used to identify similar groups of people or objects in data sets. Two-step clustering involves; (1) pre-clustering, (2) clustering. In the pre-clustering stage, data points are scanned one by one, and it is examined if the current data point can be added to a previously formed cluster or if it

starts a new one, based on the distance criterion (Satish and Bharadhwaj, 2010). In the following phase, the clustering step, sub-clusters obtained in the first step, is being used as an input for forming the desired number of clusters. It should be noted that the two-step clustering method assumes all variables as independent, continuous variables to have a normal distribution, and categorical variables to have a multinomial distribution (Norusis 2007). To sum up, the two-step clustering method addresses the scaling problem by forming pre-clusters in the first step and treating these clusters as single cases in the second step that employs hierarchical clustering (Satish and Bharadhwaj, 2010).

Two-step cluster analysis, hierarchical and *K*-means cluster are widely used clustering methods, and each of these relies on a diverse algorithm in forming clusters. Data type plays a crucial role in the determination of the clustering technique. Hierarchical data is suitable for small datasets, *K*-means is restricted to continuous values, and a two-step cluster is suitable for forming clusters with both continuous and categorical variables (Schiopu, 2010). The two-step method is also suitable for the cases where categorical variables have three or more levels, and it is capable of concurrently succeeding in forming groups that vary on criterion variables and establishing significant relationships between the clusters, the categorical variables, and the continuous variables (Satish and Bharadhwaj, 2010). Moreover, two-step clustering is considered more reliable and accurate than other clustering techniques and distinguished from others by providing full information and rich explanation (Norusis, 2007). Finally, it is primarily designed for analyzing large sets of data and is capable of determining the optimal number of clusters (Satish and Bharadhwaj, 2010).

In this thesis, the two-step clustering technique is employed due to its congruity with both the dataset and nature of this research. The main reasons for the suitability to the context and data of this research can be listed in detail as: (1) its suitability with the large dataset, (2) capability of providing reliable and accurate outcomes, and (3) its ability to provide a rich explanation on the significant relationships. Accordingly, based on the explained reasons, a two-step clustering method is used in interaction-based clustering.

4.3.2.2.2. Results of Interaction-based Clustering

Interaction-based clusters were formed by employing the two-step clustering method via the SPSS software package. Cluster analysis revealed that a three-cluster solution would be optimal. The smallest cluster contains 153 cases referring to 0,4% of the sample, whereas the largest one contains 35331 cases referring to 96,9%. In clustering, a silhouette of more than 0.0 is required for the within-cluster distance and the between-cluster distance for ensuring validity (Norusis, 2007). The analysis revealed a silhouette measure of cohesion and separation of 0.864, which can be considered a “good” level of validity. Table 14 displays the percentages of cases of each cluster, and Figure 9 depicts cluster quality. Moreover, the variables’ individual predictor importance score ranges from 0 (the least important) to 1 (the most important) was assessed. All constructs (number of replies, consecutive tweets, favorites, and retweets) were found to have a predictor importance score of 1, and thereby, the reliability of clusters is ensured.

Table 14: Cluster Distribution

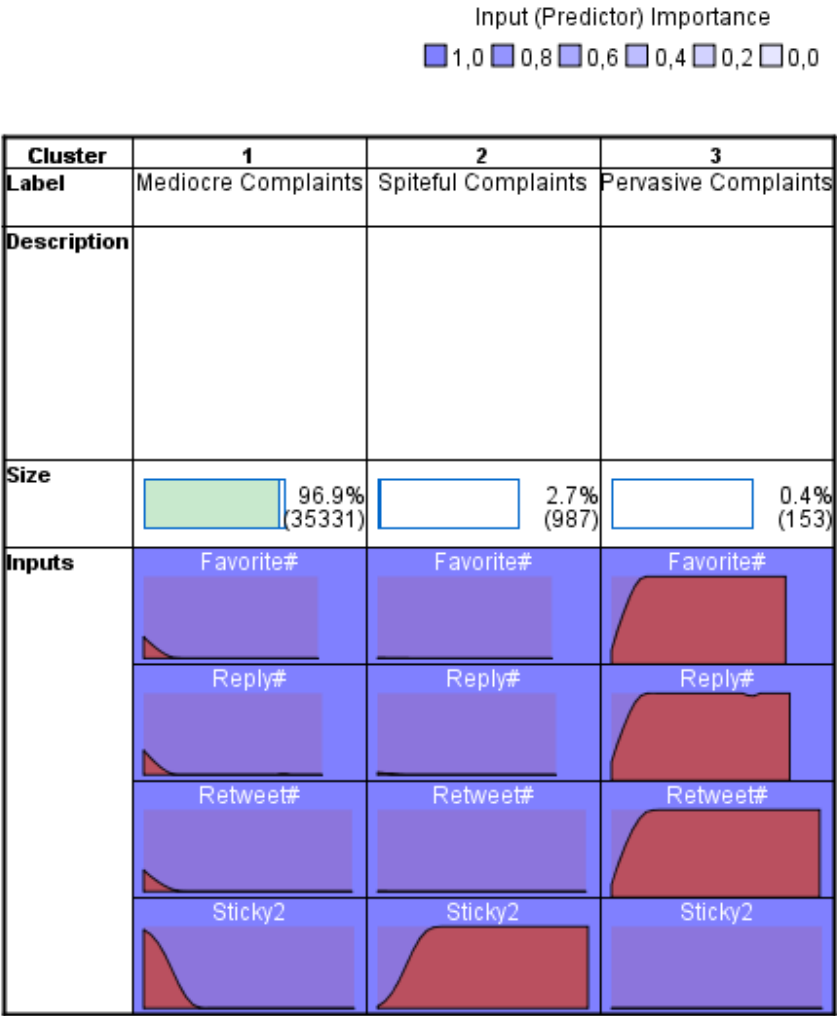
		N	% of Combined	% of Total
Cluster	1	35331	96,9%	96,9%
	2	987	2,7%	2,7%
	3	153	0,4%	0,4%
	Combined	36471	100,0%	100,0%
Total		36471		100,0%

Figure 9: Cluster Quality



Cluster analysis resulted in three complaint categories that significantly differ in interactional characteristics. Complaints belonging to the second cluster get almost no interaction in terms of favorites, replies, and retweets. However, complainants write numerous consecutive tweets about their negative experiences. Therefore, Cluster 2 is named as “spiteful complaints.” The third cluster consists of complaints that get a massive number of likes, favorites, and retweets. However, the number of consecutive complaint tweets sent by complainant mainly does not exceed one tweet. Following the high interaction rate, Cluster 3 is named as “pervasive complaints.” Complaints in the first cluster both have a low interaction rate, and the count of consecutive complaint tweets sent by the complainant is lesser when compared with the other clusters. Accordingly, Cluster 1 is named as “mediocre complaints.” Figure 10 depicts the size and input-related characteristics of the clusters, and Table 15 provides information regarding the means and standard deviations of inputs per cluster.

Figure 10: Clusters



Note: Favorite# stands for the count of favorites, Reply# for the count of replies, Retweet# for the count of retweets, and Sticky2 for the count of consecutive tweets

Table 15: Centroids

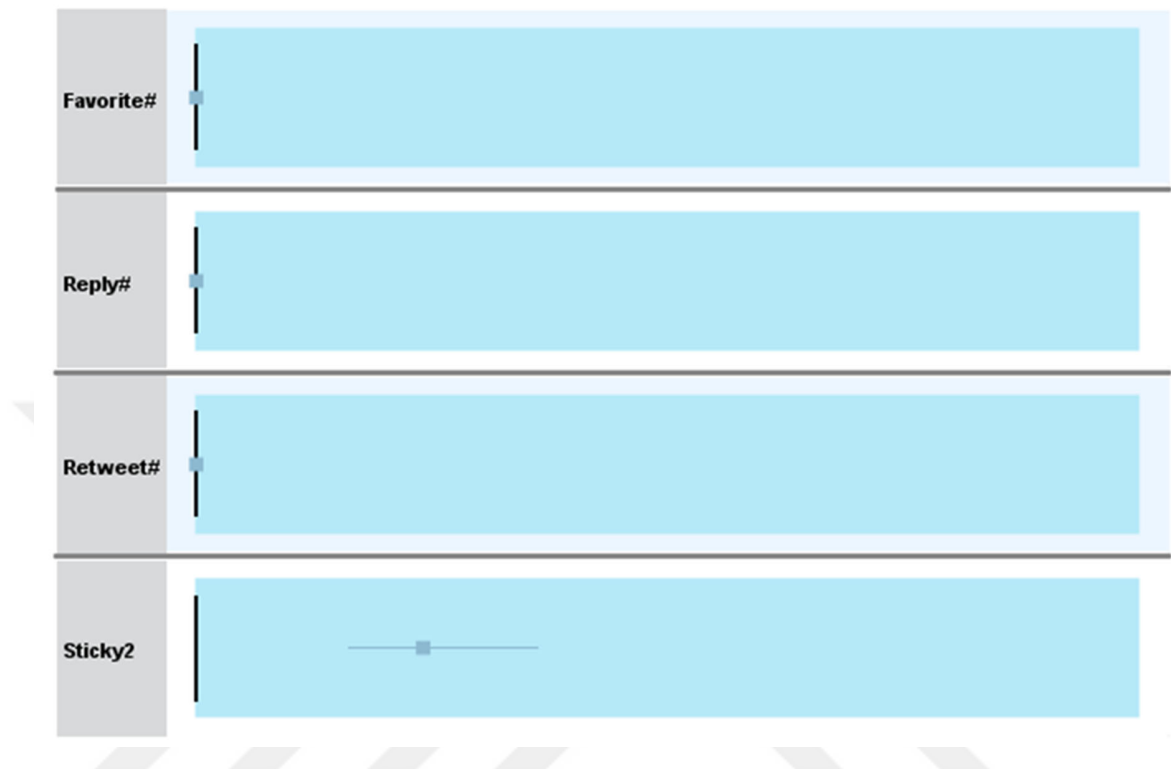
		Count of Consecutive Tweets		Count of Retweets		Count of Replies		Count of Favorites	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Cluster	1	,24	,578	,356	3,267	1,49	2,088	2,219	16,454
	2	8,08	5,210	1,035	6,765	1,94	5,350	3,923	29,204
	3	,45	,819	247,094	472,771	172,18	234,794	928,948	2458,533
	Combined	,45	1,636	1,409	34,604	2,21	18,880	6,153	170,483

4.3.2.2.2.1. Spiteful Complaints

Spiteful complaints constitute 2,7% of the sample with 987 cases. These complaints have significantly lower mean scores on the count of replies ($M=1,94$), favorites ($M=3,923$), and retweets ($M=1,035$) when compared to other complaint clusters. However, the mean score regarding the count of consecutive tweets ($M=8,08$) is significantly higher than the remaining two clusters. Figure 11 depicts the comparison of spiteful complaints cluster with other clusters in terms of the consecutive tweet, reply, favorite, and retweet counts.

Spiteful complaints score low in interaction and get a lower count of favorites, replies, and retweets than other complaints. However, what significantly distinct these complaints from others are that complainants post numerous consecutive and linked tweets about their negative experience and dissatisfaction. Some of these are a repetition of the main complaint and are being posted to attract the attention of either the service provider or the audience, whereas some are linked tweets that present a detailed description of the failure and/or dissatisfaction with the complaint handling process. For service providers that actively monitor complaints, these types of complaints incapacitate social listening by holding numerous records for a single complaint.

Figure 11: Comparison of Spiteful Complaints Cluster with Other Clusters



Note: Favorite# stands for the count of favorites, Reply# for the count of replies, Retweet# for the count of retweets, and Sticky2 for the count of consecutive tweets

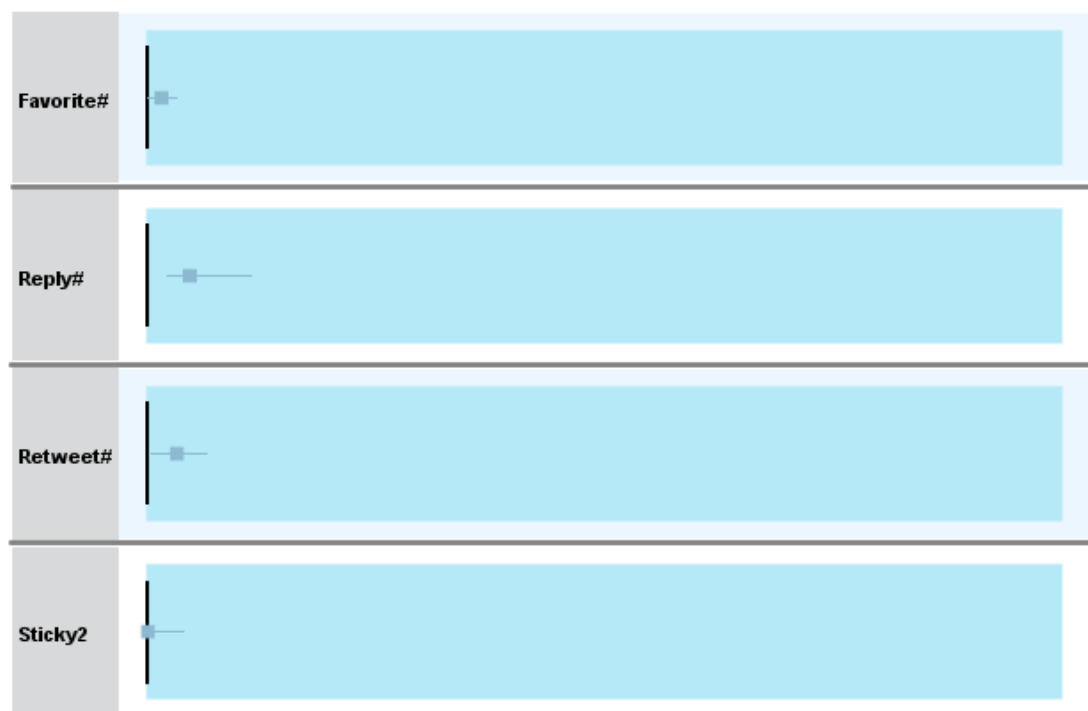
4.3.2.2.2. Pervasive Complaints

The pervasive complaints cluster is the smallest cluster that constitutes 0,4% of the sample with 153 cases. These complaints have significantly higher mean scores on the count of replies ($M=928,948$), favorites ($M=172,180$), and retweets ($M=247,094$) when compared to other complaint clusters. However, the mean score regarding the count of consecutive tweets ($M=0,45$) is significantly lower than the remaining two clusters. Figure 12 depicts the comparison of the pervasive complaints cluster with other clusters in terms of the consecutive tweet, reply, favorite, and retweet counts.

Even though pervasive complaints constitute the minority of the complaints articulated in social media, due to the enormous number of social media users they reach via a high volume of interaction, they may be considered as the most pernicious ones.

High interaction rates, the participation of other social media users through replies, retweets, and favorites lead to an exponential increase in audiences' number. Such a sharp increase may trigger further consumer movements and may result in an online firestorm.

Figure 12: Comparison of Pervasive Complaints Cluster with Other Clusters



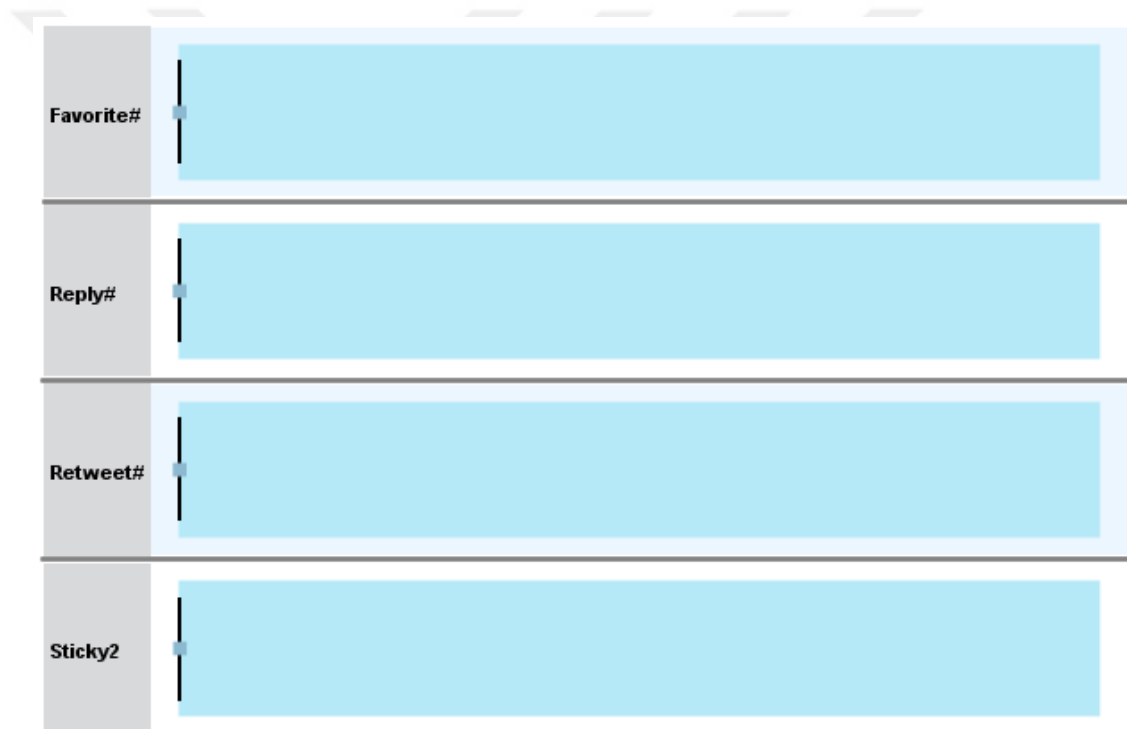
Note: Favorite# stands for the count of favorites, Reply# for the count of replies, Retweet# for the count of retweets, and Sticky2 for the count of consecutive tweets

4.3.2.2.2.3. Mediocre Complaints

Mediocre complaints cluster is the largest cluster that constitutes 96,9% of the sample with 35,331 cases. These complaints have significantly lower mean scores on the count of replies ($M=1,49$), favorites ($M=172,18$), retweets ($M=0,356$), and consecutive tweets ($M=0,24$) when compared to other complaint clusters. Figure 13 depicts the comparison of mediocre complaints cluster with other clusters in terms of the consecutive tweet, reply, favorite, and retweet counts.

Even though these complaints have the lowest interaction rate, the cluster's volume draws attention to the companies' risk of overlooking riskier complaints that contain the potential of triggering an online firestorm during social listening. Moreover, the cluster's size points that managing mediocre complaints is costlier for companies due to the required number of representatives. However, for companies that attach great importance to customer satisfaction rate, efficient handling of mediocre complaints may play a vital role.

Figure 13: Comparison of Mediocre Complaints Cluster with Other Clusters



Note: Favorite# stands for the count of favorites, Reply# for the count of replies, Retweet# for the count of retweets, and Sticky2 for the count of consecutive tweets

4.3.3. Two-Step Machine Learning Approach

As mentioned previously, two-step machine learning is employed in this thesis. The first step includes determining complainants' personality traits and complaint tweets' content features, including sentiment polarity and emotion, through text analytics. In the

second step, the relationships between the previously defined clusters and the independent variables are examined through multinomial logistic regression that is a machine learning algorithm used to evaluate the probability of categorical membership.

4.3.3.1. The First Leg of Two-Step Machine Learning Approach

In the first step of proposed approach, sentiment polarity of the complaint tweets, and emotions and personality traits of complainants are obtained via diverse text analytics tools. For revealing complainants' personality traits, their tweet history is analyzed by IBM Watson's Personality Insights. In order to obtain sentiment polarity and emotions, complaints are analyzed via ParallelDots' Sentiment API and Emotion Analysis AI API. Accordingly, this section provides a detailed explanation regarding the mentioned analyses.

4.3.3.1.1. Text Analytics

The rise of Web 2.0 technologies and intensified usage of social media provide diverse advantages to companies. Organizations utilize social media platforms and social media big data to support their business-related functions. These platforms provide high-speed access to knowledge, enhanced identification of experts, and condensed communication and operational costs to companies (Bughin and Chui 2010; Chau and Xu 2012). However, due to these platforms' nature, useful information exists in unstructured form and requires processing before being used in supporting business-related functions. Thus, text analytics tools are being employed for deriving sense-making information from social media data.

Text analytics refers to parsing text and extracting information about key entities through software tools (Russom, 2007). Information extraction involves natural language processing procedures for converting natural language into a form that can be utilized by various other computer programs. Text analytics includes three consecutive steps: text preprocessing, text representation, and knowledge discovery. Hu and Liu (2012) explain these steps as follows:

Text preprocessing: This stage includes stop word removal and stemming. In the stop word removal phase, words such as “as a, an, the, in” are eliminated. On the other hand, in the stemming phase, inflected words are reduced in their base form. For instance, “write,” “writing,” “wrote” are all represented as “write” for ensuring that the same words in variant forms can be regarded as the same feature.

Text representation: In this step, a text corpus is transformed into sparse numeric vectors (either through the bag of words approach or vector space model) to be used in linear algebraic operations.

Knowledge discovery: Machine learning or data mining methods are applied to transformed corpus for deriving knowledge.

As of the first leg of our two-step machine learning approach, complainants’ personality traits, emotion, and the sentiment polarity of complaint tweet is obtained through various tools for text analytics. Details regarding employed tools and processes are explained in the following sections.

4.3.3.1.1.1. Personality Traits

Prediction of personality traits of individuals has long been sought by scientists researching diverse areas. However, the rise of digital technologies and Web 2.0 applications carried the issue to another dimension. Text analytics tools and machine learning applications enabled deriving individuals’ personalities from their digital communications. Accordingly, various studies (e.g., Golbeck et al., 2011; Gou et al., 2014; Sumner et al., 2012) employed these tools for assessing social media users’ personalities and obtained noteworthy results.

In this thesis, IBM Watson’s Personality Insights is employed for conducting personality assessment through Twitter messages. The Personality Insights API uses a cloud service that depends on the psychometric survey-based score and employs vector representation of words obtained from large corpora established by Stanford University. The tool infers personality-related information from digital communications via machine learning techniques and provides numerical values ranging from 0.0 to 1.0 for each

characteristic. Personality Insight provides three diverse outcomes; assessment of the Big Five personality traits, needs, and values. Per our research aims, only the results obtained for the Big Five personality traits are considered.

Table 16 presents a sample of results from IBM Watson’s Personality Insights tool. It should be noted that personality traits are not derived from the complaints indicated in the “Tweet” column; instead, they are derived from individuals’ extant tweet history, complaints of whom are indicated in the “Tweet” column. Since providing the complete tweet history of complainants is infeasible, only the complaint tweets are indicated in the following table.



