FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
T Ü R K İ Y E

DEEP LEARNING-BASED SENTIMENT ANALYSIS FOR CLOUD PROVIDER SELECTION

Muhammad Raheel RAZA

Master's Thesis

DEPARTMENT OF SOFTWARE ENGINEERING

JANUARY 2022

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Software Engineering

Master's Thesis

DEEP LEARNING-BASED SENTIMENT ANALYSIS FOR CLOUD PROVIDER SELECTION

Author
Muhammad Raheel RAZA

Supervisor
Prof. Dr. Erkan TANYILDIZI

JANUARY 2022
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Software Engineering

Master's Thesis

Title:	Deep Learning-based Sentiment Analysis for Cloud Provider Selection
Author:	Muhammad Raheel RAZA
Submission Date:	02 December 2021
Defense Date:	27 January 2022

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and was unanimously approved after the defense exam made open to the academic audience.

Supervisor:	Prof. Dr. Erkan TANYILDIZI Fırat University, Faculty of Technology	[redacted]	Approved
Chair:	Prof. Dr. Fatih TALU İnönü University, Faculty of Engineering	[redacted]	Approved
Member:	Doç. Dr. Fatih ÖZKAYNAK Fırat University, Faculty of Technology	[redacted]	Approved

This thesis was approved by the Administrative Board of the Graduate School on

.... / .... / 20 ....

Signature

Prof. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesis titled "Deep Learning-based Sentiment Analysis for Cloud Provider Selection" in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Firat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

27 January 2022

Muhammad Raheel RAZA

PREFACE

Cloud Computing paradigm can maintain a trustful environment between cloud service providers and consumers using different strategies. One of the methods includes the appropriate selection of CSP based on service consumers' past experiences. The thesis entitled "Deep Learning-based Sentiment Analysis for Cloud Provider Selection" is performed by performing Aspect-Based Sentiment Analysis of extracted cloud consumer reviews using Deep Learning approaches of RNN, LSTM and GRU. The study, then, uses an Analytic Hierarchy Process model for decision making of suitable CSPs concerning the aspects considered for evaluation.

Furthermore, I would like to thank my supervisor, Prof. Dr Erkan Tanyıldızı, under the supervision of whom I was able to perform my thesis and get a lot of valuable knowledge and information about the domain field. In addition, I want to express my deep gratitude to Dr Walayat Hussain, who supported me with his extensive help during all phases of my thesis. The completion of my thesis would not have been possible without his nurturing and guidance. I would also like to thank my family, who advised and helped me study in this lovely country, Turkey.

Muhammad Raheel RAZA
ELAZIG, 2022

TABLE OF CONTENTS

Page

ABSTRACT

Deep Learning-based Sentiment Analysis for Cloud Provider Selection

Muhammad Raheel RAZA

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Software Engineering

January 2022, Page: xi + 53

The selection of a viable Cloud Service Provider (CSP) has always been a crucial task for a Cloud Service Consumer (CSC) to avail of their offered services. This selection would enable a service consumer to maintain a trustful relationship with a provider. For that purpose, consumer reviews posted on internet websites and other social media platforms need to be carefully evaluated for a proper CSP selection. Sentiment Analysis, also termed Opinion Mining, is the computational treatment of text's views, experiences, sentiments, and subjectivity. Aspect-Based Sentiment Analysis (ABSA) extracts informative aspects within the text and uses them to classify the sentiment of reviews. Nowadays, different lexicon-based, supervised learning, and un-supervised learning techniques are used for sentiment classification tasks.

Deep Learning is an AI technique used for language processing, text analysis, pattern recognition, sequence predicition tasks, etc. Its types, such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), use different strategies to carry out processings. The dissertation performs Aspect-Based Sentiment Analysis of cloud consumer reviews using Deep Learning approaches of RNN, LSTM and GRU. The cloud reviews are extracted using Harvesting-as-a-Service (HaaS) framework. Analytic Hierarchy Process (AHP) model is used to decide the importance and priorities of aspects for Cloud Service Consumers (CSCs). The evaluation would assist cloud service consumers (CSCs) choose the best CSP ideal for their requirements.

Keywords: Aspect-Based Sentiment analysis, Deep Learning, RNN, LSTM, GRU, Analytic Hierarchy Process (AHP).

ÖZET

Bulut Sağlayıcı Seçimi için Derin Öğrenmeye Dayalı Duyarlılık Analizi

Muhammad Raheel RAZA

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Yazılım Mühendisliği Anabilim Dalı

Ocak 2022, Sayfa: xi + 53

Geçerli bir Bulut Hizmeti Sağlayıcısının (CSP) seçimi, bir Bulut Hizmeti Tüketicisinin (CSC) sunulan hizmetlerden yararlanabilmesi için her zaman çok önemli bir görev olmuştur. Bu seçim, bir hizmet tüketicisinin bir sağlayıcı ile güvenilir bir ilişki sürdürmesini sağlayacaktır. Bu amaçla, uygun bir CSP seçimi için internet sitelerinde ve diğer sosyal medya platformlarında yayınlanan tüketici incelemelerinin dikkatlice değerlendirilmesi gerekir. Fikir Madenciliği olarak da adlandırılan Duygu Analizi, metnin görüşlerinin, deneyimlerinin, duygularının ve öznelliğinin sayısal olarak ele alınmasıdır. Boyut Tabanlı Duygu Analizi, metin içindeki bilgilendirici yönleri çıkarır ve bunları incelemelerin duyarlılığını sınıflandırmak için kullanır. Günümüzde duygu sınıflandırma görevleri için sözlük tabanlı, denetimli öğrenme ve denetimsiz öğrenme teknikleri gibi farklı yöntemler kullanılmaktadır.

Derin Öğrenme, dil işleme, metin analizi, örüntü tanıma ve dizi tahmini görevleri vb. için kullanılan bir yapay zeka tekniğidir. Tekrarlayan Sinir Ağı (RNN), Uzun Kısa Süreli Bellek (LSTM) ve Geçitli Tekrarlayan Birim (GRU) gibi türleri, işlemleri gerçekleştirmek için farklı stratejiler kullanır. Tez RNN, LSTM ve GRU'nun Derin Öğrenme yaklaşımlarını kullanarak bulut tüketici incelemelerinin Boyut Tabanlı Duygu Analizini gerçekleştirir. Bulut incelemeleri, Hizmet Olarak Hasat (HaaS) çerçevesi kullanılarak çıkarılır. Analitik Hiyerarşi Süreci (AHP) modeli, Bulut Hizmeti Tüketicileri (CSC'ler) için yönlerin önemine ve önceliklerine karar vermek için kullanılır. Değerlendirme, bulut hizmeti tüketicilerinin (CSC'ler) gereksinimlerine uygun en iyi CSP'yi seçmelerine yardımcı olacaktır.

Anahtar Kelimeler: Boyut Tabanlı Duygu Analizi, Derin Öğrenimi, RNN, LSTM, GRU, Analitik Hiyerarşi Süreci (AHP).

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ABSA	: Aspect-Based Sentiment