Table 16: A Sample of Results from IBM Watson's Personality Insights

Agreeableness	Conscientiousness	Extraversion	Neuroticism	Openness to experience	Tweet
1,00	0,86	1,00	0,01	0,18	@walmarthelp @Walmart can you folks Please Tell someone to Fill the @CocaCola machine at the Lakewood NY store Thanks Really wanted a Coke
1,00	0,82	0,95	0,16	0,52	@HomeDepot sad to have to leave and go to a competitor because of no customer service. Guess I need to be dressed as a guy to get any help!!
1,00	0,82	1,00	0,01	0,67	Well played #Walgreens with your Christmas Harry Potter section! I got everything but the mug (cause it was broken) I walked in for something else and walked...
0,84	0,24	1,00	0,99	0,23	I love when @FedEx says they attempted to make a delivery and no one actually comes
0,28	0,22	0,35	1,00	0,56	Waiting for two day for @UPS to deliver a package. The tracking tells me I wasn't available when the driver came. I am home, literally looking out of the window the whole day. Tried to call the number they provide online, disconnected. #upssucks

4.3.3.1.1.2. Emotion

Emotions mirror individuals' mental state of readiness resulting from cognitive evaluations regarding events or thoughts (Bagozzi et al., 1999). Following service failures, emotions mainly reflect the consequences of customers' stressful cognitive assessment (Lazarus, 1991). Some consumers articulate the consequences of this cognitive process on social media, and the complaints posted on these platforms reveal complainants' emotions either implicitly or explicitly. For deriving complainants' emotions from their digital communications, Natural Language Processing and machine learning tools yield fruitful outcomes, especially when the emotions are implicit in complainants' message.

Since sentiment polarity only reveals the tone of complaint regarding being positive, negative, and neutral, comprehending the underlying tone of messages plays a vital role for companies. Even though the extant literature focuses on negative emotions (Williams and Aaker, 2002) due to the negative valence of failure experience (Bougie et al., 2003), examining positive emotions also plays a vital role in comprehending the issue by covering all aspects. Particularly in social media, a satisfactory and timely response from the target company may alter the complainants' negative emotions. Additionally, following the criticality of the failure and personality of the complainant may impact the way the consumer feels about the incident and articulates his/her complaint. Therefore, within the context of this thesis, both negative and positive emotions are examined.

In this thesis, ParallelDots Emotion Analysis AI API is used for revealing the emotions of complainants. ParallelDots employs Convolutional Neural Networks deep learning algorithm for classifying emotions and provides numerical values ranging from 0.0 to 1.0 for the following emotions: happiness, sadness, anger, fear, excitement, and boredom. Within this context, due to the inadequacy of sentiment classes, emotions are employed for obtaining a deeper understanding of the nuances regarding the underlying tone of the complaints. Table 17 presents a sample of results obtained from the ParallelDots Emotion Analysis AI API.

Table 17: A Sample of Results from ParallelDots Emotion Analysis AI API

Happiness	Sadness	Anger	Fear	Excitement	Boredom	Tweet
0,00	0,11	0,29	0,11	0,00	0,14	I think something is wrong. Every time I jar my body my spine hurts so bad especially where my head meets my spine.
0,01	0,10	0,27	0,11	0,00	0,05	This is why I hate buying produce at @Walmart. Half the stuff has mold, the rest has bugs. Ugh. Too bad I'm broke and have no choice. *sigh*pic.twitter.com/D354e8Gpzf
0,04	0,10	0,26	0,48	0,00	0,26	This is so bad @AmazonUK your thieving driver has stolen another parcel left by another courier. https://ring.com/share/6637109573505215486 ... #SHOCKING #kent #ChristmasPresent #angry #idiot Do love your #Ring door camera though
0,00	0,11	0,23	0,16	0,00	0,17	Fedex did me bad on my shipment bruh whoever deliver my shit catching a fade
0,06	0,12	0,22	0,04	0,00	0,01	oh my goodness!! A bad situation has been made worse by poor service!! @Disney should stop using you as it reflects badly on them!

4.3.3.1.1.3. Sentiment Polarity

Sentiment polarity reveals the semantic orientation of a text (Hatzivassiloglou and McKeown, 1997) and discloses the sentiment hidden in those texts in terms of being positive (favorable), negative (unfavorable), and neutral (no indicated opinion) (Agarwal et al., 2008). Understanding the tone of complaints is critical for companies since expressions in strongly negative tones may offend service providers, company profits, and brand image (Bitner et al., 2000). Moreover, detecting such expressions and conversations vital for preventing further negative comments and may result in an online firestorm.

Compared with emotion analysis, sentiment polarity provides shallower information about the complaint and only reveals the tone of the complaint in terms of being positive, negative, and neutral. However, in this thesis, not only emotions but also sentiment polarity of complaints is investigated on purpose. In practice, for detecting dissatisfied consumers, companies classify messages about their brand in accordance with their sentiment. Hence, before investigating the emotions of dissatisfied consumers, there is a requirement to detect negative messages about the company. Therefore, revealing the sentiment polarity of complaints and comprehending the interrelation between sentiment and other variables is vital for both theory and practice.

In this thesis, ParallelDots Sentiment API is used for revealing the sentiment polarity of complaints. ParallelDots reveals the sentiment of online conversations and provides numerical values ranging from 0.0 to 1.0. Values between 0.0 and 0.5 indicate negative sentiment, whereas values between 0.6 and 1.0 indicate relatively positive sentiment. Table 18 presents a sample of results obtained from the ParallelDots Sentiment API.

Table 18: A Sample of Results from ParallelDots Sentiment API

Sentiment Score	Polarity	Tweet
0,01	Strong Negative	@AmazonHelp the new rule of having usps deliver my packages is trash. I don't get regular mail service and now my item is sitting at the post office until I can make a way there. Why am I paying for prime to now deal with usps?
0,13	Negative	@amazon Amazon - rushing things a bit? Very sloppy packaging. We are also seeing a lot more plastic used by Amazon, which is disheartening... #packaging #shipping #quality #Christmas #holidays @JeffBezos #plastics #environment #ecology #sustainable #renewablepic.twitter.com/OL3yRmLq4Z
0,98	Strong Positive	@Walmart love the grocery pick up (Aisha was great!). Just a suggestion, installing wi-fi accessable from spots while we wait???
0,00	Strong Negative	Just bought the WORST freaking fan from @Walmart . Loud despite being a comfort silent fan, doesn't blow further than 2 feet. Waste of \$\$.
0,03	Strong Negative	@Walgreens quickly approaching the point where the convenience isn't even close to being worth the ongoing hassle & ensuing frustration

4.3.3.2. The Second Leg of Two-Step Machine Learning Approach

Multinomial logistic regression was performed to assess how well the employed independent variables predict the dependent variable, which is complaint type. The model employs sentiment, personality traits, emotions, count of mentions, hashtags, followers, and followed accounts as independent variables, and contains a categorical independent variable (content-based and interaction-based complaint type). Estimations with multinomial logistic regression are conducted separately for both the entire sample and holdout samples. The decision of conducting analyses both for the entire sample and a holdout sample depends on two facts: (1) In interaction-based clusters, the ratio of pervasive and spiteful complaints with respect to mediocre ones is relatively small; and (2) In content-based categories, the ratio of each category (legal threat, call for boycott, brand switch warning) with respect to usual complaints is small. Hence, the estimated parameters are verified with holdout sampling, where all content-based categories are equalized in terms of the number of observations via random selection.

4.3.3.2.1. Multinomial Logistic Regression

Multinomial logistic regression, also known by numerous other names (polytomous logistic regression, conditional maximum entropy model, multiclass logistic regression, multiple entropy classifier, multinomial logit, and softmax regression), is a form regression used for predicting categorical membership on a dependent variable based on multiple independent variables (Starweather and Moske, 2011). This predictive analysis is mainly employed when the outcome variable is binary or dichotomous, and independent variables are continuous, categorical, or both (Wang, 2005).

Multinomial logistic regression extends the binary logistic regression to classification problems where the dependent variable can more than two possible values, and similarly, with the binary logistic regression, multinomial logistic regression employs maximum likelihood estimation for evaluating categorical membership probability (Ramadhan et al., 2017; Starweather and Moske, 2011). The main difference between the binary logistic regression model and the multinomial logistic regression model is the fact that the dependent variable in the former has only a binary choice, and in the latter

dependent variable in the latter can have more than two choices that are coded categorically (Wang, 2005).

Tabachnick et al. (2001) argue that multinomial logistic regression has numerous advantages which makes it popular amongst other models: (1) it is more robust to assumption violations of multivariate normality and equal variance-covariance matrices across groups, (2) when compared with linear regression, it provides diagnostic statistics that can be interpreted more easily, (3) and it does not assume a linear relationship between dependent and independent variables, and does not require normality, linearity, or homoscedasticity of data, (4) it does not necessitate independent variables to be an interval, (5) it does not require independent variables to be unbounded and lastly, (6) error terms that are normally distributed are not assumed in multinomial logistic regression.

Multinomial logistic regression, a supervised learning algorithm used in classification problems, is employed for two main reasons. First, multinomial logistic regression is a predictive model capable of conducting regression analysis when the dependent variable is categorical. Since the proposed research model contains a categorical dependent variable (complaint type) with more than two unordered levels, multinomial logistic regression is congruent with the proposed model. Second, during multinomial logistic regression, one level of the dependent variable is selected as the reference category, and the probability of being in any of the other levels is compared with the probability of being in the reference category. The model, then, estimates separate binary logistic regression for each level of the dependent variable, and as a result of conducting separate binary logistic regressions, independent variables' impact on each category can be observed separately. Accordingly, the model can determine the differential characteristics of each complaint type and to provide information regarding the relative importance of explanatory variables for each complaint type.

Multinomial logit models for nominal responses use separate binary logit model for each pair of response categories. Suppose, J is the number of categories Y . Let $P_1, \dots, P_J$ be the success probabilities which satisfy $\sum_j P_j = 1$. The probability distribution for this number of outcomes of the J types is multinomial. In the model of this kind, the order of

listing the categories is inappropriate, since the model treats the response scale as nominal. Furthermore, logit models for nominal response variable pair each category with a baseline category. Suppose the J is the last category, the baseline category logits are:

$$\log \frac{P_j}{P_J} = \alpha_j + \beta_j x \quad j = 1, \dots, J-1 \quad (1)$$

Therefore, the model has $J-1$ equations.

For further illustration, when $J = 3$, for instance, the model uses:

$$\log \frac{P_1}{P_3} = \alpha_1 + \beta_1 x \quad (2)$$

$$\log \frac{P_2}{P_3} = \alpha_2 + \beta_2 x \quad (3)$$

The equation for other arbitrary pair of categories of 2 and 3 can be constructed:

$$\begin{aligned} \log \frac{P_1}{P_2} &= \log \frac{P_1/P_3}{P_2/P_3} = \log \frac{P_1}{P_3} - \log \frac{P_2}{P_3} = \\ &(\alpha_1 + \beta_1 x) - (\alpha_2 + \beta_2 x) = (\alpha_1 - \alpha_2) + (\beta_1 - \beta_2) x \end{aligned} \quad (4)$$

On the other hand, success probabilities of categories of response variable can be estimated by

$$P_j = \frac{e^{\alpha_j + \beta_j x}}{\sum e^{\alpha_j + \beta_j x}}, j = 1, \dots, J \quad (5)$$

The denominator is the same for each probability, and the numerator varies according to different categories.

When the response categories are ordered, the logit considers the order by using cumulative logits. Suppose Y is response variable, the cumulative probability of response variable is the probability that Y falls at or below a specific point. Therefore, the cumulative probability is:

$$P(Y \leq j) = P_1 + \dots + P_j, j = 1, \dots, J \quad (6)$$

In addition to that the logit of the cumulative probability is:

$$\text{logit}[P(Y \leq j)] = \log \left[\frac{P(Y \leq j)}{1 - P(Y \leq j)} \right], j = 1, \dots, J-1 \quad (7)$$

Suppose that $J = 3$, the logit of the cumulative probability of $Y \leq 1$ is:

$$\text{logit } P(Y \leq 1) = \log \left[\frac{P_1}{P_2 + P_3} \right] \text{ and } \text{logit } P(Y \leq 2) = \log \left[\frac{P_1 + P_2}{P_3} \right] \quad (8)$$

For the explanatory variable x the logit model is:

$$\text{logit}[P(Y \leq j)] = \alpha_j + \beta x \quad j = 1, \dots, J-1 \quad (9)$$

The parameter β in the model describes the effects of x on the log odds of the response in category j or below. The model assumes that the effect of x on the log odds of the response categories is identical for all $J-1$ cumulative logits, in other words the coefficients are identical. Hence it assumes the $J-1$ logits have different intercept increasing in j .

The explanatory variable in cumulative logit models can be quantitative, categorical or both. When the categories are reversed in order, the result of the model fit is the same but the sign of $\hat{\beta}$ reverse.

Model interpretation can use odds ratios for the cumulative probabilities. The odds ratio comparing the cumulative probabilities for two values of x_1 and x_2 of x is:

$$\frac{P(Y \leq j | X = x_2) / P(Y > j | X = x_2)}{P(Y \leq j | X = x_1) / P(Y > j | X = x_1)} \quad (10)$$

The log of this odds ratio is the difference between the cumulative logits at these two values of X , which is proportional to the distance between the x values. For example, when difference between the two x values is 1, the odds of response at any category is multiply by e^β for each unit increase in x . This log odds ratio can be written $\beta(x_2 - x_1)$ in which the same proportionality constant (β) applies in each cumulative probability. This property is called proportional odds assumption. Proportional odds models are fit using maximum likelihood estimation.

An alternative logits for ordered categories use pairs of categories instead of using the entire response scale in forming each logit for cumulative logit models. One approach forms logits for all pairs of adjacent categories. The adjacent categories logits are the $j-1$ logits formed by the log odds of the probability of response falling in category j given that the response falls within either category j .

$$\log\left(\frac{P_{j+1}}{P_j}\right) = \alpha_j + \beta_j x \quad j = 1, \dots, J-1 \quad (11)$$

Suppose $J = 3$, the adjacent categories logit model has form:

$$\log\left(\frac{P_2}{P_1}\right) = \alpha_1 + \beta_1 x \text{ and } \log\left(\frac{P_3}{P_2}\right) = \alpha_2 + \beta_2 x \quad (12)$$

Like the baseline category logits, the adjacent categories logits resolve the logit of all pairs of response categories. The adjacent categories logit models differ from cumulative logit models in that effect of explanatory variables refer to individual response categories and not cumulative group.

Continuation-ratio logits is another approach that forms logits for ordered response categories in a chronological manner. The continuation ratio logit model is:

$$\log\left[\frac{P_1}{P_2}\right], \log\left[\frac{P_1 + P_2}{P_3}\right], \dots, \log\left[\frac{P_1 + \dots + P_{J-1}}{P_J}\right] \quad (13)$$

These are logits of conditional probabilities that the response falls in the j^{th} category for $j=1, \dots, J-1$. The likelihood is the product of the multinomial mass functions.

The pseudo code of the employed algorithm can be described as follows:

Input: Training samples: $x_1, x_2 \dots x_n$, Class labels: $y_1, y_2 \dots y_n$

Initialize: Set $k = 0$. Initialize $\tilde{w}^0$

Pre-compute: $v_i = \sum_{j=1}^n y_{ij} x_j, i \in (1, 2 \dots m)$.

Repeat:

1) Compute: $\alpha_j = \frac{1}{\sum_{i=1}^m \exp\left((w_i^k)^T x_j\right)}, j \in (1, 2 \dots n)$

2) Compute the following parallely over all the elements of $\tilde{w}$:

$$g'(w_{il} | \tilde{w}^k) = -v_{il} + \sum_{j=1}^n a_j x_{jl} \frac{\exp(dx_{jl} w_{il})}{\exp(dx_{jl} w_{il}^k - x_j^T w_i^k)}$$

Compute the value of $g'(w_{il} | \tilde{w}^k)$ at $w_{il} = 0$.

Choose the value of b based on whether the value of $g'(w_{il} | \tilde{w}^k)$ at $w_{il} = 0$ is greater than or lesser than zero.

w_{il}^{k+1} is obtained by solving:

$$\underset{w_{il}}{\text{minimize}} - \sum_{i=1}^m \sum_{l=1}^d w_{il} v_{il} + \sum_{j=1}^n a_j \sum_{i=1}^m \sum_{l=1}^d \frac{1}{d} \left(\frac{\exp(dx_{jl} w_{il})}{\exp(dx_{jl} w_{il}^k - x_j^T w_i^k)} \right)$$

using bisection method with $a = 0$ and updated b .

3) $k \leftarrow k + 1$, **until** $\left| \frac{l_{MLR}(\tilde{w}^k) - l_{MLR}(\tilde{w}^{k-1})}{l_{MLR}(\tilde{w}^{k-1})} \right| \leq 10^{-3}$

4.3.3.2.1.1. Results for Content-Based Categories

4.3.3.2.1.1.1. Full Sample Results for Content-Based Categories

As explained previously, three types of complaints (complaints containing the legal threat, call for boycott, and switching intention) were revealed during content-based classification. During multinomial logistic regression, instead of setting the largest cluster as the reference category, a fourth category named “usual complaints” is employed as the reference category for two main reasons. First, according to the nature of this thesis' context, cluster memberships are stable, and a transition between complaint type clusters is not possible. Since what is said is said and what is posted is posted, there is no possibility for a posted complaint tweet containing a legal threat to become a complaint tweet denoting switching intention. In this regard, assessing the probability of a complaint containing legal threat being in the cluster consisting of complaints containing a call for a boycott or switching intention is not required. Second, instead of comparing the probability of belonging to one of the content-based categories with the probability of being in the reference category, the primary motivation is to examine whether the employed independent variables can reveal what makes a regular complaint tweet to fall into one of the content-based categories. Hence, a reference category other than existing content-based categories is employed.

The likelihood ratio test results, shown in Table 19, indicate that the model is statistically significant (Sig. = $p < 0.0001$) and denote that employed independent variables can predict the type of complaint with statistical significance. The final model is statistically significant, $\chi^2 = 493.655$ with 48 degrees of freedom ($p < 0.0001$), and hence it may be stated that the model can significantly distinguish the type of complaint based on the indicated independent variables. Pseudo R-square results indicate that the final model explains 0.075 (Nagelkerke R^2) of the variances in content-based complaint types.

Table 19: Model Fitting Information for Full Sample Results of Content-Based Categories

Independent variables	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept only	7221.047			
Final	6727.392	493,655	48	.000

The results, shown in Table 20, revealed that each complaint cluster retains a diverse set of explanatory variables. It is found that complaints containing legal threat are characterized by boredom, sadness, excitement, agreeableness, conscientiousness, and the count of mentioned accounts. Those containing a call for boycott are characterized by the count of mentioned accounts, sadness, boredom, and sentiment polarity of the complaint. Complaints that denote switching intention are found to be characterized by sadness, anger, the count of mentioned accounts and hashtags used, sentiment polarity, agreeableness, neuroticism, and openness to the experience of the complainant. It should also be noted that the count of used followers and accounts followed, happiness, fear, and extraversion of the complainant are found to have no significant distinguishing impact on any of the content-based categories.

Boredom and conscientiousness of the complainant are found to have a statistically significant positive impact on distinguishing complaints that contain a legal threat from the usual ones. Boredom has a significant positive impact with the logit coefficient of 12.374 and an exponential beta of 131.641 ($\beta=12.374$; $p<.001$; $\text{Exp } (\beta)=131.641$). This suggests that other things being equal, the odds of a posted complaint to contain legal threat would be 131.641 times more than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain legal threat increases by 99.9% with an increase in the indicated boredom. The conscientiousness of the complainant also has a positive impact with the logit coefficient of 0.869 and an exponential beta of 2.385 ($\beta=0.869$; $p<.001$; $\text{Exp } (\beta)=2.385$), suggesting that other things being equal, the odds of a posted complaint to contain legal threat would be 2.385 times more than the odds of a complaint belonging to the usual complaints

cluster, and the probability of a posted complaint to contain legal threat increases by 99.9% with an increase in the conscientiousness level of the complainant. On the other hand, the count of mentioned accounts, sadness, excitement, and agreeableness of the complainant is found to negatively impact distinguishing complaints that contain a legal threat from the usual ones. The count of mentioned accounts has a negative impact with the logit coefficient of -0.340 and an exponential beta of 0.711 ($\beta = -0.340$; $p < .001$; $\text{Exp}(\beta) = 0.711$), suggesting that other things being equal, the odds of a posted complaint to contain legal threat would be 0.711 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain legal threat decreases by 99.9% with an increase in the count of mentioned accounts. Excitement also has a negative impact with the logit coefficient of -7.740 and an exponential beta of 0 ($\beta = -7.740$; $p < .001$; $\text{Exp}(\beta) = 0$), suggesting that the probability of a posted complaint to contain legal threat decreases by 99.9% with an increase in the indicated excitement. In a similar vein, agreeableness also has a negative impact with the logit coefficient of -1.423 and an exponential beta of 0.241 ($\beta = -1.423$; $p < .001$; $\text{Exp}(\beta) = 0.241$), suggesting that other things being equal, the odds of a posted complaint to contain legal threat would be 0.241 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain legal threat decreases by 99.9% as the agreeableness of the complainant increases. Finally, sadness has a negative impact with the logit coefficient of -4.481 and an exponential beta of 0.011 ($\beta = -4.481$; $p < .05$; $\text{Exp}(\beta) = 0.011$), suggesting that other things being equal, the odds of a posted complaint to contain legal threat would be 0.011 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain legal threat decreases by 95% with an increase in the indicated sadness.

The count of mentioned accounts and sadness are found to have a statistically significant positive impact on distinguishing complaints that contain a call for a boycott from the usual ones. The count of mentioned accounts has a positive impact with the logit coefficient of 0.162 and an exponential beta of 1.176 ($\beta = 0.162$; $p < .05$; $\text{Exp}(\beta) = 1.176$), suggesting that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 1.176 times more than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain a call for

boycott increases by 95% as the count of mentioned accounts increases. Sadness also has a positive impact with the logit coefficient of 10.172 and an exponential beta of 26158.755 ($\beta=10.172$; $p<.001$; $\text{Exp}(\beta)= 26158.755$), suggesting that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 26158.755 times more than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain a call for boycott increases by 99.9% as the indicated sadness increases. On the other hand, sentiment and boredom are found to have a negative impact on distinguishing complaints that contain a call for a boycott from the usual ones. Sentiment has a negative impact with the logit coefficient of -2.194, and an exponential beta of 0.111 ($\beta=-2.194$; $p<.001$; $\text{Exp}(\beta)= 0.111$), suggesting that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 0.111 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain a call for boycott decreases by 99.9% as the sentiment of the complaint converges to the positive polarity. Finally, boredom has a negative impact with the logit coefficient of -10.919 and an exponential beta of 1.811E-5 ($\beta=-10.919$; $p<.001$; $\text{Exp}(\beta)= 1.811\text{E}-5$), suggesting that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 1.811E-5 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain a call for boycott decreases by 99.9% as the indicated boredom increases.

Sadness is found to have a statistically significant positive impact on distinguishing complaints that contain brand switch warning from the usual complaints. Sadness has a positive impact with the logit coefficient of 6.131 and an exponential beta of 459.752 ($\beta=6.131$; $p<.001$; $\text{Exp}(\beta)= 459.752$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 459.752 times more than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning increases by 99.9% as the indicated sadness increases. On the other hand, the count of mentioned accounts and hashtags used, sentiment, anger, agreeableness, neuroticism, and openness to experience of a complainant are found to have a negative impact on distinguishing complaint that contains brand switch warning from the usual complaints. The count of mentioned accounts has a negative impact with the logit coefficient of -0.171 and an

exponential beta of 0.462 ($\beta=-0.171$; $p<.001$; $\text{Exp}(\beta)= 0.462$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 0.462 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 99.9% as the count of mentioned accounts increases. Similarly, the count of hashtags used has a negative impact with the logit coefficient of -0.312 and an exponential beta of 0.462 ($\beta=-0.312$; $p<.05$; $\text{Exp}(\beta)= 0.732$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 0.732 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 95% as the count of hashtags used increases. Sentiment also has a negative impact with the logit coefficient of -0.664 and an exponential beta of 0.515 ($\beta=-0.664$; $p<.001$; $\text{Exp}(\beta)= 0.515$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 0.515 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 99.9% as the sentiment of the complaint converges to positive polarity. Anger has a negative impact with the logit coefficient of -12.635 and an exponential beta of 3.257E-6 ($\beta=-12.635$; $p<.05$; $\text{Exp}(\beta)= 3.257\text{E}-6$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 3.257E-6 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 95% as the indicated anger increases. In a similar vein, agreeableness has a negative impact with the logit coefficient of -0.674 and an exponential beta of 0.509 ($\beta=-0.674$; $p<.05$; $\text{Exp}(\beta)= 0.509$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 0.509 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 95% as agreeableness of complainant increases. Neuroticism also has a negative impact with the logit coefficient of -0.587 and an exponential beta of 0.556 ($\beta=-0.587$; $p<.05$; $\text{Exp}(\beta)= 0.556$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 0.556 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 95% as neuroticism of complainant

increases. Finally, openness to experience has a negative impact with the logit coefficient of -0.776 and an exponential beta of 0.465 ($\beta=-0.776$; $p<.001$; $\text{Exp}(\beta)= 0.465$), suggesting that other things being equal, the odds of a posted complaint to contain brand switch warning would be 0.465 times less than the odds of a complaint belonging to the usual complaints cluster, and the probability of a posted complaint to contain brand switch warning decreases by 99.9% as openness to experience of complainant increases.



Table 20: Full Observation Parameter Estimates for Content-Based Categories

Comparison	<i>Usual vs. Legal Threat</i>				Usual vs. Call for Boycott				Usual vs. Brand Switch Warning			
	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)
Intercept	(5.369)***	.372	208		(5.122)***	.290	312		(3.27)***	.295	122	
Count of mentions	(.340)***	.127	7.21	.711	.162**	.067	5.81	1.176	(.771)***	.122	40.0	.462
Count of hashtags	(.201)	.131	2.37	.818	.029	.073	.157	1.029	(.312)**	.138	5.08	.732
Sentiment	.434	.294	2.17	1.543	(2.194)***	.321	46.8	.111	(.664)***	.257	6.65	.515
Happiness	.667	2.29	.084	1.949	1.260	1.57	.638	3.525	(2.438)	1.96	1.54	.087
Sadness	(4.481)**	1.85	5.86	.011	10.172***	1.37	55.0	2615	6.131***	1.86	10.7	459.75
Anger	.527	2.22	.056	1.693	3.585	2.49	2.06	36.06	(12.63)**	497	6.44	0.0003
Fear	1.945	1.96	.979	6.996	(1.429)	3.04	.220	.240	(8.193)	5.15	2.53	.000
Excitement	(7.740)***	2.52	9.39	.000	.805	.842	.913	2.236	.326	.873	.139	1.385
Boredom	12.374***	1.07	131	2365	(10.91)***	2.31	22.3	0.001	(1.053)	2.20	.227	.349
Agreeableness	(1.423)***	.405	12.3	.241	(.154)	.295	.272	.857	(.674)**	.328	4.23	.509
Conscientiousness	.869***	.339	6.56	2.385	.328	.268	1.4	1.388	.046	.293	.025	1.047
Extraversion	(.001)	.351	.000	.999	.071	.277	.066	1.073	.237	.293	.652	1.267
Neuroticism	(.132)	.290	.207	.877	(.178)	.233	.583	.837	(.587)**	.242	5.89	.556
Openness to Exp.	.015	.326	.002	1.015	.208	.262	.629	1.231	(.766)***	.270	8.05	.465
Following	.000	.000	.065	1.000	.000	.000	.627	1.000	.000	.000	.724	1.000
Followers	.000	.000	.149	1.000	.000	.000	.234	1.000	.000	.000	.087	1.000

Note: *B* – Estimated multinomial logistic regression coefficients; Std. Error – Standard errors of the individual regression coefficients; Wald – Wald chi-square test that examines the null hypothesis that the estimate equals 0; Exp (*B*) – An exponential beta; ****p*<0.01, ***p*<0.05, **p*<0.1

4.3.3.2.1.1.2. Holdout Sample Results for Content-Based Categories

The likelihood ratio test results, shown in Table 21, indicate that the model is statistically significant (Sig. = $p < 0.0001$) and denote that employed independent variables can predict the type of complaint with statistical significance. The final model is statistically significant, $\chi^2 = 281.039$ with 48 degrees of freedom ($p < 0.0001$), and hence it may be stated that the model can significantly distinguish the type of complaint based on the five indicated independent variables. Pseudo R-square results indicate that the final model explains 0.405 (Nagelkerke R^2) of the variances in content-based complaint types.

Table 21: Model Fitting Information for Holdout Sample Results of Content-Based Categories

Independent variables	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept only	1633.036			
Final	1351.997	281.039	48	.000

The results, shown in Table 22, revealed that each complaint cluster retains a diverse set of explanatory variables. It is found that complaints containing legal threat are characterized by boredom, fear, anger, excitement, count of mentioned accounts, and agreeableness of complainant. In contrast, those containing a boycott call are characterized by sadness, negative sentiment, and neuroticism of the complainant. Complaints that denote switching intention are characterized by sadness, anger, and counts of mentioned accounts. Boredom is highly, whereas fear, anger, excitement, count of mentioned accounts, and agreeableness of the complainant are slightly significant in distinguishing complaints from the legal threat cluster. The sentiment is high, whereas the complainant's sadness and neuroticism are slightly significant in distinguishing complaints belonging to the call for boycott cluster. Finally, the count of mentioned

accounts is high, whereas the complainant's sadness and anger are slightly significant in distinguishing complaints belonging to the brand switch cluster. It should also be noted that the count of used hashtags, followers, and accounts followed, happiness, and personality traits of conscientiousness, extraversion, and openness to experience are found to have no significant distinguishing impact on any of the content-based categories.

As mentioned, boredom has a highly significant positive impact on distinguishing legal threat cluster. The logit coefficient for boredom is 18.548, and its exponential beta is 113558044.400 ($\beta=18.548$; $p<.001$; $\text{Exp } (\beta)= 113558044.400$). This suggests that other things being equal, the odds of a posted complaint to contain a legal threat would be 113558044.400 times more than the odds of a complaint belonging to the usual complaints cluster. Thus, the probability of a posted complaint to contain legal threat increases by 99.9% with an increase in the indicated boredom. Another factor having a significant positive impact on distinguishing legal threat cluster is fear. The logit coefficient for fear is 14.245, and its exponential beta is 1537149.783 ($\beta=14.245$; $p<.1$; $\text{Exp } (\beta)= 1537149.783$), suggesting that other things being equal, the odds of a posted complaint to contain legal threat would be 1537149.783 times more than the odds of a complaint belonging to the usual complaints cluster. Thus, the probability of a posted complaint to contain a legal threat increases by 90% with an increase in the indicated fear. Contrary to boredom and fear, anger, excitement, the agreeableness of the complainant, and the count of mentioned accounts negatively impact distinguishing the legal threat cluster. The logit coefficient for anger is -10.949, and its exponential beta is 1.757E-5 ($\beta=-10.949$; $p<.1$; $\text{Exp } (\beta)= 1.757\text{E}-5$), suggesting that other things being equal, the odds of a posted complaint to contain legal threat would be 1.757E-5 times less than the odds of a complaint belonging to the usual complaints cluster. Accordingly, the probability of a posted complaint to contain a legal threat decreases by 90%, with an increase in the indicated anger. For excitement, the logit coefficient is -7.235, and the exponential beta is 0.001 ($\beta=-7.235$; $p<.05$; $\text{Exp } (\beta)= .001$), indicating that, other things being equal, the odds of a posted complaint to contain legal threat would be 0.001 times less than the odds of a complaint belonging to the usual complaints cluster. Hence, the probability of a posted complaint to contain a legal threat decreases by 95%, with an increase in the indicated excitement. The logit coefficient for the count of mentioned accounts is -0.372, and the exponential beta is 0.689 ($\beta=-0.372$; $p<.05$; $\text{Exp } (\beta)= .689$), denoting that other

things being equal, the odds of a posted complaint to contain legal threat would be 0.689 times less than the odds of a complaint belonging to the usual complaints cluster. Therefore, the probability of a posted complaint to contain a legal threat decreases by 95%, with an increase in the count of mentioned accounts in the complaint tweet. Finally, for agreeableness, the logit coefficient is -1.623, and the exponential beta is 0.197 ($\beta = -1.623$; $p < .05$; $\text{Exp}(\beta) = .197$), indicating that, other things being equal, the odds of a posted complaint to contain legal threat would be 0.197 times less than the odds of a complaint belonging to the usual complaints cluster. Accordingly, the probability of a posted complaint to contain a legal threat decreases by 95% as the complainant's agreeableness increases.

Sadness has a highly significant positive impact on distinguishing call for boycott cluster with a logit coefficient of 7.033 and an exponential beta of 1133.878 ($\beta = 7.033$; $p < .05$; $\text{Exp}(\beta) = 1133.878$). This suggests that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 1133.878 times more than the odds of a complaint belonging to the usual complaints cluster. Thus, the probability of a posted complaint to contain a call for boycott increases by 95%, with an increase in the indicated sadness. Neuroticism and sentiment have a negative impact on distinguishing the call for boycott cluster. The logit coefficient for neuroticism is -0.776, and its exponential beta is 0.460 ($\beta = -0.776$; $p < .1$; $\text{Exp}(\beta) = 0.460$), suggesting that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 0.460 times less than the odds of a complaint belonging to the usual complaints cluster. Accordingly, the probability of a posted complaint to contain a call for boycott decreases by 90% as the complainant's neuroticism increases. Finally, for sentiment, the logit coefficient is -2.287, and the exponential beta is 0.102 ($\beta = -2.287$; $p < .001$; $\text{Exp}(\beta) = .102$), indicating that, other things being equal, the odds of a posted complaint to contain a call for a boycott would be 0.102 times less than the odds of a complaint belonging to the usual complaints cluster. Accordingly, the probability of a posted complaint to contain a call for boycott decreases by 99.9% as the positivity of the complaint tweet's sentiment increases.

Sadness has a significant positive impact on distinguishing brand switch cluster with a logit coefficient of 7.123 and an exponential beta of 1472845.156 ($\beta = 7.123$; $p < .05$; $\text{Exp}(\beta) = 1472845.156$). This suggests that other things being equal, the odds of a posted

complaint to contain a brand switch warning would be 1472845.156 times more than the odds of a complaint belonging to the usual complaints cluster. Hence, the probability of a posted complaint to contain brand switch warning increases by 95%, with an increase in the indicated sadness. The count of accounts mentioned in the complaint tweet and anger has a negative impact on distinguishing the brand switch cluster. The logit coefficient for the count of mentions is -0.401, and the exponential beta is 0.894 ($\beta = -0.401$; $p < .001$; $\text{Exp}(\beta) = 0.894$), suggesting that other things being equal, the odds of a posted complaint to contain a call for a boycott would be 0.894 times less than the odds of a complaint belonging to the usual complaints cluster. Accordingly, the probability of a posted complaint to contain brand switch warning decreases by 99.9% as the count of mentioned accounts in the complaint tweet increases.

Table 22: Holdout Sample Parameter Estimates for Content-Based Categories

Comparison	<i>Usual vs. Legal Threat</i>				Usual vs. Call for Boycott				Usual vs. Brand Switch Warning			
	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)
Intercept	(.272)	.590	.212		.096	.530	.033		1.311	.527	6.202	
Count of mentions	(.372)**	.172	4.668	.689	.021	.126	.027	1.021	(.401)***	.148	7.378	.894
Count of hashtags	(.305)	.201	2.298	.737	.081	.105	.603	1.085	(.258)	.163	2.493	1.064
Sentiment	.231	.484	.227	1259	(2.287)***	.507	20.338	.102	(.592)	.429	1.903	1.283
Happiness	4.506	3.613	1.555	90.575	1.854	3.185	.339	6.388	(.486)	3.205	.023	328
Sadness	2.959	3.890	.579	19.286	7.033**	3.428	4.209	1133	7.123**	3.612	3.889	14728
Anger	(10.949)*	6.352	2.972	0.0075	(3.715)	6.106	.370	.024	(18.38)**	7.622	5.817	.032
Fear	14.245*	7.907	3.246	15371	12.008	8.095	2.200	1640	2.339	9.362	.062	96623
Excitement	(7.235)**	3.047	5.640	.001	.421	1.543	.074	1.523	(.512)	1.429	.128	9.855
Boredom	18.548***	3.927	22.30	11355	(2.356)	4.268	.305	.095	.669	4.596	.021	15928
Agreeableness	(1.623)**	.671	5.841	.197	(.259)	.563	.211	.772	(.686)	.562	1.491	1.514
Conscientiousness	.088	.565	.024	1.092	.354	.485	.533	1.425	(.096)	.490	.038	2.373
Extraversion	.280	.586	.229	1.324	.224	.518	.187	1.251	.303	.511	.350	3.686
Neuroticism	(.290)	.481	.364	.748	(.776)*	.438	3.135	.460	(.837)*	.431	3.769	1.008
Openness to Exp.	.317	.529	.360	1.374	.225	.479	.220	1.252	(.492)	.468	1.102	1.532
Following	.000	.000	.918	1.000	.000	.000	1.418	1.000	.000	.000	.429	1.000
Followers	.000	.000	.245	1.000	.000	.000	.197	1.000	.000	.000	.297	1.000

Note: *B* – Estimated multinomial logistic regression coefficients; Std. Error – Standard errors of the individual regression coefficients; Wald – Wald chi-square test that examines the null hypothesis that the estimate equals 0; Exp (*B*) – An exponential beta; ****p*<0.01, ***p*<0.05, **p*<0.1

4.3.3.2.1.2. Results for Interaction-Based Clusters

4.3.3.2.1.2.1. Full Sample Results for Interaction-Based Clusters

Multinomial regression analysis for the full sample is conducted with 36471 observations. Among the previously revealed interaction-based clusters (spiteful, mediocre, and pervasive), mediocre complaints are selected as the reference category since it is the largest cluster. The likelihood ratio test results, shown in Table 23, indicate that the model is statistically significant (Sig. = $p < 0.0001$) and denote that employed independent variables can predict the type of complaint with statistical significance. The final model is statistically significant, $\chi^2 = 644.656$ with 32 degrees of freedom ($p < 0.0001$), and hence it may be stated that the model can significantly distinguish the type of complaint based on the indicated independent variables. Pseudo R-square results indicate that the final model explains 0.067 (Nagelkerke R^2) of the variance in content-based complaint types.

Table 23: Model Fitting Information for Full Sample Results of Interaction-Based Clusters

Independent variables	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept only	11044.338			
Final	10399.682	644.656	32	.000

The results, shown in Table 24, revealed that each complaint cluster retains a diverse set of explanatory variables. It is found that spiteful complaints differentiate from mediocre complaints in indicated anger, sentiment, count of hashtags used and followed accounts, and extraversion of the complainant. In contrast, pervasive complaints differentiate from mediocre complaints in indicated boredom, fear, sentiment, personality traits of conscientiousness and extraversion, and the count of mentioned accounts, followers, and followings. Indicated anger, sentiment, count of hashtags used and followed accounts, and extraversion of the complainant are highly significant in

distinguishing spiteful complaints from mediocre complaints. On the other hand, the count mentioned accounts, followers, followed accounts, sentiment, and conscientiousness of the complainant are highly significant, whereas indicated fear, boredom, and extraversion of the complainant are slightly significant in distinguishing pervasive from mediocre complaints. It should also be noted that happiness, sadness, excitement, and personality traits of agreeableness, neuroticism, and openness to experience are found to have no significant impact on distinguishing any of the interaction-based clusters from mediocre complaints.

The count of hashtags used has a statistically significant positive impact on distinguishing spiteful complaints from mediocre complaints with the logit coefficient of 0.120 and an exponential beta of 1.128 ($\beta=0.120$; $p<.001$; $\text{Exp}(\beta)= 1.128$). This suggests that other things being equal, the odds of a posted complaint to become spiteful would be 1.128 times more than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become spiteful increases by 99.9 % with an increase in the hashtags used. Sentiment also has a statistically significant positive impact with the logit coefficient of 0.498 and an exponential beta of 1.645 ($\beta=0.498$; $p<.001$; $\text{Exp}(\beta)= 1.645$), indicating that other things being equal, the odds of a posted complaint to become spiteful would be 1.645 times more than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become spiteful increases by 99.9 % as the sentiment of the complaint converges to positive polarity. The anger of the complainant has a statistically significant positive impact with the logit coefficient of 9.790 and an exponential beta of 17848.201 ($\beta=9.790$; $p<.001$; $\text{Exp}(\beta)= 17848.201$), indicating that other things being equal, the odds of a posted complaint to become spiteful would be 17848.201 times more than the odds of a complaint belonging to become a mediocre complaint and the probability of a posted complaint to become spiteful increases by 99.9 % as the indicated anger increases. The count of accounts followed has a statistically significant impact with the logit coefficient of 0 and an exponential beta of 1 ($\beta=0$; $p<.001$; $\text{Exp}(\beta)= 1$) indicating that other things being equal, the odds of a posted complaint to become spiteful would be 100% more than the odds of a complaint belonging to become a mediocre complaint. Accordingly, the probability of a posted complaint becoming spiteful increases by 99.9%, with an increase in the count of accounts followed by the complainant. Finally, extraversion has statistically

significant a negative impact on distinguishing spiteful complaints from mediocre complaints with the logit coefficient of -0.459 and an exponential beta of 0.632 ($\beta = -0.459$; $p < .001$; $\text{Exp}(\beta) = 0.632$). This suggests that other things being equal, the odds of a posted complaint to become spiteful would be 0.632 times less than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become spiteful decreases by 99.9 % with an increase in the extraversion of the complainant.

The count of followers followed accounts, sentiment, boredom, conscientiousness, and the complainant's extraversion have a statistically significant positive impact on distinguishing pervasive complaints from mediocre ones. The logit coefficient for the count of followers and following is 0, and their exponential betas are 1 ($\beta = 0$; $p < .01$; $\text{Exp}(\beta) = 1$). This suggests that other things being equal, the odds of a posted complaint to become pervasive would be 100% more than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become pervasive increases by 99.9% with an increase in the count of followers and followed accounts. Sentiment also has a statistically significant positive impact with the logit coefficient of 1.304 and an exponential beta of 3.682 ($\beta = 1.304$; $p < .001$; $\text{Exp}(\beta) = 3.682$), indicating that other things being equal, the odds of a posted complaint to become pervasive would be 3.682 times more than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become pervasive increases by 99.9 % as the sentiment of the complaint converges to positive polarity. Conscientiousness has a statistically significant positive impact with the logit coefficient of 1.836 and an exponential beta of 6.274 ($\beta = 1.836$; $p < .001$; $\text{Exp}(\beta) = 6.274$), indicating that other things being equal, the odds of a posted complaint to become pervasive would be 6.274 times more than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become pervasive increases by 99.9 % as the conscientiousness of complainant increases. Another factor that has a positive impact is boredom with a logit coefficient of 5.791 and an exponential beta of 327.285 ($\beta = 5.791$; $p < .05$; $\text{Exp}(\beta) = 327.285$), indicating that other things being equal, the odds of a posted complaint to become pervasive would be 327.285 times more than the odds of a complaint belonging to become a mediocre complaint and the probability of a posted complaint to become pervasive increases by 95 % as the boredom of complainant

increases. Extraversion also has a statistically significant positive impact with the logit coefficient of 0.695 and an exponential beta of 2.003 ($\beta=0.695$; $p<.005$; $\text{Exp}(\beta)= 2.003$), indicating that other things being equal, the odds of a posted complaint to become pervasive would be 2.003 times more than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become pervasive increases by 95 % as the extraversion of the complainant increases. The count of mentions and indicated fear has a negative impact on distinguishing pervasive complaints from mediocre complaints with the logit coefficient of -1.625 and an exponential beta of 0.197 ($\beta=-1.625$; $p<.001$; $\text{Exp}(\beta)= 0.197$). This suggests that other things being equal, the odds of a posted complaint to become pervasive would be 0.197 times less than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become pervasive decreases by 99.9 % with an increase in the count of mentioned accounts in complaint tweet. Finally, fear has statistically significant a negative impact on distinguishing pervasive complaints from mediocre complaints with the logit coefficient of -8.295, and an exponential beta of 0 ($\beta=-0.459$; $p<0.1$; $\text{Exp}(\beta)= 0$), suggesting that 0 times less than the odds of a complaint belonging to become a mediocre complaint, and the probability of a posted complaint to become pervasive decreases by 90% with an increase in the indicated fear in the complaint.

Table 24: Full Observation Parameter Estimates for Content-Based Categories

Comparison	<i>Mediocre vs. Spiteful</i>				Mediocre vs. Pervasive			
	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)
Intercept	(3.770)***	.142	703.714		(6.268)***	.405	239.907	
Count of mentions	(.039)	.041	.929	.962	(1.625)***	.158	106.167	.197
Count of hashtags	.120***	.028	18.353	1.128	.088	.075	1.374	1.093
Sentiment	.498***	.107	21.759	1.645	1.304***	.272	22.901	3.682
Happiness	(.517)	.844	.375	.596	(3.593)	2.500	2.066	.028
Sadness	.850	.861	.976	2.339	2.579	2.300	1.256	13.178
Anger	9.790***	1.058	85.574	17848.201	(4.069)	4.303	.894	.017
Fear	(1.119)	1.183	.894	.327	(8.295)*	4.566	3.300	.000
Excitement	(.121)	.409	.088	.886	.684	.866	.624	1.982
Boredom	.557	.766	.529	1.745	5.791**	1.878	9.510	327.285
Agreeableness	.212	.143	2.194	1.237	(.082)	.364	.051	.921
Conscientiousness	(.004)	.135	.001	.996	1.836***	.334	30.202	6.274
Extraversion	(.459)***	.135	11.487	.632	.695**	.360	3.719	2.003
Neuroticism	(.069)	.112	.384	.933	.403	.303	1.766	1.496
Openness to Exp.	.032	.125	.065	1.032	.155	.335	.215	1.168
Following	.000***	.000	54.881	1.000	.000***	.000	33.608	1.000
Followers	.000	.000	.020	1.000	.000***	.000	47.357	1.000

Note: *B* – Estimated multinomial logistic regression coefficients; Std. Error – Standard errors of the individual regression coefficients; Wald – Wald chi-square test that examines the null hypothesis that the estimate equals 0; Exp (*B*) – An exponential beta; ****p*<0.01, ***p*<0.05, **p*<0.1

4.3.3.2.1.2.2. Holdout Sample Results for Interaction-Based Clusters

The likelihood ratio test results, shown in Table 25, indicate that the model is statistically significant (Sig. = $p < 0.0001$) and denote that employed independent variables can predict the type of complaint with statistical significance. The final model is statistically significant, $\chi^2 = 370.133$ with 32 degrees of freedom ($p < 0.0001$), and hence it may be stated that the model can significantly distinguish the type of complaint based on the indicated independent variables. Pseudo R -square results indicate that the final model explains 0.623 (Nagelkerke R^2) of the variances in interaction-based complaint types.

Table 25: Model Fitting Information for Holdout Sample Results of Interaction-Based Clusters

Independent variables	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept only	1008.526			
Final	638.393	370.133	32	.000

The results, shown in Table 26, revealed that each complaint cluster retains a diverse set of explanatory variables. It is found that spiteful complaints differentiate from mediocre complaints in indicated anger, count of mentioned accounts and followers, and extraversion of the complainant. In contrast, pervasive complaints differentiate from mediocre complaints in indicated boredom, fear, and the count of hashtags used mentioned accounts and followers. The count of followers is high, whereas the number of hashtags used, anger, and extraversion of the complainant are slightly significant in distinguishing spiteful complaints from mediocre complaints. On the other hand, the count mentioned accounts are high, whereas the number of hashtags used and followers, fear, and boredom are slightly significant in distinguishing pervasive complaints from mediocre complaints. It should also be noted that the count of accounts followed, sentiment, happiness, sadness, excitement, and personality traits of conscientiousness, agreeableness, neuroticism, and openness to experience, is found to have no significant impact on distinguishing any of the interaction-based clusters from mediocre complaints.

As mentioned, the count of followers has a statistically significant positive impact on distinguishing spiteful complaints from mediocre complaints. The logit coefficient for the count of followers is 0, and its exponential beta is 1 ($\beta=.0$; $p<.05$; $\text{Exp}(\beta)= 1$). This suggests that other things being equal, the odds of a posted complaint to become spiteful would be 100% more than the odds of a complaint belonging to become a mediocre complaint. Thus, the probability of a posted complaint becoming spiteful increases by 95% with an increase in followers' count. Anger also has a significant positive impact on distinguishing spiteful complaints from mediocre complaints with a logit coefficient of 8.018 and an exponential beta of 3035.250 ($\beta=.8.018$; $p<.1$; $\text{Exp}(\beta)= 3035.250$). This suggests that other things being equal, the odds of a posted complaint to be spiteful would be 3035.250 times more than the odds of being a mediocre complaint. Therefore, the probability of a posted complaint to be spiteful increases by 90%, with an increase in the indicated anger. The logit coefficient for the count of hashtags used is 0.277 with an exponential beta of 1.319 ($\beta=0.277$; $p<.1$; $\text{Exp}(\beta)= 1.319$), suggesting that other things being equal, the odds of a posted complaint to be spiteful would be 1.319 times more than the odds of being a mediocre complaint. Accordingly, the probability of a posted complaint to be spiteful increases by 90%, with an increase in the count of hashtags used in the complaint tweet. Extraversion has a negative impact on distinguishing spiteful complaints from mediocre complaints. The logit coefficient for extraversion is -0.872, and the exponential beta is 0.418 ($\beta=-0.872$; $p<.1$; $\text{Exp}(\beta)= 0.418$), suggesting that other things being equal, the odds of a posted complaint be spiteful would be 0.418 times less than the odds being a mediocre complaint. Accordingly, the probability of a posted complaint to be spiteful decreases by 90% as the complainant's extraversion increases.

The count of followers, hashtags used, and boredom has a statistically significant positive impact on distinguishing pervasive complaints from mediocre ones. The logit coefficient for the count of followers is 0, and its exponential beta is 1 ($\beta=.0$; $p<.05$; $\text{Exp}(\beta)= 1$). This suggests that other things being equal, the odds of a posted complaint to become pervasive would be 100% more than the odds of a complaint belonging to become a mediocre complaint. Thus, the probability of a posted complaint becoming pervasive increases by 95% with an increase in followers' count. The count of hashtags used has a positive impact with a logit coefficient of 0.439 and an exponential beta of 1.551 ($\beta=0.439$; $p<.05$; $\text{Exp}(\beta)= 1.551$), suggesting that other things being equal, the

odds of a posted complaint to be pervasive would be 1.551 times more than the odds of being a mediocre complaint. Therefore, the probability of a posted complaint to be pervasive increases by 95%, with an increase in the count of hashtags used in the complaint tweet. Indicated boredom also has a positive impact with a logit coefficient of 9.638 and an exponential beta of 15332.632 ($\beta=9.638$; $p<.05$; $\text{Exp } (\beta)= 15332.632$), suggesting that other things being equal, the odds of a posted complaint to be pervasive would be 15332.632 times more than the odds of being a mediocre complaint. Accordingly, the probability of a posted complaint to be pervasive increases by 95%, with an increase in the complainant's indicated boredom. The count of mentions in the complaint and the indicated fear of the complainant have a negative impact on distinguishing pervasive complaints from mediocre complaints. The logit coefficient for the count of mention is -0.764, and the exponential beta is 0.466 ($\beta=-0.764$; $p<.001$; $\text{Exp } (\beta)= 0.466$), suggesting that other things being equal, the odds of a posted complaint be spiteful would be 0.466 times less than the odds being a mediocre complaint. Accordingly, the probability of a posted complaint to be spiteful decreases by 99.9% as the count of mentioned accounts in the complaint tweet increases. Finally, the logit coefficient for boredom is -20.696, and its exponential beta is 1.027E-9 ($\beta=-20.696$; $p<.05$; $\text{Exp } (\beta)= 1.027\text{E-}9$), suggesting that other things being equal, the odds of a posted complaint be spiteful would be 1.027E-9 times less than the odds being a mediocre complaint. Therefore, the probability of a posted complaint being spiteful decreases by 95% as the complainant's indicated boredom increases.

Table 26: Parameter Estimates for Interaction-Based Clusters

Comparison	<i>Mediocre vs. Spiteful</i>				Mediocre vs. Pervasive			
	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)	<i>B</i>	Std. Error	Wald	Exp (<i>B</i>)
Intercept	(.030)	.546	.003		(2.318)**	.914	6.433	
Count of mentions	.136	.103	1.731	1.146	(.764)***	.290	6.954	.466
Count of hashtags	.277*	.143	3.733	1.319	.439**	.198	4.931	1.551
Sentiment	(.508)	.393	1.670	.602	(.150)	.589	.065	.861
Happiness	(3.510)	2.964	1.402	.030	(3.060)	4.334	.498	.047
Sadness	(1.041)	3.642	.082	.353	(.350)	5.342	.004	.705
Anger	8.018*	4.744	2.857	3035.250	(1.365)	8.511	.026	.255
Fear	(2.761)	6.470	.182	.063	(20.696)**	10.280	4.053	1.027E-9
Excitement	1.124	1.306	.741	3.078	.317	2.250	.020	.728
Boredom	.709	3.616	.038	2.031	9.638**	4.543	4.501	15332.632
Agreeableness	.403	.549	.539	1.497	.267	.794	.113	1.307
Conscientiousness	(.176)	.522	.114	.839	.796	.756	1.108	2.217
Extraversion	(.872)*	.527	2.744	.418	.902	.767	1.384	2.465
Neuroticism	(.038)	.424	.008	.963	.294	.649	.205	1.341
Openness to Exp.	.132	.461	.082	1.141	.474	.704	.453	1.606
Following	.000	.000	.014	1.000	.000	.000	.563	1.000
Followers	.000**	.000	5.753	1.000	.000**	.000	11.390	1.000

Note: *B* – Estimated multinomial logistic regression coefficients; Std. Error – Standard errors of the individual regression coefficients; Wald – Wald chi-square test that examines the null hypothesis that the estimate equals 0; Exp (*B*) – An exponential beta; ****p*<0.01, ***p*<0.05, **p*<0.

4.3.4. Model Validation and Testing

In machine learning applications, following the training, models are required to be validated and tested on a holdout data set. For measuring the model's effectiveness, accuracy, precision and recall metrics are employed. Accuracy refers to the proportion of the correct predictions to the total number of predictions (Kotsiantis, 2007). Additionally, recall refers to the fraction of correct predictions over the total number of true positives and false negatives, whereas precision denotes the fraction of correct predictions over the total number of true positive and true negative predictions (Kolchyna et al., 2015).

For model validation, precision, recall and accuracy metrics are accessed for holdout samples of interaction-based clusters and content-based categories. Each cluster and category are altered to contain equal number of observations, referring to 148 observations for each content-based category and 153 observations for each interaction-based cluster. Results revealed that the proposed model predicts content-based categories with 47.9% accuracy. Recall for legal threat, call for boycott, and brand switch warning categories is respectively 62.8%, 44.5%, and 39.5%. Moreover, precision for legal threat, call for boycott, and brand switch warning categories is respectively 69.4%, 43.6%, and 37.3%. The accuracy for interaction-based clusters is 65.8%, whereas recall for mediocre, spiteful, pervasive clusters is respectively 67.3%, 43.8%, and 86.3%. Finally, precision mediocre, spiteful, pervasive clusters is respectively 52.8%, 53.6%, and 95%.

Table 27: Precision, Recall and Accuracy Results for The Validation Set

Content-Based Categories					
Observed	Predicted				Recall
	Legal Threat	Boycott	Brand Switch	Normal	
Legal Threat	93	18	23	14	62.8%
Boycott	14	65	39	30	44.5%
Brand Switch	18	31	58	41	39.5%
Normal	12	35	35	66	44.6%
Precision	69.4%	43.6%	37.4%	43.7%	47.9%

Interaction-Based Clusters				
Observed	Predicted			Recall
	Mediocre	Spiteful	Pervasive	
Mediocre	103	47	3	67.3%
Spiteful	82	67	4	43.8%
Pervasive	10	11	132	86.3%
Precision	52.8%	53.6%	95.0%	65.8%

As the next step, the proposed model is tested with a test set that contains observations that are previously unseen by the model. Accordingly, a new data set is formed by gathering complaints about investigated companies that are posted on September 18, 2020. Data gathering resulted in 1,983 observations, and among which 500 were randomly selected. Results revealed that the proposed model predicts content-based categories with 61.7% accuracy. Recall for legal threat, call for boycott, and brand switch warning categories is respectively 66.7%, 40%, and 50%. Moreover, precision for legal threat, call for boycott, and brand switch warning categories is respectively 22.2%, 9.5%, and 10.7%. The accuracy for interaction-based clusters is 70.5%, whereas recall for mediocre, spiteful, pervasive clusters is respectively 78.3%, 50%, and 83.3%. Finally, precision of mediocre, spiteful, pervasive clusters is respectively 97.7%, 9.5%, and 54.8%.

Among content-based categories, even though it exceeds 99% for complaints that do not indicate a behavioral intention (labelled as 'normal' in related tables), precision is

lowest for the call for boycott category and is followed by brand switch warning and legal threat categories. The main reason behind this fact is that number of observations for these categories are slightly low when compared with the observations where no behavioral intention is indicated. Similarly, due to the low number of observations, pervasive cluster holds lowest precision rate among interaction-based clusters. It should be noted that, precision of the predictions for these clusters and categories will improve as more observations will be seen by the model.

Table 28: Precision, Recall and Accuracy Results for The Test Set

Content-Based Categories					
Observed	Predicted				Recall
	Legal Threat	Boycott	Brand Switch	Normal	
Legal Threat	2	0	1	0	66.7%
Boycott	0	2	2	1	40.0%
Brand Switch	0	1	3	2	50.0%
Normal	7	18	22	439	90.3%
Precision	22.2%	9.5%	10.7%	99.3%	61.7%
Interaction-Based Clusters					
Observed	Predicted			Recall	
	Mediocre	Spiteful	Pervasive		
Mediocre	335	63	30		78.3%
Spiteful	4	7	3		50.0%
Pervasive	4	4	40		83.3%
Precision	97.7%	9.5%	54.8%		70.5%

4.4. DISCUSSION OF THE RESULTS AND MANAGERIAL IMPLICATIONS

Ubiquity, ease of use, and speed of social media gave rise to the trend of sharing consumption experiences online. The upsurge in electronic WOM created both an advantageous and dangerous field for the companies. An effectively handled online complaint may result in winning the complainant back and attract new customers through shaping reputation (Einwiller and Steilen, 2015), but it can also dangerously backfire on the company when not handled effectively (Lee and Song, 2010). Therefore, understanding online complaints, identifying typologies of both online complaints and online complainants, and excelling in online complaint management plays a crucial role for the companies.

Although consumer complaint behavior is widely studied, there is no considerable amount of research on online complainant types and online complaint handling. Due to the differences between online and offline environments, traditional complaint typologies cannot be fully integrated into online complaints. Hence, for obtaining a more accurate estimation of online complaint's magnitude, a typology for online complaints is needed. Moreover, though the extant literature validates the relationship between personality traits, emotions, and complaining intention through the surveying method, the method is prone to respondents' risk of honesty. Therefore, for obtaining both more accurate estimation and a deeper understanding, analysis of unstructured, autogenous, textual data is required. Also, companies need an effective method of handling complaints in a way to balance responsiveness and cost. Based on these gaps, the aims of this thesis have been determined as to both identify content-based and interaction-based online consumer complaint types and predict complaint types according to the complaint magnitude rooted in complainants' personality traits, emotion, Twitter usage activity, and complaints sentiment polarity and feedback frequency.

Initiated by the objectives of this thesis, online complaints were classified according to their content. For content-based categorization of the online complaints, widely accepted CCB models of Hirschman (1970), Day and Landon (1977), Singh (1988), and Tronvoll (2012) were revisited in accordance with the nature of the online environment. A distinction based on the previous models was put forth by eliminating the distinction between public and private actions. Rather, the indicated further action (i.e.,

legal threat, boycott, brand switching) was considered for determining content-based categories. Also, since posting a complaint on social media actually refers to “voicing” the experience that enables a constant “WOM,” voicing the complaint or taking no action conditions that are included in the traditional CCB models were not included in our content-based categorization. Accordingly, three categories of online complaint contents were examined: taking legal action, calling for a boycott, and brand switch warning (referred to as “exit” in previous models). Results revealed that 20% of the complaints (584 out of 297,374) denote a behavioral intention, and 39% of these contain a call for boycott, 35% brand switch warning, and 26% legal threat.

Initiated by the objectives of this thesis and the gaps in the literature, the first research question was: *What are the characteristics of content-based categories of online complaints?* Complaints that contain call for a boycott are characterized by negative sentiment polarity. Additionally, complainants who denote a boycott call are mainly even-tempered individuals and feel sad about the incident. In other words, what differentiates a boycott indicating complaint from the complaints that do not indicate a behavioral intention is the negative sentiment polarity of the complaint, as well as the sadness articulated in the complaint tweet and even-tempered nature of the complainant. Complaints that contain brand switch warnings are characterized by a low number of mentioned accounts in the complaint post. Complainants who denote a brand switch warning are mainly calm and even-tempered individuals who feel sad about the incident. To put it another way, what differentiates a brand switch indicating complaint from the complaints that do not indicate a behavioral intention is the low number of mentioned accounts in the complaint post and the sadness articulated in the complaint tweet, calm and even-tempered nature of the complainant. Finally, complaints that contain legal threats are characterized by a low number of mentioned accounts in the complaint post. Complainants who denote a legal threat are mainly peaceful, unagitated but disagreeable individuals and feel afraid and bored about the incident. In other words, what differentiates a legal threat indicating complaint from the complaints that do not indicate a behavioral intention is a low number of mentioned accounts in the complaint post, as well as the fear and boredom articulated in complaint tweet, and peaceful, unagitated and disagreeable nature of the complainant.

Following the gaps in the literature, the second and third research questions were: *What are the interaction-based types of online complaints regarding complainants' personality traits, emotions, and Twitter usage activity, as well as the sentiment and feedback frequency of their complaint? What are the characteristics of interaction-based types of online complaints?* Two-step cluster analysis revealed three types of interaction-based clusters: (1) *Spiteful* complaints that contain numerous consecutive tweets but get almost no interaction in terms of favorites, replies, and retweets; (2) *Pervasive* complaints that do not exceed one tweet but get a massive number of likes, favorites, and retweets; (3) *Mediocre* complaints that have a low interaction rate, and the count of consecutive complaint tweets sent by the complainant is lesser when compared with the other clusters. Results revealed that 0,4% of the complaints are pervasive, 2,7% are spiteful, and 96,9% are mediocre. Findings point out that what differentiates a spiteful complaint from a mediocre one is the high number of hashtags in the post, the anger articulated in the complaint and introversion of the complainant. In other words, spiteful complaints are characterized by a high number of hashtags, articulated anger, and the low level of extraversion of the complainant. Pervasive complaints are characterized by a high number of hashtags and a low number of mentions. Complainants who post pervasive complaints are intrepid individuals and feel bored about the incident. To put it another way, what differentiates pervasive complaints from mediocre complaints is low number of mentions, high level of hashtags, boredom, and intrepid nature of the complainant.

Originating from the literature gaps, the fourth research question was: *How can the complaint type be predicted via machine learning and text analytics?* In accordance with its aim, this thesis employs a multi-level approach for identifying interaction-based clusters and content-based categories and revealing their characteristics. Findings verify that the proposed model provides a novel approach for predicting complaint types via machine learning and text analytics. In other words, by depending on the results of the thesis, companies can predict the complaint type in terms of content and interaction. By employing text analytics tools, the proposed model can reveal the sentiment polarity of the complaint, the personality of the complainant, and how the complainant feels about the incident. Through the utilization of these information together with the count of mentioned accounts and used hashtags in the complaint, and the count of followers and followings of the complainant, the proposed model can predict whether the complaint

contains a legal threat, call for boycott, or brand switch warning. In a similar vein, the future feedback frequency in terms of interaction (likes, retweets, favorites, and replies) can be predicted through our prediction model employing multinomial logistic regression which is a supervised machine learning algorithm. For instance, if a complaint tweet contains numerous hashtags, the employed text analytics tools reveal that the complainant is more of an introvert than an extrovert and feels angry about the incident, the model labels the complaint as a “spiteful” one and predicts that the complaint will get almost no interaction in terms of favorites, retweets, and replies, but the complainant will keep on posting about the same incident. Similarly, if a complaint tweet contains few mentioned accounts, and the employed text analytics tools reveal that the complainant is an even-tempered person who feels less of angry but more of sad about the incident, the model predicts that the complainant holds a switching intention.

Following the aims of this thesis and the gaps in the literature, the last research question was: *How can online complaint typology be used for prioritizing customer complaints to be answered?* As explained in detail above, the proposed model estimates the content and feedback frequency of complaints separately. Accordingly, findings disclose three content-based categories (legal threat, call for boycott, switching intention containing complaints) and three interaction-based clusters (mediocre, pervasive, and spiteful complaints). Companies can utilize the proposed model and complaint types for complaint prioritization by integrating the model into a CRM module. However, it should be noted that the proposed model provides a flexible framework to be applicable by companies with different motivations. Therefore, for ensuring the validity of prioritization, determining the “highly critical” complaint types and “less critical” complaint types is left to companies that will utilize the proposed model. In other words, the model provides a flexible framework for companies by allowing them to determine which type of complaints are more critical and require an urgent answer. Since interaction-based and content-based complaint types are mutually exclusive, a complaint that contains legal threat can be mediocre, spiteful, or pervasive. A pervasive complaint may contain a call for boycott, legal threat, or switching intention in a similar vein. Therefore, while integrating the proposed model to a CRM module, companies should rank the following nine complaint types from highly critical to less critical: (1) Pervasive complaints that contain a legal threat, (2) Mediocre complaints that contain a legal threat, (3) Spiteful complaints that

contain a legal threat, (4) Pervasive complaints that contain a call for boycott, (5) Mediocre complaints that contain a call for boycott, (6) Spiteful complaints that contain a call for boycott, (7) Pervasive complaints that contain switching intention, (8) Mediocre complaints that contain switching intention, (9) Spiteful complaints that contain switching intention. Since priorities may vary from company to company and from time to time, the proposed model has two main strengths. First, due to its flexible nature that enables tailored priority ranking, the model can be utilized in various settings and industries. Second, priority ranking can be updated when needed. The ability to customize the model according to the company goals and needs provides companies the opportunity to utilize a tailored, proactive, and enhanced social CRM and online complaint handling process.

The findings of this thesis and proposed model contribute to both marketing literature and practice. The broad utilization of social networks creates vast amounts of data in various formats, providing us the ability to reach microscopic behavioral patterns on a macroscopic global scale (Golder and Macy, 2014). To reach microscopic patterns on a macroscopic scale both in academia and practice, text analytics and machine learning are gaining significant currency due to their capability to process high volume data in a cost-effective way. Based on the gaps in the literature, findings of this thesis contribute to social CRM, CCB, and online complaint handling literatures. The main contributions of this thesis can be listed as; (1) identifying types of online complaints on Twitter; (2) proposing a set of content and profile features that can be employed for revealing complaint type; (3) revealing the relationship between content features, profile features, and online complaint type (4) proposing a novel approach for employing text analytics and machine learning for excelling online complaint through prioritization.

From the practitioners' viewpoint, social media is a fruitful landscape containing high volume consumer data that can provide significant insights. However, overlooking this fertile landscape's idiosyncratic dynamics and drawbacks can dangerously backfire on companies. Due to the volume of interaction and data, some complaints fall off the companies' radar, and some remain unanswered for a long time. Since customer experience is a holistic process covering all touch-points customers have with an organization (Bolton et al., 2014), customers' overall experience with the company may worsen once their complaint remains unanswered. Moreover, such complaints contain a high risk of reaching a wider audience and causing outrage (Pfeffer et al., 2014).

Accordingly, the proposed model brings a novel and more proactive approach that ensures efficient online complaint handling and efficient customer experience management.

Moreover, understanding microscopic, individual behavior patterns on a macroscopic social media scale is critical for companies to survive in the digital landscape. In addition to ensuring high responsiveness, it is vital to understand every complainant and estimate each complaint's magnitude to build a sufficient complaint handling procedure and prevent a possible online firestorm. Since they are only capable of capturing occurred online firestorms, social listening and social CRM remain incapable of estimating the future magnitude of a complaint. However, surviving an online firestorm without any damage is almost impossible, and hence, instead of capturing already occurred ones, companies should be able to estimate possible firestorms and take preventive actions before the occurrence. Therefore, it can be stated that responsiveness is necessary but not sufficient for effective web care and should be supported with additional individual-level, predictive, and proactive CRM support modules. Therefore, the proposed model provides a cost-effective massive antenna to estimate each complaint's magnitude and handle complaints effectively.

This thesis provides a novel approach to social CRM and online complaint handling by utilizing text analytics and machine learning. Following this thesis's context and findings, five options for utilizing text analytics and machine learning in excelling online complaint management can be listed: (1) Complainant's personality can be revealed through analyzing tweet history, (2) Complainants' feelings about the incident can be obtained through analyzing complaint tweet, (3) The sentiment polarity of complaint tweet can be attained, (4) The behavioral intention (call for boycott, legal threat, switching to another brand) embedded in the content of complaint tweet can be revealed, (5) Future interaction magnitude of the complaint can be predicted. A flexible model for prioritizing online complaints is presented to improve social CRM effectiveness and online complaint management procedures. Accordingly, by applying some or all of the options mentioned above, companies can customize and improve their online complaint management procedures and alleviate online complaints' detrimental effects.

The modular and flexible nature of the proposed model enables customization in accordance with the needs and dynamics of a company. By depending upon both content-based classes and interaction-based clusters, decision-makers may break online complaints into nine groups, as mentioned previously. As of the following step, companies may rank these nine groups of complaint types in subjective urgency or danger order by conducting a multi-criteria decision analysis. If a company decides to utilize either of the interaction-based or content-based types, the urgency ranking can be assessed by ordering related types (i.e., mediocre, pervasive and spiteful complaints for interaction-based types) via multi-criteria decision analysis.

4.5. LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

The findings of this thesis should be noticed within the context of three main limitations. First, the analyzed data consists of complaint tweets that are posted in English. Hence, the employed keyword sets for content-based classification, and the results obtained from text analytics tools are only valid for complaints posted in English. In order to employ the proposed model for analyzing complaints that are posted in a different language, a new set of keywords should be revealed via n-gram analysis. Moreover, the availability of the employed text analytics tools for the given language should be checked and replaced with other tools if they are not available in the given language. Therefore, it would be fruitful to test the model with complaints posted in a different language to ensure the validity of the model for other languages and re-test the significance of the employed content and profile features in different languages.

Second, the scope of this thesis in terms of employed companies and industries is limited to two companies with the largest market cap from each economic sector listed in the S&P 500 Index. Accordingly, complaints about JPMorgan Chase and Bank of America from “financials” industry, Amazon and Home Depot from “cyclicals,” UnitedHealth and CVS Health from “healthcare,” UPS and FedEx from “industrials,” Walmart and Walgreens from “non-cyclicals,” and AT&T and Verizon from “telecommunication” are analyzed. Moreover, companies operating in B2B domain are not included in this study. For broadening the scope, analyzed companies may be diversified in further studies. Additionally, even though the clients of some companies

(i.e., advertising agencies) are other companies, due to the nature of the sector in which they operate, obtaining feedback from clients' customers may yield significant insights regarding the quality and/or impact of their services.

Third, since this thesis analyzes complaints that are gathered from Twitter, employed profile features are determined accordingly. In future studies, the scope of proposed model may be broadened by modifying employed independent variables in accordance with the analyzed social media platform. For instance, the nature of Facebook is quite different than Twitter. Twitter users mainly use fictional names, do not indicate their occupation, or do not write place of residence in their profile. Moreover, text content of tweets can only obtain up to 280 characters. However, on Facebook, users mainly sign in with their real names, indicate their place of birth, place of residence, educational background, and occupation on their pages. In addition, Facebook do not impose any character limits per post. Accordingly, by employing additional profile features and/or modifying the existing ones, the scope of this thesis may be broadened for other social media platforms.

As mentioned previously, the modular nature of the proposed model enables customization in accordance with the needs and dynamics of a company. A further application study may be conducted for ranking revealed complaint types in subjective urgency order by employing a multi-criteria decision analysis method for a given company or industry. By including managers from different departments with varied priorities, the urgency ranking may be constituted.

REFERENCES

- Abney, A. K., Pelletier, M. J., Ford, T.-R. S., & Horky, A. B. (2017). #IHateYourBrand: Adaptive Service Recovery Strategies on Twitter. *Journal of Services Marketing*, 31(3), 281–294. <https://doi.org/10.1108/JSM-02-2016-0079>
- Acker, O., Gröne, F., Akkad, F., Pötscher, F., & Yazbek, R. (2011). Social CRM: How Companies Can Link into The Social Web of Consumers. *Journal of Direct, Data and Digital Marketing Practice*, 13(1), 3-10. <https://doi.org/10.1057/dddmp.2011.17>
- Adamopoulos, P., Ghose, A., & Todri, V. (2018). The Impact of User Personality Traits on Word of Mouth: Text-Mining Social Media Platforms. *Information Systems Research*, 29(3), 612-640. <https://doi.org/10.1287/isre.2017.0768>
- Adams, M. (2000). *More Fliers Land on The Internet to Air Complaints: Government, Some Airlines Post Links to E-Mail Forms*. USA Today, p. B12.
- Agarwal, N. (2016). A Study of Personality Traits of B. Ed Students. *Students (June 15, 2020)*. Agarwal, Nidhi and Pundir Neelam, 127-131.
- Agarwal, R., Prabhakar, T. V., & Chakrabarty, S. (2008, August). "I Know What You Feel": Analyzing the Role of Conjunctions in Automatic Sentiment Analysis. In *International Conference on Natural Language Processing* (pp. 28-39). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-85287-2_4
- Aggarwal, C. C. (2018). *Machine Learning for Text*. Cham, New York: Springer International Publishing.
- Agnihotri, A., & Bhattacharya, S. (2016). Online Review Helpfulness: Role of Qualitative Factors. *Psychology & Marketing*, 33(11), 1006-1017. <https://doi.org/10.1002/mar.20934>
- Ahlqvist, T., Bäck, A., Heinonen, S., & Halonen, M. (2010). Road-Mapping the Societal Transformation Potential of Social Media. *Foresight*, 12(5), 3–26. <https://doi.org/10.1108/14636681011075687>

- Ahmad, S. (2002). Service Failures and Customer Defection: A Closer Look at Online Shopping Experiences. *Managing Service Quality*, 12, (1), 19-29. <https://doi.org/10.1108/09604520210415362>
- Ali, F., Kwak, K. S., & Kim, Y. G. (2016). Opinion Mining Based on Fuzzy Domain Ontology and Support Vector Machine: A Proposal to Automate Online Review Classification. *Applied Soft Computing*, 47, 235-250. <https://doi.org/10.1016/j.asoc.2016.06.003>
- Allbeck, J. M., & Badler, N. I. (1998). Avatars A/Spl Grave/La Snow Crash. In Proceedings Computer Animation'98, Cat. No. 98EX169, 19-24, IEEE.
- Allen, J. (2017, October 5). The History of Social Media. Future Marketing. Retrieved December 5, 2020, from <https://www.future-marketing.co.uk/the-history-of-social-media/>
- Allport, G. W. (1961). Pattern and growth in personality. New York, NY: Holt, Rinehart and Winston.
- Allport, G. W., & Odbert, H. S. (1936). Trait-Names: A Psycho-Lexical Study. *Psychological Monographs*, 47(1), i.
- Alpaydin, E. (2014). *Introduction to Machine Learning*. S.I.: MIT Press.
- Alt, R., & Reinhold, O. (2016). *Social Customer Relationship Management*. Grundlagen, Anwen.
- Andreasen, A. R. (1985), Consumer Response to Dissatisfaction in Loose Monopolies: The Case of Medical Care. *Journal of Consumer Research*, 12 (2), 135-41. <https://doi.org/10.1086/208502>
- Andreasen, A. R. (1988). *Consumer Complaints and Redress: What We Know and What We Don't Know*. In the Frontier of Research in the Consumer interest, E.S. Maynes, Ed. Columbia, MO: American Council on Consumer Interests.
- Andreasen, A. R. and A. Best (1977). Consumers Complain – Does Business Respond? *Harvard Business Review*, 55 (July-Aug), 93-101.

- Andreassen, T. W., & Streukens, S. (2013). Online Complaining. *Managing Service Quality: An International Journal*, 23(1), 4-24. <https://doi.org/10.1108/09604521311287632>
- Apte, C., Damerau, F. J., and Weiss, S. M. (1994). Automated Learning of Decision Rules for Text Categorization. *ACM Transactions on Information Systems*, 12(3): 233-251. <https://doi.org/10.1145/183422.183423>
- Apuke, O. D. (2016). Social and Traditional Mainstream Media of Communication: Synergy and Variance Perspective. *Journal of New Media and Mass Communication*, 53(4), 83-86.
- Arndt, J. (1967). *Word of Mouth Advertising: A Review of The Literature*. New York: Advertising Research Foundation.
- Ash, S. B. (1978). *A Comprehensive Study of Consumer Satisfaction with Durable Goods*. In *Advances in Consumer Research*, H. Keith Hunt, Ed. Ann Arbor, MI: Association for Consumer Research.
- Askool, S., & Nakata, K. (2011). A Conceptual Model for Acceptance of Social CRM Systems Based on A Scoping Study. *AI & Society*, 26(3), 205-220. <https://doi.org/10.1007/s00146-010-0311-5>
- Asur, S., & Huberman, B. A. (2010). Predicting the Future with Social Media. *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* 1, 492-499. <https://doi.org/10.1109/WI-IAT.2010.63>
- Au, N., Buhalis, D., & Law, R. (2009). Complaints on The Online Environment—The Case of Hong Kong Hotels. *Information and Communication Technologies in Tourism* 2009, 73-85.
- Averill, J. (1983). Studies on Anger and Aggression: Implications for Theories of Emotion. *American Psychologist*, 38(11), 1145–1160.
- Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The Role of Emotions in Marketing. *Journal of The Academy of Marketing Science*, 27(2), 184-206. <https://doi.org/10.1177/0092070399272005>

- Bailey, A. A. (2004). Thiscompanysucks.Com: The Use of The Internet in Negative Consumer-To-Consumer Articulations. *Journal of Marketing Communications*, 10(3), 169-182. <https://doi.org/10.1080/1352726042000186634>
- Balaji, M. S., Khong, K. W., & Chong, A. Y. L. (2016). Determinants of Negative Word-Of-Mouth Communication Using Social Networking Sites. *Information & Management*, 53(4), 528-540. <https://doi.org/10.1016/j.im.2015.12.002>
- Band, W. and Petouhoff, N. L. (2009). *Topic Overview: Social CRM Goes Mainstream*. Forrester Research Report, No 55884.
- Bearden, W. O. (1983). Profiling Consumers Who Register Complaints Against Auto Repair Services. *The Journal of Consumer Affairs*, 17 (2 Winter), 315-35. <https://doi.org/10.1111/j.1745-6606.1983.tb00306.x>
- Bearden, W. O. and R. L. Oliver (1985). The Role of Public and Private Complaining in Satisfaction with Problem Resolution. *The Journal of Consumer Affairs*, 19 (2), 222-40. <https://doi.org/10.1111/j.1745-6606.1985.tb00353.x>
- Bearden, W.O. and Teel, J.E. (1983). Selected Determinants of Consumer Satisfaction and Complaint Reports. *Journal of Marketing Research*, 20, 21-8. <https://doi.org/10.1177/002224378302000103>
- Belkin, M., Niyogi, P., & Sindhvani, V. (2006). Manifold Regularization: A Geometric Framework for Learning from Labeled and Unlabeled Examples. *Journal of Machine Learning Research*, 7(Nov), 2399-2434.
- Bernoff, J., & Li, C. (2008). *Groundswell: Winning in A World Transformed by Social Technologies*. Boston: Harvard Business School Publishing.
- Berthon, P. R., Pitt, L. F., Plangger, K., & Shapiro, D. (2012). Marketing Meets Web 2.0, Social Media, And Creative Consumers: Implications for International Marketing Strategy. *Business Horizons*, 55(3), 261-271. <https://doi.org/10.1016/j.bushor.2012.01.007>
- Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The Fuzzy C-Means Clustering Algorithm. *Computers & Geosciences*, 10(2-3), 191-203.

- Bigne, E., Oltra, E., & Andreu, L. (2019). Harnessing Stakeholder Input on Twitter: A Case Study of Short Breaks in Spanish Tourist Cities. *Tourism Management*, 71, 490-503. <https://doi.org/10.1016/j.tourman.2018.10.013>
- Bilro, R. G., Loureiro, S. M. C., & Guerreiro, J. (2019). Exploring Online Customer Engagement with Hospitality Products and Its Relationship with Involvement, Emotional States, Experience and Brand Advocacy. *Journal of Hospitality Marketing & Management*, 28(2), 147-171. <https://doi.org/10.1080/19368623.2018.1506375>
- Bitner, M. J., Booms, B. H., & Tetreault, M. S. (1990). The Service Encounter: Diagnosing Favorable and Unfavorable Incidents. *Journal of Marketing*, 54(1), 71-84. <https://doi.org/10.1177/002224299005400105>
- Block, J. (1995). A Contrarian View of The Five-Factor Approach to Personality Description. *Psychological Bulletin*, 117, 187-215.
- Blodgett, J. G., & Anderson, R. D. (2000). A Bayesian Network Model of The Consumer Complaint Process. *Journal of Service Research*, 2(4), 321-338. <https://doi.org/10.1177/109467050024002>
- Blodgett, J. G., D. H. Granbois, and R. G. Walters (1993). The Effects of Perceived Justice on Complainants Negative Word-of-Mouth Behavior and Repatronage Intentions. *Journal of Retailing*, 69 (4), 399-428.
- Blodgett, J. G., K. L. Wakefield, and J. H. Barnes (1995). The Effects of Customer Service on Consumer Complaining Behavior. *Journal of Services Marketing*, 9 (4), 31-42. <https://doi.org/10.1108/08876049510094487>
- Blodgett, J.G. and Granbois, D.H. (1992). Toward an Integrated Conceptual Model of Consumer Complaint Behavior. *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behaviour*, 5, 93-103.
- Bolfing, C. P. (1989). How Do Consumers Express Dissatisfaction and What Can Service Marketers Do About It? *Journal of Services Marketing*, 3 (Spring), 5-23. <https://doi.org/10.1108/EUM0000000002483>

- Bolfing, C.P. (1989). How Do Customers Express Dissatisfaction and What Can Service Marketers Do About It? *Journal of Services Marketing*, 3(2), 5-23. <https://doi.org/10.1108/EUM00000000002483>
- Bolton, R. N., Gustafsson, A., McColl-Kennedy, J., Sirianni, N. J., & Tse, D. K. (2014). Small Details That Make Big Differences: A Radical Approach to Consumption Experience as A Firm's Differentiating Strategy. *Journal of Service Management*, 25(2), 253-274. <https://doi.org/10.1108/JOSM-01-2014-0034>
- Borders, B. (2009). A Brief History of Social Media. Retrieved November 31, 2019, from <http://socialmediarockstar.com/history-of-social-media>
- Boshoff, C. (1999). RECOVSAT: An Instrument to Measure Satisfaction with Transaction-Specific Service Recovery. *Journal of Service Research*, 1(3), 236-249. <https://doi.org/10.1177/109467059913005>
- Bougie, R., Pieters, R., & Zeelenberg, M. (2003). Angry Customers Don't Come Back, They Get Back: The Experience and Behavioral Implications of Anger and Dissatisfaction in Services. *Journal of The Academy of Marketing Science*, 31(4), 377-393. <https://doi.org/10.1177/0092070303254412>
- Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, And Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Breidbach, C., Brodie, R., & Hollebeek, L. (2014). Beyond Virtuality: From Engagement Platforms to Engagement Ecosystems. *Managing Service Quality*, 24(6), 592-611. <https://doi.org/10.1108/MSQ-08-2013-0158>
- Bruce, M., & Solomon, M. R. (2013). Managing for Media Anarchy: A Corporate Marketing Perspective. *Journal of Marketing Theory and Practice*, 21(3), 307-318. <https://doi.org/10.2753/MTP1069-6679210305>
- Bughin, J., & Chui, M. (2010). The Rise of The Networked Enterprise: Web 2.0 Finds Its Payday. *Mckinsey Quarterly*, 4, 3-8.
- Buzzient. (n.d.). Buzzient Features. Retrieved November 11, 2019, from <http://www.buzzient.com/features/>

- Calheiros, A. C., Moro, S., & Rita, P. (2017). Sentiment Classification of Consumer-Generated Online Reviews Using Topic Modeling. *Journal of Hospitality Marketing & Management*, 26(7), 675-693. <https://doi.org/10.1080/19368623.2017.1310075>
- Camastra, F., & Vinciarelli, A. (2015). *Machine Learning for Audio, Image and Video Analysis: Theory and Applications*. 2nd ed. London: Springer.
- Cambra-Fierro, J., Melero-Polo, I., & Sese, F. J. (2016). Can Complaint-Handling Efforts Promote Customer Engagement? *Service Business*, 10(4), 847-866. <https://doi.org/10.1007/s11628-015-0295-9>
- Cattell, R. B. (1957) Handbook for The IPAT Anxiety Scale Questionnaire (Self-Analysis Form): Brief Verbal Questionnaire, Q-Form, As Distinct from Objective T-Battery. Savoy, IL: Institute for Personality and Ability Testing.
- Cattell, R.B. (1943) The Description of Personality: Basic Traits Resolved into Clusters. *Journal of Abnormal Social Psychology*, 38, 476 – 506.
- Chaiken, S. (1987). The Heuristic Model of Persuasion. In *Social Influence: The Ontario Symposium*, 5, 3-39.
- Champoux, V., Durgee, J., & McGlynn, L. (2012). Corporate Facebook Pages: When “Fans” Attack. *Journal of Business Strategy*, 33(2), 22-30. <https://doi.org/10.1108/02756661211206717>
- Chau, M., & Xu, J. (2012). Business Intelligence in Blogs: Understanding Consumer Interactions and Communities. *MIS Quarterly*, 36(4), 1189-1216. <https://doi.org/10.2307/41703504>
- Chesbrough, H. (2007). Business Model Innovation: It's Not Just About Technology Anymore. *Strategy & Leadership*, 35(6), 12-17. <https://doi.org/10.1108/10878570710833714>
- Christodoulides, G., Jevon, C., & Bonhomme, J. (2012). Memo to Marketers: Quantitative Evidence for Change: How User-Generated Content Really Affects Brands. *Journal of Advertising Research*, 52 (1), 53–64. <https://doi.org/10.2501/JAR-52-1-053-064>

- Connor-Smith, J. K., & Flachsbart, C. (2007). Relations Between Personality and Coping: A Meta-Analysis. *Journal of Personality and Social Psychology*, 93(6), 1080–1107. <https://doi.org/10.1037/0022-3514.93.6.1080>
- Cook, S. (2012). *Complaint Management Excellence: Creating Customer Loyalty Through Service Recovery*. London: Kogan Page Publishers.
- Cooke, M., & Buckley, N. (2008). Web 2.0, Social Networks and The Future of Market Research. *International Journal of Market Research*, 50(2), 267-292.
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273-297. <https://doi.org/10.1007/BF00994018>
- Costa, A., Guerreiro, J., Moro, S., & Henriques, R. (2019). Unfolding the Characteristics of Incentivized Online Reviews. *Journal of Retailing and Consumer Services*, 47, 272-281. <https://doi.org/10.1016/j.jretconser.2018.12.006>
- Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO P–R) and NEO Five-Factor Inventory (NEO–FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources
- Crosier, K. and B. Z. Erdogan (2001). Advertising Complainants: Who and Where Are They? *Journal of Marketing Communications*, 7 (2), 109-20. <https://doi.org/10.1080/13527260121943>
- Crosier, K., T. Hernandez, S. Mohabir-Collins, and B. Z. Erdogan (1999). The Risk of Collateral Damage in Advertising Campaignes. *Journal of Marketing Management*, 15 (8), 837-55. <https://doi.org/10.1362/026725799784772701>
- D'Avanzo, E., Pilato, G., & Lytras, M. (2017). Using Twitter Sentiment and Emotions Analysis of Google Trends for Decisions Making. Program? *Electronic Library and Information Systems*, 51(3), 322-350. <https://doi.org/10.1108/PROG-02-2016-0015>

- David, J. P., & Suls, J. (1999). Coping Efforts in Daily Life: Role of Big Five Traits and Problem Appraisals. *Journal of Personality*. <https://doi.org/10.1111/1467-6494.00056>
- Davidow, M. (2000). The Bottom-Line Impact of Organizational Responses to Customer Complaints. *Journal of Hospitality & Tourism Research*, 24(4), 473-490. <https://doi.org/10.1177/109634800002400404>
- Davidow, M. (2003). Organizational Responses to Customer Complaints: What Works and What Doesn't. *Journal of Service Research*, 5(3), 225-250. <https://doi.org/10.1177/1094670502238917>
- Davidow, M., & Dacin, P. A. (1997). Understanding and Influencing Consumer Complaint Behavior: Improving Organizational Complaint Management. *ACR North American Advances*, 24, 450-456.
- Day, R. L. (1984). *Modeling Choices Among Alternative Responses to Dissatisfaction*. In *Advances in Consumer Research*, Thomas C. Kinnear, Ed. Vol. II. Provo, UT: Association for Consumer Research.
- Day, R. L., Grabicke, K., Schaetzle, T., & Staubach, F. (1981). The Hidden Agenda Of Consumer Complaining. *Journal of Retailing*, 57(3), 86–106.
- Day, R.L. & Landon, E.L. Jr. (1977). *Toward A Theory of Consumer Complaining Behavior*. In *Consumer and Industrial Buying Behavior*. Edited by A.G. Woodside, I.N. Sheth, & P.D. Bennet. New York: North-Holland, 425-437.
- Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Swaim, R. C. (2002). The Driving Anger Expression Inventory: A Measure of How People Express Their Anger on The Road. *Behaviour Research and Therapy*, 40(6), 717-737. [https://doi.org/10.1016/S0005-7967\(01\)00063-8](https://doi.org/10.1016/S0005-7967(01)00063-8)
- Deighton, J., & Kornfeld, L. (2010). Netzwerk: Digital Marketing in the Music Industry. HBS Case, (510-055).
- DeWitt, T., & Brady, M. K. (2003). Rethinking Service Recovery Strategies: The Effect of Rapport on Consumer Responses to Service Failure. *Journal of Service Research*, 6(2), 193-207. <https://doi.org/10.1177/1094670503257048>

- DeWitt, T., Nguyen, D. T., & Marshall, R. (2008). Exploring Customer Loyalty Following Service Recovery: The Mediating Effects of Trust and Emotions. *Journal of Service Research*, 10(3), 269-281. <https://doi.org/10.1177/1094670507310767>
- Dey, A. (2016). Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technologies*, 7(3), 1174-1179.
- Dhaoui, C., Webster, C. M., & Tan, L. P. (2017). Social Media Sentiment Analysis: Lexicon Versus Machine Learning. *Journal of Consumer Marketing*, 34(6), 480-488. <https://doi.org/10.1108/JCM-03-2017-2141>
- Digman, J. M. (1990). Personality Structure: Emergence of The Five Factors Model. *Annual Review of Psychology*, 41, 417-440.
- Doerr, N., Mattoni, A., & Teune, S. (2013). *Toward A Visual Analysis of Social Movements, Conflict, And Political Mobilization*. In *Advances in The Visual Analysis of Social Movements*. Emerald Group Publishing Limited.
- Donoghue, S., & de Klerk, H. M. (2006). Dissatisfied Consumers' Complaint Behaviour Concerning Product Failure of Major Electrical Household Appliances—A Conceptual Framework. *Journal of Consumer Sciences*, 34, 41-55.
- Ebner, M. (2009). Introducing Live Microblogging: How Single Presentations Can Be Enhanced by The Mass. *Journal of Research in Innovative Teaching*, 2(1), 91-100.
- Eddy, S. R. (1996). Hidden Markov Models. *Current Opinion in Structural Biology*, 6(3), 361-365.
- Edosomwan, S., Prakasan, S. K., Kouame, D., Watson, J., & Seymour, T. (2011). The History of Social Media and Its Impact on Business. *Journal of Applied Management and Entrepreneurship*, 16(3), 79-91.
- Edvardsson, B., Enquist, B. and Johnston, R. (2005). Cocreating Customer Value Through Hyperreality In the Prepurchase Service Experience. *Journal of Service Research*, 8(2), 149-61. <https://doi.org/10.1177/1094670505279729>
- Einwiller, S. A., & Steilen, S. (2015). Handling Complaints on Social Network Sites—An Analysis of Complaints and Complaint Responses on Facebook And Twitter Pages

- of Large US Companies. *Public Relations Review*, 41(2), 195-204. <https://doi.org/10.1016/j.pubrev.2014.11.012>
- Ekinci, Y., & Dawes, P. L. (2009). Consumer Perceptions of Frontline Service Employee Personality Traits, Interaction Quality, And Consumer Satisfaction. *The Service Industries Journal*, 29(4), 503-521. <https://doi.org/10.1080/02642060802283113>
- Ekinci, Y., Calderon, J., & Siala, H. (2016). Do Personality Traits Predict Complaining Consumers? *International Journal of Business Environment*, 8(1), 32-42. <https://doi.org/10.1504/IJBE.2016.074793>
- El Naqa, I., & Murphy, M. J. (2015). *What Is Machine Learning?* In Machine Learning in Radiation Oncology. Springer, Cham.
- Enge, E. (2014), "Twitter Engagement Unmasked: A Study of More Than 4M Tweets", Available At: www.stonetemple.com/twitter-engagement-unmasked/
- Ennew, C., & Schoefer, K. (2003). Service Failure and Service Recovery in Tourism: A Review. Retrieved January 8, 2021, from https://www.researchgate.net/profile/Christine_Ennew/publication/252577031_Service_Failure_and_Service_Recovery_in_Tourism_A_Review/links/00b4953604d8999e28000000/Service-Failure-and-Service-Recovery-in-Tourism-A-Review.pdf
- Eslami, S. P., & Ghasemaghaei, M. (2018). Effects of Online Review Positiveness and Review Score Inconsistency on Sales: A Comparison by Product Involvement. *Journal of Retailing and Consumer Services*, 45, 74-80. <https://doi.org/10.1016/j.jretconser.2018.08.003>
- Estelami, H. (2000). Competitive and Procedural Determinants of Delight and Disappointment in Consumer Complaint Outcomes. *Journal of Service Research*, 2(3), 285-300. <https://doi.org/10.1177/109467050023006>
- Etzel, M. J., & Silverman, B. I. (1981). A Managerial Perspective on Directions for Retail Customer Dissatisfaction Research. *Journal of Retailing*, 57(3), 124–136.
- Faase, R., Helms, R., & Spruit, M. (2011). Web 2.0 In the CRM Domain: Defining Social CRM. *International Journal of Electronic Customer Relationship Management*, 5(1), 1-22. <https://doi.org/10.1504/IJECRM.2011.039797>

- Feick, L. F., & Price, L. L. (1987). The Market Maven: A Diffuser of Marketplace Information. *Journal of Marketing*, 51(1), 83-97. <https://doi.org/10.1177/002224298705100107>
- Fiske, D.S. 1949. Consistency of The Factorial Structures of Personality Ratings from Different Sources. *Journal of Abnormal Social Psychology*, 44,329–440.
- Fitness, J. (2000). Anger in the Workplace: An Emotion Script Approach to Anger Episodes Between Workers and Their Superiors, Co-Workers and Subordinates. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 21(2), 147-162. [https://doi.org/10.1002/\(SICI\)1099-1379\(200003\)21:2<147::AID-JOB35>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1099-1379(200003)21:2<147::AID-JOB35>3.0.CO;2-T)
- Flach, P. (2012). *Machine Learning: The Art and Science of Algorithms That Make Sense of Data*. Cambridge: Cambridge University Press.
- Folkes, V. S., S. Koletsky, and J. L. Graham (1987). A Field Study of Causal Inferences and Consumer Reaction: The View from the Airport. *Journal of Consumer Research*, 13 (4), 534-39. <https://doi.org/10.1086/209086>
- Fornell, C. and N. M. Didow (1980). Economic Constraints on Consumer Complaining Behavior. *Advances in Consumer Research*, 7 (1), 318-23.
- Fornell, C., & Westbrook, R. A. (1979). An Exploratory Study of Assertiveness, Aggressiveness, and Consumer Complaining Behavior. *ACR North American Advances*.
- Francken, D. A. and W. F. van Raaij (1985). Socio-Economic and Demographic Determinants of Consumer Problem Perception. *Journal of Consumer Policy*, 8 (3), 303-15. <https://doi.org/10.1007/BF00380387>
- Francken, D. A., & Van Raaij, W. F. (1985). Socio-Economic and Demographic Determinants of Consumer Problem Perception. *Journal of Consumer Policy*, 8(3), 303-314. <https://doi.org/10.1007/BF00380387>

- Fricker, R. D., & Schonlau, M. (2002). Advantages and Disadvantages of Internet Research Surveys: Evidence from The Literature. *Field Methods*, 14(4), 347-367. <https://doi.org/10.1177/152582202237725>
- Friedman, M. (1985) Consumer Boycotts in the United States, 1970-1980: Contemporary Events in Historical Perspective. *Journal of Consumer Affairs*, 5, 1-23.
- Gelb, B. D., & Sundaram, S. (2002). Adapting To “Word of Mouse”. *Business Horizons*, 45(4), 21-25. [https://doi.org/10.1016/S0007-6813\(02\)00222-7](https://doi.org/10.1016/S0007-6813(02)00222-7)
- Gelbrich, K., & Roschk, H. (2011). A Meta-Analysis of Organizational Complaint Handling and Customer Responses. *Journal of Service Research*, 14(1), 24-43. <https://doi.org/10.1177/1094670510387914>
- Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter Brand Sentiment Analysis: A Hybrid System Using N-Gram Analysis and Dynamic Artificial Neural Network. *Expert Systems with Applications*, 40(16), 6266-6282. <https://doi.org/10.1016/j.eswa.2013.05.057>
- Giglio, S., Bertacchini, F., Bilotta, E., & Pantano, P. (2019). Using Social Media to Identify Tourism Attractiveness in Six Italian Cities. *Tourism Management*, 72, 306-312. <https://doi.org/10.1016/j.tourman.2018.12.007>
- Gilly, M. C., & Gelb, B. D. (1982). Post-Purchase Consumer Processes and The Complaining Consumer. *Journal of Consumer Research*, 9(3), 323-328. <https://doi.org/10.1086/208927>
- Global Digital Report 2019. (2019). Retrieved January 07, 2021, from <https://wearesocial.com/global-digital-report-2019>
- Golbeck, J., Robles, C., & Turner, K. (2011). Predicting Personality with Social Media. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, 253-262.
- Goldberg, L. R. (1992). The Development of Markers for the Big-Five Factor Structure. *Psychological Assessment*, 4, 26–42.

- Golder, S. A., & Macy, M. W. (2014). Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*, 40, 129-152. <https://doi.org/10.1146/annurev-soc-071913-043145>
- Gou, L., Zhou, M. X., & Yang, H. (2014, April). Knowme And Shareme: Understanding Automatically Discovered Personality Traits from Social Media and User Sharing Preferences. In *Proceedings of The SIGCHI Conference on Human Factors in Computing Systems*, 955-964.
- Gountas, J., & Gountas, S. (2007). Personality Orientations, Emotional States, Customer Satisfaction, and Intention to Repurchase. *Journal of Business Research*, 60(1), 72-75. <https://doi.org/10.1016/j.jbusres.2006.08.007>
- Graham, J. W., & Keeley, M. (1992). Hirschman's Loyalty Construct. *Employee Responsibilities and Rights Journal*, 5(3), 191-200. <https://doi.org/10.1007/BF01385047>
- Grainer, M. (2003). *Customer care: The Multibillion Dollar Sinkhole: A Case of Customer Rage Unassuaged*. Alexandria, VA: Customer Care Alliance.
- Granbois, D., J. O. Summers, and G. L. Frazier (1977). *Correlates of Consumer Expectation and Complaining Behavior*. In *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, R. L. Day, Ed. Bloomington, IN: Indiana University Press.
- Graziano, W. G., & Tobin, R. M. (2009). *Agreeableness*. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (p. 46-61). The Guilford Press.
- Greenberg, P. (2009). *Social CRM Comes of Age*. Oracle White Paper.
- Greenberg, P. (2010). The Impact of CRM 2.0 On Customer Insight. *Journal of Business & Industrial Marketing*, 25(6), 410-419. <https://doi.org/10.1108/08858621011066008>
- Greenwood, B. N., & Gopal, A. (2015). Research Note—Tigerblood: Newspapers, Blogs, And the Founding of Information Technology Firms. *Information Systems Research*, 26(4), 812-828. <https://doi.org/10.1287/isre.2015.0603>

- Grover, D. (2011). Effective Customer Relationship Management Through E-CRM. *Viewpoint*, 2, 27-38.
- Gunarathne, P., Rui, H., & Seidmann, A. (2017). Whose and What Social Media Complaints Have Happier Resolutions? Evidence from Twitter. *Journal of Management Information Systems*, 34(2), 314–340. <https://doi.org/10.1080/07421222.2017.1334465>
- Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60-76.
- Hagel, J., III, & Armstrong, A.G. (1997). *Net Gain: Expanding Markets Through Virtual Communities*. Boston, MA: Harvard Business School Press.
- Harrigan, P., Soutar, G., Choudhury, M. M., & Lowe, M. (2015). Modelling CRM in A Social Media Age. *Australasian Marketing Journal (AMJ)*, 23(1), 27-37. <https://doi.org/10.1016/j.ausmj.2014.11.001>
- Harrington, P. (2012). *Machine Learning in Action*. Manning Publications Co.
- Harris, E. G., & Mowen, J. C. (2001). The Influence of Cardinal-, Central-, And Surface-Level Personality Traits on Consumers' Bargaining and Complaint Intentions. *Psychology & Marketing*, 18(11), 1155-1185. <https://doi.org/10.1002/mar.1048>
- Hart, C. W., Heskett, J. L., & Sasser, J. W. (1990). The Profitable Art of Service Recovery. *Harvard Business Review*, 68(4), 148-156.
- Hatzivassiloglou, V., & McKeown, K. (1997). Predicting the Semantic Orientation of Adjectives. In *35th Annual Meeting of The Association for Computational Linguistics And 8th Conference of The European Chapter of The Association for Computational Linguistics*, 174-181.
- Hawkins, J. B., Brownstein, J. S., Tuli, G., Runels, T., Broecker, K., Nsoesie, E. O. & Greaves, F. (2016). Measuring Patient-Perceived Quality of Care in US Hospitals Using Twitter. *BMJ Qual Saf*, 25(6), 404-413. <http://dx.doi.org/10.1136/bmjqs-2015-004309>

- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley.
- Heggde, G., & Shainesh, G. (Eds.). (2018). *Social Media Marketing: Emerging Concepts and Applications*. New York: Springer.
- Heng, Y., Gao, Z., Jiang, Y., & Chen, X. (2018). Exploring Hidden Factors Behind Online Food Shopping from Amazon Reviews: A Topic Mining Approach. *Journal of Retailing and Consumer Services*, 42, 161-168. <https://doi.org/10.1016/j.jretconser.2018.02.006>
- Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic Word-Of-Mouth Via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on The Internet? *Journal of Interactive Marketing*, 18(1), 38-52. <https://doi.org/10.1002/dir.10073>
- Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., & Skiera, B. (2010). The Impact of New Media on Customer Relationships. *Journal of Service Research*, 13(3), 311–330. <https://doi.org/10.1177/1094670510375460>
- Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2015). Does Twitter Matter? The Impact of Microblogging Word of Mouth on Consumers' Adoption of New Movies. *Journal of the Academy of Marketing Science*, 43(3), 375-394. <https://doi.org/10.1007/s11747-014-0388-3>
- Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hogarth, J. M. and M. P. English (2002). Consumer Complaints and Redress: An Important Mechanism for Protection and Empowering Consumers. *International*

- Journal of Consumer Studies*, 26 (3), 217-26. <https://doi.org/10.1046/j.1470-6431.2002.00239.x>
- Homburg, C., & Fürst, A. (2005). How Organizational Complaint Handling Drives Customer Loyalty: An Analysis of the Mechanistic and the Organic Approach. *Journal of Marketing*, 69(3), 95–114. <https://doi.org/10.1509/jmkg.69.3.95.66367>
- Hu, X., & Liu, H. (2012). Text Analytics in Social Media. In *Mining Text Data* (pp. 385-414). Springer, Boston, MA: Springer.
- Hu, Y. H., Shiau, W. M., Shih, S. P., & Chen, C. J. (2018). Considering Online Consumer Reviews to Predict Movie Box-Office Performance Between the Years 2009 And 2014. *The US. The Electronic Library*, 36(6), 1010-1026. <https://doi.org/10.1108/EL-02-2018-0040>
- Huang, C. Y., Shen, Y. Z., Lin, H. X., & Chang, S. S. (2007). Bloggers' Motivations and Behaviors: A Model. *Journal of Advertising Research*, 47(4), 472–484. <https://doi.org/10.2501/S0021849907070493>
- Hung, M. C., & Yang, D. L. (2001). An Efficient Fuzzy C-Means Clustering Algorithm. *Proceedings 2001 IEEE International Conference on Data Mining*, 225-232. <https://doi.org/10.1109/ICDM.2001.989523>
- Hustad, T. P. and E. A. Pessemier (1973). Will the Real Consumer Activist Please Stand Up: An Examination of Consumers' Opinions About Marketing Practices. *Journal of Marketing Research*, 10 (August), 319-24. <https://doi.org/10.1177/002224377301000316>
- Ikeda, K., Hattori, G., Ono, C., Asoh, H., & Higashino, T. (2013). Twitter User Profiling Based on Text and Community Mining for Market Analysis. *Knowledge-Based Systems*, 51, 35-47. <https://doi.org/10.1016/j.knosys.2013.06.020>
- Ilhan, B. E., Kübler, R. V., & Pauwels, K. H. (2018). Battle of The Brand Fans: Impact of Brand Attack and Defense on Social Media. *Journal of Interactive Marketing*, 43, 33-51. <https://doi.org/10.1016/j.intmar.2018.01.003>

- Istanbulluoglu, D. (2017). Complaint Handling on Social Media: The Impact of Multiple Response Times on Consumer Satisfaction. *Computers in Human Behavior*, 74, 72-82. <https://doi.org/10.1016/j.chb.2017.04.016>
- Jacoby, J. and Jaccard, J. (1981). The Sources, Meaning, and Validity of Consumer Complaint Behavior: A Psychological Analysis. *Journal of Retailing*, 57(3), 4- 24.
- Jang, H. J., Sim, J., Lee, Y., & Kwon, O. (2013). Deep Sentiment Analysis: Mining the Causality Between Personality-Value-Attitude for Analyzing Business Ads in Social Media. *Expert Systems with Applications*, 40(18), 7492-7503. <https://doi.org/10.1016/j.eswa.2013.06.069>
- Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter Power: Tweets as Electronic Word of Mouth. *Journal of The American Society for Information Science and Technology*, 60(11), 2169-2188. <https://doi.org/10.1002/asi.21149>
- Jimenez-Marquez, J. L., Gonzalez-Carrasco, I., Lopez-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). Towards A Big Data Framework for Analyzing Social Media Content. *International Journal of Information Management*, 44, 1-12. <https://doi.org/10.1016/j.ijinfomgt.2018.09.003>
- John, O. J., & Srivastava, S. (1999). *The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives*. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (pp. 102–138). New York: Guilford.
- Kaplan, A. M., & Haenlein, M. (2009). The Fairyland of Second Life: Virtual Social Worlds and How to Use Them. *Business Horizons*, 52(6), 563-572. <https://doi.org/10.1016/j.bushor.2009.07.002>
- Kaplan, A. M., & Haenlein, M. (2010). Users of The World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons*, 53(1), 59-68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in Social Media Research: Past, Present and Future. *Information Systems Frontiers*, 20(3), 531-558. <https://doi.org/10.1007/s10796-017-9810-y>

- Keaveney, S. M. (1995). Customer Switching Behavior in Service Industries: An Exploratory Study. *Journal of Marketing*, 59(2), 71-82. <https://doi.org/10.1177/002224299505900206>
- Keim, M. E., & Noji, E. (2011). Emergent Use of Social Media: A New Age of Opportunity for Disaster Resilience. *American Journal of Disaster Medicine*, 6(1), 47-54.
- Kelleher, J. D., Mac Namee, B., & D'arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, And Case Studies*. S.I.: MIT Press.
- Kelley, S. W., & Davis, M. A. (1994). Antecedents to Customer Expectations for Service Recovery. *Journal of The Academy of Marketing Science*, 22(1), 52-61. <https://doi.org/10.1177/0092070394221005>
- Keng, K. A., Richmond, D., & Han, S. (1995). Determinants of Consumer Complaint Behaviour: A Study of Singapore Consumers. *Journal of International Consumer Marketing*, 8(2), 59-76. https://doi.org/10.1300/J046v08n02_05
- Khan, A., Baharudin, B., Lee, L. H., & Khan, K. (2010). A Review of Machine Learning Algorithms for Text-Documents Classification. *Journal of Advances in Information Technology*, 1(1), 4-20. <https://doi.org/10.4304/jait.1.1.4-20>
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media. *Business Horizons*, 54(3), 241-251. <https://doi.org/10.1016/j.bushor.2011.01.005>
- Kim, Y., Dwivedi, R., Zhang, J., & Jeong, S. R. (2016). Competitive Intelligence in Social Media Twitter: Iphone 6 Vs. Galaxy S5. *Online Information Review*, 40(1), 42-61. <https://doi.org/10.1108/OIR-03-2015-0068>
- Klein, J., & Dawar, N. (2004). Corporate Social Responsibility and Consumers' Attributions and Brand Evaluations in A Product–Harm Crisis. *International Journal of research in Marketing*, 21(3), 203-217. <https://doi.org/10.1016/j.ijresmar.2003.12.003>

- Klostermann, J., Plumeyer, A., Böger, D., & Decker, R. (2018). Extracting Brand Information from Social Networks: Integrating Image, Text, And Social Tagging Data. *International Journal of Research in Marketing*, 35(4), 538-556. <https://doi.org/10.1016/j.ijresmar.2018.08.002>
- Kolchyna, O., Souza, T. T., Treleven, P., & Aste, T. (2015). Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination. *Arxiv Preprint arXiv:1507.00955*.
- Kolodinsky, J. (1993). Complaints, Redress and Subsequent Purchases of Medical Services by Dissatisfied Consumers. *Journal of Consumer Policy*, 16 (2), 193-214. <https://doi.org/10.1007/BF01418376>
- Kolodinsky, J. (1995). Usefulness of Economics and Explaining Consumer Complaints. *Journal of Consumer Affairs*, 29 (1), 29-54. <https://doi.org/10.1111/j.1745-6606.1995.tb00038.x>
- Komunda, M. and Osarenkhoe, A. (2012), Remedy or Cure for Service Failure? Effects of Service Recovery on Customer Satisfaction and Loyalty. *Business Process Management Journal*, 18(1), 82-103. <https://doi.org/10.1108/14637151211215028>
- Kotsiantis, S. B. (2007). *Supervised Machine Learning: A Review of Classification Techniques*. In Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in Ehealth, HCI, Information Retrieval and Pervasive Technologies, 3–24, Amsterdam, The Netherlands: IOS Press.
- Kowalski, R. M., & Cantrell, C. C. (1995). Psychometric Properties of The Propensity to Complain Scale. *Unpublished Manuscript, Western Carolina University, Cullowhee, NC*, 33.
- Kowalski, R.M. (1996). Complaints and Complaining: Functions, Antecedents, And Consequences. *Psychological Bulletin*, 119, 179-96.
- Kraft, F. B. (1977). *Characteristics of Consumer Complainers and Complaint Repatronage Behavior*. In Consumer Satisfaction, Dissatisfaction and Complaining

Behavior Proceedings, Ralph L. Day and H. Keith Hunt, Ed. Bloomington: Indiana University.

Kumar, A., & Sebastian, T. M. (2012). Sentiment Analysis: A Perspective on Its Past, Present and Future. *International Journal of Intelligent Systems and Applications*, 4(10), 1-14. <https://doi.org/10.5815/ijisa.2012.10.01>

Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T., & Tillmanns, S. (2010). Undervalued or Overvalued Customers: Capturing Total Customer Engagement Value. *Journal of Service Research*, 13(3), 297-310. <https://doi.org/10.1177/1094670510375602>

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What Is Twitter, A Social Network or A News Media? In *Proceedings of the 19th International Conference On World Wide Web*, 591-600. <https://doi.org/10.1145/1772690.1772751>

Kwok, L., & Yu, B. (2013). Spreading Social Media Messages on Facebook: An Analysis of Restaurant Business-To-Consumer Communications. *Cornell Hospitality Quarterly*, 54(1), 84-94. <https://doi.org/10.1177/1938965512458360>

Lahuerta-Otero, E., Cordero-Gutiérrez, R., & De la Prieta-Pintado, F. (2018). Retweet or like? That is the question. *Online Information Review*, 42(5), 562-578. <https://doi.org/10.1108/OIR-04-2017-0135>

Lam, T., & Zhang, H. Q. (1999). Service Quality of Travel Agents: The Case of Travel Agents in Hong Kong. *Tourism Management*, 20(3), 341-349. [https://doi.org/10.1016/S0261-5177\(98\)00118-6](https://doi.org/10.1016/S0261-5177(98)00118-6)

Landon, E. L. (1977). *A Model of Consumer Complaint Behavior*. In *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, Ralph L. Day, Ed. Bloomington, IN: Indiana University Department of Marketing.

Landon, E. L. (1980). The Direction of Consumer Complaint Research. *Advances in Consumer Research*, 7, 335-38.

Landon, Jr. (1977). A Model of Consumer Complaint Behavior, in *Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*. Ralph L. Day and H. Keith Hunt, Eds., Bloomington: Indiana University.

- Lapidus, R. S., & Pinkerton, L. (1995). Customer Complaint Situations: An Equity Theory Perspective. *Psychology & Marketing*, 12(2), 105-122. <https://doi.org/10.1002/mar.4220120203>
- Lazarus, R. (1991). *Stress and Coping: An Anthology*. New York: Columbia University Press.
- Lazarus, R. and S. Folkman (1984). *Stress, Appraisal and Coping*. New York: Springer.
- Lee, D., Hosanagar, K., & Nair, H. S. (2018). Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook. *Management Science*, 64(11), 5105-5131. <https://doi.org/10.1287/mnsc.2017.2902>
- Lee, S. and Cude, B.J. (2011). Consumer Complaint Channel Choice in Online and Offline Purchases. *International Journal of Consumer Studies*, 36 (1), 90-96. <https://doi.org/10.1111/j.1470-6431.2010.00992.x>
- Lee, Y. L., & Song, S. (2010). An Empirical Investigation of Electronic Word-Of-Mouth: Informational Motive and Corporate Response Strategy. *Computers in Human Behavior*, 26(5), 1073-1080. <https://doi.org/10.1016/j.chb.2010.03.009>
- Lee, Y. L., & Song, S. (2010). An Empirical Investigation of Electronic Word-Of-Mouth: Informational Motive and Corporate Response Strategy. *Computers in Human Behavior*, 26(5), 1073-1080. <https://doi.org/10.1016/j.chb.2010.03.009>
- Levy, M. (2009). Web 2.0 Implications on Knowledge Management. *Journal of Knowledge Management*, 13(1), 120-134. <https://doi.org/10.1108/13673270910931215>
- Lewis, B. K. (2010). Social Media and Strategic Communication: Attitudes and Perceptions Among College Students. *Public Relations Journal*, 4(3), 1–23.
- Li, W., Lee, B., Krausz, F., and Sahin, K. (1991). *Text Classification by a Neural Network*. In Proceedings of the 23rd Annual Summer Computer Simulation Conference. D. Pace, ed. Baltimore, Society for Computer Simulation, San Diego, CA.

- Liefeld, J. P., H. C. Edgecombe, and L. Wolfe (1975). Demographic Characteristics of Canadian Consumer Complainers. *Journal of Consumer Affairs*, 9 (Summer), 73-80. <https://doi.org/10.1111/j.1745-6606.1975.tb00550.x>
- Lin, W. B. (2010). Service Failure and Consumer Switching Behaviors: Evidence from The Insurance Industry. *Expert Systems with Applications*, 37(4), 3209-3218. <https://doi.org/10.1016/j.eswa.2009.09.059>
- Lin, Y. (2020, January 05). 10 Twitter Statistics Every Marketer Should Know in 2020 [Infographic]. Retrieved January 07, 2021, from <https://www.oberlo.com/blog/twitter-statistics>
- Littman, M. L., & Szepesvári, C. (1996). *A Generalized Reinforcement-Learning Model: Convergence and Applications*. In ICML, 96, 310-318.
- Liu, R. R. and P. McClure (2001). Recognizing Cross-Cultural Differences in Consumer Complaint Behavior and Intentions: An Empirical Examination. *Journal of Consumer Marketing*, 18 (1), 54-74. <https://doi.org/10.1108/07363760110365813>
- Liu, R. R., H. S. Watkins, and Y. Yi (1997). Taxonomy of Consumer Complaint Behaviour: Replication and Extension. *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behaviour*, 10 (92-103).
- Liu, X., Singh, P. V., & Srinivasan, K. (2016). A Structured Analysis of Unstructured Big Data by Leveraging Cloud Computing. *Marketing Science*, 35(3), 363-388. <https://doi.org/10.1287/mksc.2015.0972>
- Lombart, C., & Louis, D. (2012). Consumer Satisfaction and Loyalty: Two Main Consequences of Retailer Personality. *Journal of Retailing and Consumer Services*, 19(6), 644-652. <https://doi.org/10.1016/j.jretconser.2012.08.007>
- Lutz, R. J. (1975). Changing Brand Attitudes Through Modification of Cognitive Structure. *Journal of Consumer Research*, 1(4), 49-59. <https://doi.org/10.1086/208607>
- Ma, L., Sun, B., & Kekre, S. (2015). The Squeaky Wheel Gets the Grease—An Empirical Analysis of Customer Voice and Firm Intervention on Twitter. *Marketing Science*, 34(5), 627-645. <https://doi.org/10.1287/mksc.2015.0912>

- Malafi, T. N., M. A. Cini, S. L. Taub, and J. Bertoami (1993). Social Influence and The Decision to Complain: Investigations on The Role of Advice. *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behaviour*, 6, 81-89.
- Malhotra, N. K., & Malhotra, N. K. (2012). *Basic Marketing Research: Integration of Social Media*. Boston: Pearson.
- Malthouse, E. C., Calder, B. J., Kim, S. J., & Vandenberg, M. (2016). Evidence That User-Generated Content That Produces Engagement Increases Purchase Behaviours. *Journal of Marketing Management*, 32(5-6), 427-444. <https://doi.org/10.1080/0267257X.2016.1148066>
- Malthouse, E. C., Haenlein, M., Skiera, B., Wege, E., & Zhang, M. (2013). Managing Customer Relationships in the Social Media Era: Introducing the Social CRM House. *Journal of Interactive Marketing*, 27(4), 270-280. <https://doi.org/10.1016/j.intmar.2013.09.008>
- Mano, H., & Oliver, R. L. (1993). Assessing the Dimensionality and Structure of The Consumption Experience: Evaluation, Feeling, And Satisfaction. *Journal of Consumer research*, 20(3), 451-466. <https://doi.org/10.1086/209361>
- Maritz (2011). Maritz Research and Evolve24 – Twitter Study. Retrieved December 5, 2020, from <http://www.maritzresearch.com/~media/Files/MaritzResearch/e24/ExecutiveSummaryTwitterPoll.pdf>
- Markus, H. R. and S. Kitayama (1990). Culture and the Self: Implications for Cognition, Emotion and Motivation. *Psychological Review*, 98 (2), 224-53. <https://doi.org/10.1037/0033-295X.98.2.224>
- Martínez, A. M., & Kak, A. C. (2001). Pca versus Ida. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2), 228-233. <https://doi.org/10.1109/34.908974>
- Mason, J. B. and S. H. j. Himes (1973). An Explanatory Behavioral and Socioeconomic Profile of Consumer Action About Dissatisfaction with Selected Household

- Appliances. *Journal of Consumer Affairs*, 7 (Winter), 121-27.
<https://doi.org/10.1111/j.1745-6606.1973.tb00528.x>
- Mattila, A. S. (2001). The Effectiveness of Service Recovery in A Multi-Industry Setting. *Journal of Services Marketing*, 15(7), 583-596.
<https://doi.org/10.1108/08876040110407509>
- Mattila, A. S., & Mount, D. J. (2003). The Impact of Selected Customer Characteristics and Response Time On E-Complaint Satisfaction and Return Intent. *International Journal of Hospitality Management*, 22(2), 135-145. [https://doi.org/10.1016/S0278-4319\(03\)00014-8](https://doi.org/10.1016/S0278-4319(03)00014-8)
- Mattila, A. S., & Wirtz, J. (2004). Consumer Complaining to Firms: The Determinants of Channel Choice. *Journal of Services Marketing*, 18(2), 147-155.
<https://doi.org/10.1108/08876040410528746>
- Matzler, K., Faullant, R., Renzl, B., & Leiter, V. (2005). The Relationship Between Personality Traits (Extraversion and Neuroticism), Emotions and Customer Self-Satisfaction. *Innovative Marketing*, 1(2), 32-39.
- Maurer, C., & Schaich, S. (2011). Online Customer Reviews Used as Complaint Management Tool. *ENTER*, 499-511.
- McClendon, L., & Meghanathan, N. (2015). Using Machine Learning Algorithms to Analyze Crime Data. *Machine Learning and Applications: An International Journal (MLAIJ)*, 2(1), 1-12.
- McColl-Kennedy, J.R. and Smith, A.K. (2006). *Customer Emotions in Service Failure and Recovery Encounters*, Zerbe, W.J., Ashkanasy, N.M. and Härtel, C.E.J. (Ed.) Individual and Organizational Perspectives on Emotion Management and Display (Research on Emotion in Organizations, Vol. 2), Emerald Group Publishing Limited, Bingley, pp. 237-268. [https://doi.org/10.1016/S1746-9791\(06\)02010-4](https://doi.org/10.1016/S1746-9791(06)02010-4)
- McCrae, R. R. & John, O. P. (1992). An Introduction to The Five-Factor Model and Its Applications. *Journal of Personality*, 60, 175-215.

- McCrae, R. R. (2009). The Five-Factor Model of Personality Traits: Consensus and Controversy. In *the Cambridge Handbook of Personality Psychology*, pp.148-161, Cambridge: Cambridge University Press.
- Melancon, J. P., & Dalakas, V. (2018). Consumer Social Voice in The Age of Social Media: Segmentation Profiles and Relationship Marketing Strategies. *Business Horizons*, 61(1), 157–167. <https://doi.org/10.1016/j.bushor.2017.09.015>
- Mendes-Filho, L., & Tan, F. B. (2009). User-Generated Content and Consumer Empowerment in The Travel Industry: A Uses & Gratifications and Dual-Process Conceptualization. *PACIS 2009 Proceedings*, 28. <http://aisel.aisnet.org/pacis2009/28>
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication*, 60(3), 413-439. <https://doi.org/10.1111/j.1460-2466.2010.01488.x>
- Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (2013). *Machine Learning: An Artificial Intelligence Approach*. New York: Springer.
- Michalski, R., Carbonell, J., & Mitchell, T. (1985). Machine Learning: An Artificial Intelligence Approach. *Artificial Intelligence*, 25(2), 236–238.
- Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*. Amsterdam, The Netherlands: Academic Press.
- Mitchell, T. M. (2006). *The Discipline of Machine Learning*. Machine Learning Department Technical Report CMU-ML-06-108, Carnegie Mellon University.
- Mooradian, T. A., & Olver, J. M. (1997). "I Can't Get No Satisfaction:" The Impact of Personality and Emotion on Postpurchase Processes. *Psychology & Marketing*, 14(4), 379-393. [https://doi.org/10.1002/\(SICI\)1520-6793\(199707\)14:4<379::AID-MAR5>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1520-6793(199707)14:4<379::AID-MAR5>3.0.CO;2-6)
- Mooradian, T. A., & Swan, K. S. (2006). Personality-And-Culture: The Case of National Extraversion and Word-Of-Mouth. *Journal of Business Research*, 59(6), 778-785. <https://doi.org/10.1016/j.jbusres.2006.01.015>

- Moreno, A., & Redondo, T. (2016). Text Analytics: The Convergence of Big Data and Artificial Intelligence. *IJIMAI*, 3(6), 57-64.
- Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, 58(3), 20-38. <https://doi.org/10.1177/002224299405800302>
- Moro, S., Rita, P., & Vala, B. (2016). Predicting Social Media Performance Metrics and Evaluation of The Impact on Brand Building: A Data Mining Approach. *Journal of Business Research*, 69(9), 3341-3351. <https://doi.org/10.1016/j.jbusres.2016.02.010>
- Moyer, M. S. (1985). Characteristics of Consumer Complaints: Implications for Marketing and Public Policy. *Journal of Public Policy and Marketing*, 3 (1), 67-84. <https://doi.org/10.1177/074391568400300105>
- Murray, C.E. (2011). Marriott's Corporate Twitter Strategy. Retrieved January 27, 2020, from www.hotelinteractive.com/article.aspx?articleid=19614
- Musser, J., & O'Reilly, T. (2006). Web 2.0 Principles and Best Practices. (Electronic Version). O'Reilly Radar. Retrieved January 20, 2020, from http://www.oreilly.com/catalog/web2report/chapter/web20_report_except.pdf.
- Nave, M., Rita, P., & Guerreiro, J. (2018). A Decision Support System Framework to Track Consumer Sentiments in Social Media. *Journal of Hospitality Marketing & Management*, 27(6), 693-710. <https://doi.org/10.1080/19368623.2018.1435327>
- Nguyen, K. A., & Coudounaris, D. N. (2015). The Mechanism of Online Review Management: A Qualitative Study. *Tourism Management Perspectives*, 16, 163-175. <https://doi.org/10.1016/j.tmp.2015.08.002>
- Nguyen, K. A., & Coudounaris, D. N. (2015). The Mechanism of Online Review Management: A Qualitative Study. *Tourism Management Perspectives*, 16, 163-175. <https://doi.org/10.1016/j.tmp.2015.08.002>
- Nguyen, M. T., Tran, D. V., & Nguyen, L. M. (2018). Social Context Summarization Using User-Generated Content and Third-Party Sources. *Knowledge-Based Systems*, 144, 51-64. <https://doi.org/10.1016/j.knosys.2017.12.023>

- Nilashi, M., Ibrahim, O., Yadegaridehkordi, E., Samad, S., Akbari, E., & Alizadeh, A. (2018). Travelers Decision Making Using Online Review in Social Network Sites: A Case on Tripadvisor. *Journal of Computational Science*, 28, 168-179. <https://doi.org/10.1016/j.jocs.2018.09.006>
- Nilsson, N. J. (1982). *Principles of Artificial Intelligence (Symbolic Computation/Artificial Intelligence)*. New York: Springer.
- Nilsson, N. J. (2009). *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press.
- Norusis, M. J. (2007). *SPSS 15.0 Advanced Statistical Procedures Companion*. Chicago, IL: Prentice Hall.
- O'Shea, W. (2003, July 4-10). Six Degrees of Sexual Frustration: Connecting the Dots with Friendster.Com. Village Voice. Retrieved December 5, 2020, from <http://www.villagevoice.com/news/0323,oshea,44576,1.html>
- OECD (2007). Participative Web and User-Created Content: Web 2.0, Wikis, And Social Networking. Paris, France: Organization for Economic Co-Operation and Development. Retrieved December 7, 2020, from http://www.oecd.org/document/40/0,3343,en_2649_34223_39428648_1_1_1_1,00.html
- Oliver, R. L. (1980). A Cognitive Model of Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, 17 (November), 460- 69. <https://doi.org/10.1177/002224378001700405>
- Oliver, R. L. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*. New York: McGraw-Hill Companies Inc.
- Olivier, C., Bernhard, S., & Alexander, Z. (2006). Semi-Supervised Learning. *IEEE Transactions on Neural Networks*, 20(3), 542-542.
- Oracle. (2011). Social Media Connector for Siebel Customer Relationship Management. Retrieved November 11, 2019, from <http://www.oracle.com/us/social-connector-siebel-crm-ds-507941.pdf>

- Orsingher, C., Valentini, S., & De Angelis, M. (2010). A Meta-Analysis of Satisfaction with Complaint Handling in Services. *Journal of The Academy of Marketing Science*, 38(2), 169-186. <https://doi.org/10.1007/s11747-009-0155-z>
- Oster, S. (1980). The Determinants of Consumer Complaints. *Review of Economics and Statistics*, 62 (4), 603-09. <https://doi.org/10.2307/1924785>
- Ostrom, A. and D. Iacobucci (1995). Consumer Trade-Offs and The Evaluation of Services. *Journal of Marketing*, 59 (1), 17-28. <https://doi.org/10.1177/002224299505900102>
- Park, S. B., Kim, H. J., & Ok, C. M. (2018). Linking Emotion and Place on Twitter at Disneyland. *Journal of Travel & Tourism Marketing*, 35(5), 664-677. <https://doi.org/10.1080/10548408.2017.1401508>
- Parr, B. (2008). It's Time We Defined Social Media. No More Arguing. Here's the Definition. Retrieved December 5, 2020, from <http://benparr.com/2008/08/its-time-we-defined-social-media-no-more-arguing-heres-the-definition/>
- Parra-López, E., Bulchand-Gidumal, J., Gutiérrez-Taño, D., & Díaz-Armas, R. (2011). Intentions to Use Social Media in Organizing and Taking Vacation Trips. *Computers in Human Behavior*, 27(2), 640-654. <https://doi.org/10.1016/j.chb.2010.05.022>
- Passant, A., Hastrup, T., Bojars, U., & Breslin, J. (2008). *Microblogging: A Semantic Web and Distributed Approach*. Proc. SFSM.
- Peláez, J. I., Martínez, E. A., & Vargas, L. G. (2019). Decision Making in Social Media with Consistent Data. *Knowledge-Based Systems*, 172, 33-41. <https://doi.org/10.1016/j.knosys.2019.02.009>
- Pfaff, M. and S. Blivice Eds. (1977). *Socioeconomic Correlates of Consumer and Citizen Dissatisfaction and Activism*. Bloomington, IN: Bureau of Business Research.
- Pfeffer, J., Zorbach, T., & Carley, K. M. (2014). Understanding Online Firestorms: Negative Word-Of-Mouth Dynamics in Social Media Networks. *Journal of*

- Marketing Communications*, 20(1-2), 117-128.
<https://doi.org/10.1080/13527266.2013.797778>
- Phau, I. and Puspita Sari, R. (2004), Engaging in Complaint Behaviour: An Indonesian Perspective. *Marketing Intelligence & Planning*, 22(4), 407-426.
<https://doi.org/10.1108/02634500410542770>
- Piñeiro-Chousa, J., Vizcaíno-González, M., & Pérez-Pico, A. M. (2017). Influence of Social Media Over the Stock Market. *Psychology & Marketing*, 34(1), 101-108.
<https://doi.org/10.1002/mar.20976>
- Pinto, M. B., & Mansfield, P. (2012). Facebook As A Complaint Mechanism: An Investigation of Millennials. *Journal of Behavioral Studies in Business*, 5, 1.
- Plutchik, R. (1980). A General Psychoevolutionary Theory of Emotion. In *Theories of Emotion*, 3-33. Academic Press.
- Portugal, I., Alencar, P., & Cowan, D. (2018). The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review. *Expert Systems with Applications*, 97, 205-227. <https://doi.org/10.1016/j.eswa.2017.12.020>
- Pournarakis, D. E., Sotiropoulos, D. N., & Giaglis, G. M. (2017). A Computational Model for Mining Consumer Perceptions in Social Media. *Decision Support Systems*, 93, 98-110. <https://doi.org/10.1016/j.dss.2016.09.018>
- Presi, C., Saridakis, C., & Hartmans, S. (2014). User-Generated Content Behaviour Of the Dissatisfied Service Customer. *European Journal of Marketing*, 48(9), 1600-1625. <https://doi.org/10.1108/EJM-07-2012-0400>
- Quesenberry, K. A. (2018). *Social Media Strategy: Marketing, Advertising, and Public Relations in the Consumer Revolution*. Rowman & Littlefield.
- Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning*, 1(1), 81-106.
<https://doi.org/10.1007/BF00116251>
- Raacke, J., & Bonds-Raacke, J. (2008). Myspace And Facebook: Applying the Uses and Gratifications Theory to Exploring Friend-Networking Sites. *Cyberpsychology & Behavior*, 11(2), 169-174. <https://doi.org/10.1089/cpb.2007.0056>

- Raacke, J., & Bonds-Raacke, J. (2008). Myspace And Facebook: Applying the Uses and Gratifications Theory to Exploring Friend-Networking Sites. *Cyberpsychology & Behavior*, 11(2), 169-174. <https://doi.org/10.1089/cpb.2007.0056>
- Rabiner, L. R., & Juang, B. H. (1986). An Introduction to Hidden Markov Models. *IEEE ASSP Magazine*, 3(1), 4-16.
- Ramadhan, W. P., Novianty, S. A., & Setianingsih, S. C. (2017). Sentiment Analysis Using Multinomial Logistic Regression. In *2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, 46-49. <https://doi.org/10.1109/ICCEREC.2017.8226700>
- Rathan, M., Hulipalled, V. R., Venugopal, K. R., & Patnaik, L. M. (2018). Consumer Insight Mining: Aspect Based Twitter Opinion Mining of Mobile Phone Reviews. *Applied Soft Computing*, 68, 765-773. <https://doi.org/10.1016/j.asoc.2017.07.056>
- Riaz, Z., & Khan, M. I. (2016). Impact of Service Failure Severity and Agreeableness on Consumer Switchover Intention. *Asia Pacific Journal of Marketing and Logistics*, 28(3), 420-434. <https://doi.org/10.1108/APJML-07-2015-0106>
- Richins, M. L. (1979). *Consumer Complaining Process: A Comprehensive Model*. In: R.L. Day and H.K. Hunt (Eds), *New Dimensions of Consumer Satisfaction and Complaining Behavior*, Bloomington, IN: Indiana University Press.
- Richins, M. L. (1983b). An Analysis of Consumer Interaction Styles in the Marketplace. *Journal of Consumer Research*, 10 (1), 73-83. <https://doi.org/10.1086/208946>
- Richins, M. L. (1987). A Multivariate Analysis of Responses to Dissatisfaction. *Journal of the Academy of Marketing Science*, 15 (3), 24-32. <https://doi.org/10.1177/009207038701500303>
- Richins, M. L. and B. J. Verhage (1985). Cross-Cultural Differences in Consumer Attitudes and Their Implications for Complaint Management. *International Journal of Research in Marketing*, 2 (3), 197-206. [https://doi.org/10.1016/0167-8116\(85\)90012-6](https://doi.org/10.1016/0167-8116(85)90012-6)

- Richins, M. L., & Verhage, B. J. (1985). Cross-Cultural Differences in Consumer Attitudes and Their Implications for Complaint Management. *International Journal of Research in Marketing*, 2(3), 197-206. [https://doi.org/10.1016/0167-8116\(85\)90012-6](https://doi.org/10.1016/0167-8116(85)90012-6)
- Richins, M.L. (1983a). Negative Word-Of-Mouth by Dissatisfied Consumers: A Pilot Study. *Journal of Marketing*, 47(1), 68-78. <https://doi.org/10.1177/002224298304700107>
- Ritholz, B. (2010, December 03). History of Social Media. Retrieved January 07, 2021, from <http://www.ritholtz.com/blog/2010/12/history-of-social-media/>
- Robinson, L. M., G. Trebbi, and R. D. Adler (1982). *Taking It to The Top: A Profile of Customers Who Complain to the CEO*. In International Fare in Consumer Satisfaction and Complaining Behavior, R. Day and K. Hunt, Eds. Bloomington, Indiana: Indiana University.
- Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The Big Five Personality Factors and Personal Values. *Personality and Social Psychology Bulletin*, 28(6), 789-801. <https://doi.org/10.1177/0146167202289008>
- Rogers, A., Daunt, K. L., Morgan, P., & Beynon, M. (2017). Examining the Existence of Double Jeopardy and Negative Double Jeopardy Within Twitter. *European Journal of Marketing*, 51(7/8), 1224-1247. <https://doi.org/10.1108/EJM-03-2015-0126>
- Rogers, J. C. and T. G. Williams (1990). Consumer Personal Values as Antecedents to Dyadic and Third-Party Public Consumer Complaint Behaviour: An Exploratory Study. *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behaviour*, 3, 71-81.
- Rogers, S., & Girolami, M. (2016). *A First Course in Machine Learning*. Chapman and Hall/CRC.
- Rosenblatt, F. (1961). *Principles of Neurodynamics: Perceptrons And the Theory of Brain Mechanisms*. Buffalo NY: Cornell Aeronautical Lab Inc.
- Russell, S. J. (2003). *Artificial Intelligence: A Modern Approach* (2nd Edition ed.). Upper Saddle River, NJ, NJ, USA: Prentice Hall.

- Russom, P. (2007). BI Search and Text Analytics. *TDWI Best Practices Report*, 9-11.
- Ryoo, J., & Bendle, N. (2017). Understanding the Social Media Strategies of US Primary Candidates. *Journal of Political Marketing*, 16(3-4), 244-266. <https://doi.org/10.1080/15377857.2017.1338207>
- Samuel, A. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210-29. <https://doi.org/10.1147/rd.33.0210>
- Satish, S. M., & Bharadhwaj, S. (2010). Information Search Behaviour Among New Car Buyers: A Two-Step Cluster Analysis. *IIMB Management Review*, 22(1-2), 5-15. <https://doi.org/10.1016/j.iimb.2010.03.005>
- Schapiro, R. E. (1990). The Strength of Weak Learnability. *Machine Learning*, 5(2), 197-227.
- Scherer, K. R. (1984). On the Nature and Function of Emotion: A Component Process Approach. *Approaches to Emotion*, 2293(317), 31.
- Schiopu, D. (2010). Applying Twostep Cluster Analysis for Identifying Bank Customers' Profile. *Buletinul*, 62(3), 66-75.
- Schniederjans, D., Cao, E. S., & Schniederjans, M. (2013). Enhancing Financial Performance with Social Media: An Impression Management Perspective. *Decision Support Systems*, 55(4), 911-918. <https://doi.org/10.1016/j.dss.2012.12.027>
- Sekaran U. 2003. *Research Methods for Business: A Skill Building Approach*. Hoboken, NJ: John Wiley.
- Sen, S., Gürhan-Canli, Z., & Morwitz, V. (2001). Withholding Consumption: A Social Dilemma Perspective on Consumer Boycotts. *Journal of Consumer Research*, 28(3), 399-417. <https://doi.org/10.1086/323729>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press.
- Shaoul, C., Westbury, C. F., & Baayen, H. R. (2013). The Subjective Frequency of Word N-Grams. *Psihologija*, 46(4), 497-537.

- Shareef, M. A., Mukerji, B., Alryalat, M. A. A., Wright, A., & Dwivedi, Y. K. (2018). Advertisements on Facebook: Identifying the Persuasive Elements in The Development of Positive Attitudes in Consumers. *Journal of Retailing and Consumer Services*, 43, 258-268. <https://doi.org/10.1016/j.jretconser.2018.04.006>
- Shi, W. H., Liu, J. Y., & Zhang, Y. (2011). The Effect of Service Failure Attribution on Consumer Complaint Behavior: The Mediating Role of Negative Emotion. *The Journal of China Universities of Posts and Telecommunications*, 18, 169-173. [https://doi.org/10.1016/S1005-8885\(10\)60155-0](https://doi.org/10.1016/S1005-8885(10)60155-0)
- Shimp B (2009). Social CRM. Say What? Retrieved December 5, 2020, from <http://allbizanswers.com/social-crm-say-what/>
- Shiner, R. L., & DeYoung, C. G. (2013). *The Structure of Temperament and Personality Traits: A Developmental Perspective*. In P. D. Zelazo (Ed.), *Oxford Library of Psychology. The Oxford Handbook of Developmental Psychology, Vol. 2. Self and Other* (p. 113–141). Oxford University Press.
- Shokohyar, S., Tavalae, R., & Karamatnia, K. (2017). Identifying Effective Indicators in The Assessment of Organizational Readiness for Accepting Social CRM. *International Journal of Information, Business and Management*, 9(4), 209.
- Singh, J. (1990a). A Typology of Consumer Dissatisfaction Response Styles. *Journal of Retailing*, 66 (1), 57-99.
- Singh, J. (1990b). Voice, Exit and Negative Word-Of-Mouth Behaviors: An Investigation Across Three Service Categories. *Journal of the Academy of Marketing Science*, Vol. 18, pp. 1-15.
- Singh, J. (1991). Industry Characteristics and Consumer Dissatisfaction. *The Journal of Consumer Affairs*, 25 (1), 19-56. <https://doi.org/10.1111/j.1745-6606.1991.tb00279.x>
- Singh, J. and Pandya, S. (1991). Exploring the Effects of Consumers' Dissatisfaction Level on Complaint Behaviours. *European Journal of Marketing*, 25, 7-21. <https://doi.org/10.1108/EUM00000000000621>

- Singh, J. and R. E. Wilkes (1996). When Consumers Complain: A Path Analysis of the Key Antecedents of Consumer Complaint Response Estimates. *Journal of the Academy of Marketing Science*, 24 (4), 350-67. <https://doi.org/10.1177/0092070396244006>
- Singh, J., & Howell, T. D. (1985). *Consumer Complaining Behavior: A Review and Prospectus*. In H. K. Hunt, & R. L. Day (Eds.), *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 8, 41-49.
- Singh, J., & Wilkes, R. E. (1996). When Consumers Complain: A Path Analysis of The Key Antecedents of Consumer Complaint Response Estimates. *Journal of the Academy of Marketing science*, 24(4), 350. <https://doi.org/10.1177/0092070396244006>
- Slama, M. and K. Celuch (1994). Assertion and Attention to Social Comparison Information as Influences on Consumer Complaint Intentions. *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behaviour*, 7, 246-51.
- Slegg, J. (2020, January 30). Hashtags Come to Google Search Results. Retrieved January 07, 2021, from <http://searchenginewatch.com/article/2297394/Hashtags-Come-to-Google-Search-Results>
- Smith, A. K., & Bolton, R. N. (1998). An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounters: Paradox or Peril? *Journal of Service Research*, 1(1), 65-81. <https://doi.org/10.1177/109467059800100106>
- Smith, A. K., R. N. Bolton, and J. Wagner (1999). A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research*, 36 (3), 356-72. <https://doi.org/10.1177/002224379903600305>
- Solnick, S. J. and D. Hemenway (1992). Complaints and Disenrollment at A Health Maintenance Organization. *Journal of Consumer Affairs*, 26 (1), 90- 104. <https://doi.org/10.1111/j.1745-6606.1992.tb00017.x>
- Spalding, J. B., Jr. and N. Marcus (1981). *Postal and Telephone Complaint Handling Procedures: A Comparative Study of the U.S. and the U.K.* In *Consumer*

- Satisfaction*. In Dissatisfaction and Complaining Behavior Proceedings, Ralph L. Day and H. Keith Hunt, Eds. Bloomington: Indiana University.
- Starkweather, J. and Moske, A.K. (2011). Multinomial Logistic Regression. URL: http://www.unt.edu/rss/class/Jon/Benchmarks/MLR_JDS_Aug2011.pdf
- Stauss, B., & Seidel, W. (2004). *Complaint Management: The Heart of CRM*. Thomson/South-Western.
- Stephens, N. 2000. *Complaining*. In Teresa Swartz and Dawn Iacobucci (Eds), Handbook of Service Marketing and Management, California: Sage Publications.
- Stephens, N. and K. P. Gwinner (1998). Why Don't Some People Complain? A Cognitive-Emotive Process Model of Consumer Complaint Behavior. *Journal of the Academy of Marketing Science*, 26 (3), 172-89. <https://doi.org/10.1177/0092070398263001>
- Stephens, N., & Gwinner, K. P. (1998). Why Don't Some People Complain? A Cognitive-Emotive Process Model of Consumer Complaint Behavior. *Journal of the Academy of Marketing science*, 26(3), 172-189. <https://doi.org/10.1177/0092070398263001>
- Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social Media Analytics—Challenges in Topic Discovery, Data Collection, And Data Preparation. *International Journal of Information Management*, 39, 156-168. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>
- Strauss, J., & Pesce, A. (1998). Corporate Responses to Consumer E-Mail Complaints: A Pilot Study. *Marketing Management Association 1998 Proceedings*, 46-50.
- Su, L., Wan, L. C., & Wyer Jr, R. S. (2018). The Contrasting Influences of Incidental Anger and Fear on Responses to A Service Failure. *Psychology & Marketing*, 35(9), 666-675. <https://doi.org/10.1002/mar.21114>
- Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting Dark Triad Personality Traits from Twitter Usage and A Linguistic Analysis of Tweets. In *2012 11th International Conference on Machine Learning and Applications 2*, 386-393.

- Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning*. Cambridge: MIT Press.
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using Multivariate Statistics*. Boston, MA: Pearson.
- Talib, R., Hanif, M. K., Ayesha, S., & Fatima, F. (2016). Text Mining: Techniques, Applications and Issues. *International Journal of Advanced Computer Science and Applications*, 7(11), 414-418.
- Tax, S. S., Brown, S. W., & Chandrashekar, M. (1998). Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing. *Journal of Marketing*, 62(2), 60-76. <https://doi.org/10.1177/109467050023006>
- Thevenot, G. (2007). Blogging as A Social Media. *Tourism and Hospitality Research*, 7(3-4), 287–289. <https://doi.org/10.1057/palgrave.thr.6050062>
- Thøgersen, J., Juhl, H. J., & Poulsen, C. S. (2009). Complaining: A Function of Attitude, Personality, And Situation. *Psychology & Marketing*, 26(8), 760-777. <https://doi.org/10.1002/mar.20298>
- Tiedens, L. Z., & Linton, S. (2001). Judgment Under Emotional Certainty and Uncertainty: The Effects of Specific Emotions on Information Processing. *Journal of Personality and Social Psychology*, 81(6), 973. <https://doi.org/10.1037/0022-3514.81.6.973>
- Tracy, P. (2020). Stock Market Index. Retrieved January 07, 2021, from <https://investinganswers.com/dictionary/s/stock-market-index>
- Trainor, K. J. (2012). Relating Social Media Technologies to Performance: A Capabilities-Based Perspective. *Journal of Personal Selling and Sales Management*, 32(3), 317–33. <https://doi.org/10.2753/PSS0885-3134320303>
- Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of Sentiment Reviews Using N-Gram Machine Learning Approach. *Expert Systems with Applications*, 57, 117-126. <https://doi.org/10.1016/j.eswa.2016.03.028>
- Tripp, T. M., & Grégoire, Y. (2011). When Unhappy Customers Strike Back on The Internet. *MIT Sloan Management Review*, 52(3), 37-44.

- Tronvoll, B. (2007a). Complainer Characteristics When Exit is Closed. *International Journal of Service Industry Management*, 18 (1), 25-51. <https://doi.org/10.1108/09564230710732885>
- Tronvoll, B. (2007b). Customer Complaint Behaviour From the Perspective of The Service-Dominant Logic of Marketing. *Managing Service Quality: An International Journal*, 17(6), 601–620. <https://doi.org/10.1108/09604520710834966>
- Tronvoll, B. (2008). *Customer Complaint Behaviour in Service*, Doctoral Dissertation, Karlstads Universitet.
- Tronvoll, B. (2011). Negative Emotions and Their Effect on Customer Complaint Behaviour. *Journal of Service Management*, 22(1), 111–134. <https://doi.org/10.1108/09564231111106947>
- Tronvoll, B. (2012). A Dynamic Model of Customer Complaining Behaviour From the Perspective of Service-Dominant Logic. *European Journal of Marketing*, 46(1/2), 284-305. <https://doi.org/10.1108/03090561211189338>
- Tschohl, J. (2007). Why Great Customer Service Crushes the Opposition. *Customer Management*, 15(4), 18.
- Tse, D.K. and Wilton, P.C. (1988). Models of Consumer Satisfaction Formation: An Extension. *Journal of Marketing Research*, 25(May), 204-12. <https://doi.org/10.1177/002224378802500209>
- Tupes, E. C., & Christal, R. E. (1961). *Recurrent Personality Factors Based on Trait Ratings*. (USAF Tech. Rep. No. 61-97). Lackland Air Force Base, TX: U.S. Air Force.
- Turney, P. D., & Littman, M. L. (2003). Measuring Praise and Criticism: Inference of Semantic Orientation from Association. *ACM Transactions on Information Systems (TOIS)*, 21(4), 315-346. <https://doi.org/10.1145/944012.944013>
- Tuten, T. L., & Solomon, M. R. (2018). *Social Media Marketing*. New Delhi: Sage Texts.
- Van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press.

- Vaughan, P. (2014). 72% of People Who Complain on Twitter Expect a Response Within an Hour. Retrieved January 07, 2021, from <https://blog.hubspot.com/marketing/twitter-response-time-data>
- Vázquez, S., Muñoz-García, Ó., Campanella, I., Poch, M., Fisas, B., Bel, N., & Andreu, G. (2014). A Classification of User-Generated Content into Consumer Decision Journey Stages. *Neural Networks*, 58, 68-81. <https://doi.org/10.1016/j.neunet.2014.05.026>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward A Unified View. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, 36(1), 157-178. <https://doi.org/10.2307/41410412>
- Verhagen, T., Nauta, A., & Feldberg, F. (2013). Negative Online Word-Of-Mouth: Behavioral Indicator or Emotional Release? *Computers in Human Behavior*, 29(4), 1430-1440. <https://doi.org/10.1016/j.chb.2013.01.043>
- Vermeulen, I. E., & Seegers, D. (2009). Tried and Tested: The Impact of Online Hotel Reviews on Consumer Consideration. *Tourism Management*, 30(1), 123-127. <https://doi.org/10.1016/j.tourman.2008.04.008>
- Viens, A. (2019). This graph tells us who's using social media the most. Retrieved January 07, 2021, from <https://www.weforum.org/agenda/2019/10/social-media-use-by-generation/>
- Volkov, M., D. Harker, and M. Harker (2002). Opinions About Advertising in Australia: A Study of Complainants. *Journal of Marketing Communications*, 8 (4), 229-42. <https://doi.org/10.1080/13527260210142329>
- Walker, L., Baines, P. R., Dimitriu, R., & Macdonald, E. K. (2017). Antecedents of Retweeting in A (Political) Marketing Context. *Psychology & Marketing*, 34(3), 275-293. <https://doi.org/10.1002/mar.20988>

- Wang, X., Baesens, B., & Zhu, Z. (2019). On the Optimal Marketing Aggressiveness Level of C2C Sellers in Social Media: Evidence from China. *Omega*, 85, 83-93. <https://doi.org/10.1016/j.omega.2018.05.014>
- Wang, Y. (2005). A Multinomial Logistic Regression Modeling Approach for Anomaly Intrusion Detection. *Computers & Security*, 24(8), 662-674. <https://doi.org/10.1016/j.cose.2005.05.003>
- Ward, J. C., & Ostrom, A. L. (2006). Complaining to The Masses: The Role of Protest Framing in Customer-Created Complaint Web Sites. *Journal of Consumer Research*, 33(2), 220-230. <https://doi.org/10.1086/506303>
- Warland, R. H., R. O. Herrmann, and J. Willits (1975). Dissatisfied Consumers: Who Gets Upset and Who Takes Action. *Journal of Consumer Affairs*, 9 (Winter), 148-63. <https://doi.org/10.1111/j.1745-6606.1975.tb00559.x>
- Watson, D., & Clark, L. A. (1992). On Traits and Temperament: General and Specific Factors of Emotional Experience and Their Relation to the Five-Factor Model. *Journal of Personality*, 60(2), 441-476. <https://doi.org/10.1111/j.1467-6494.1992.tb00980.x>
- Watts, S. A., & Zhang, W. (2008). Capitalizing on Content: Information Adoption in Two Online Communities. *Journal of The Association for Information Systems*, 9(2), 3. <https://doi.org/10.17705/1jais.00149>
- Webster, C. and D. S. Sundaram (1998). Service Consumption Criticality in Failure Recovery. *Journal of Business Research*, 41 (2), 153-59. [https://doi.org/10.1016/S0148-2963\(97\)00004-0](https://doi.org/10.1016/S0148-2963(97)00004-0)
- Weiss, S. M., Indurkha, N., Zhang, T., & Damerau, F. (2010). *Text Mining: Predictive Methods for Analyzing Unstructured Information*. New York: Springer.
- Werbos, P.J. (1974) *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Doctoral Dissertation, Harvard University, Cambridge.
- Westbrook, R. A. (1987). Product/Consumption-Based Affective Responses and Postpurchase Processes. *Journal of Marketing Research*, 24(3), 258-270. <https://doi.org/10.1177/002224378702400302>

- Westbrook, R. A., & Newman, J. W. (1978). An Analysis of Shopper Dissatisfaction for Major Household Appliances. *Journal of Marketing Research*, 15(3), 456-466. <https://doi.org/10.1177/002224377801500316>
- Westbrook, R. A., & Oliver, R. L. (1991). The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction. *Journal of Consumer Research*, 18(1), 84-91. <https://doi.org/10.1086/209243>
- Williams, P., & Aaker, J. L. (2002). Can Mixed Emotions Peacefully Coexist? *Journal of Consumer Research*, 28(4), 636-649. <https://doi.org/10.1086/338206>
- Williams, P., & Aaker, J. L. (2002). Can Mixed Emotions Peacefully Coexist? *Journal of Consumer Research*, 28(4), 636-649. <https://doi.org/10.1086/338206>
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2015). *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco, CA: Morgan Kaufmann.
- Witten, I.H. and Frank, E. (2000). *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. San Francisco, CA: Morgan Kaufmann.
- Wolrich, C. (2013). Top Corporate Hate Web Sites. Retrieved January 07, 2021, from https://www.forbes.com/2005/03/07/cx_cw_0308hate.html#2b79f6194ed5
- Wong, T. C., Chan, H. K., & Lacka, E. (2017). An ANN-Based Approach of Interpreting User-Generated Comments from Social Media. *Applied Soft Computing*, 52, 1169-1180. <https://doi.org/10.1016/j.asoc.2016.09.011>
- Woodcock, N., Broomfield, N., Downer, G., & Starkey, M. (2011). The Evolving Data Architecture of Social Customer Relationship Management. *Journal of Direct, Data and Digital Marketing Practice*, 12(3), 249-266. <https://doi.org/10.1057/dddmp.2010.45>
- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A Comparative Analysis of Major Online Review Platforms: Implications for Social Media Analytics in Hospitality and Tourism. *Tourism Management*, 58, 51-65. <https://doi.org/10.1016/j.tourman.2016.10.001>

- Yao X., Liu Y. (2014). *Machine Learning*. In: Burke E., Kendall G. (Eds) Search Methodologies. Boston, MA: Springer.
- Ye, J., Janardan, R., Park, C. H., & Park, H. (2004). An Optimization Criterion for Generalized Discriminant Analysis on Undersampled Problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8), 982-994. <https://doi.org/10.1109/TPAMI.2004.37>
- Ye, J., Xiong, T., Li, Q., Janardan, R., Bi, J., Cherkassky, V., & Kambhamettu, C. (2006). Efficient Model Selection for Regularized Linear Discriminant Analysis. *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*, 532-539. <https://doi.org/10.1145/1183614.1183691>
- Youngs, G. (2009, March). Blogging and Globalization: The Blurring of The Public/Private Spheres. *Aslib Proceedings*, 61(2), 127-138. <https://doi.org/10.1108/00012530910946884>
- Zarrella, D. (2009). *The Social Media Marketing Book*. O'Reilly Media, Inc.
- Zeithaml, V. A., L. L. Berry, and A. Parasuraman (1988). Communication and Control Processes in the Delivery of Service Quality. *Journal of Marketing*, 52 (2), 35-48. <https://doi.org/10.1177/002224298805200203>
- Zerfass, A., & Schramm, D. M. (2014). Social Media Newsrooms in Public Relations: A Conceptual Framework and Corporate Practices in Three Countries. *Public Relations Review*, 40(1), 79-91. <https://doi.org/10.1016/j.pubrev.2013.12.003>
- Zhang, Y. (2010). *New Advances in Machine Learning*. InTech.
- Zhang, L. and Peng, T.Q. (2015), Breadth, Depth, And Speed: Diffusion of Advertising Messages on Microblogging Sites. *Internet Research*, 25(3), 453-470
- Zhao, Y., & Karypis, G. (2002). Evaluation of Hierarchical Clustering Algorithms for Document Datasets. *Proceedings of The Eleventh International Conference on Information and Knowledge Management*, 515-524. <https://doi.org/10.1145/584792.584877>

Zhu, X. (2011). *Semi-Supervised Learning*. Technical Report Computer Sciences 1530, University of Wisconsin – Madison.

Zhu, X., & Goldberg, A. B. (2009). Introduction to Semi-Supervised Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 3(1), 1-130.

Zocca, V., Spacagna, G., Slater, D. ve Roelants, P. (2017). *Python Deep Learning*. Birmingham: Packt Publishing.

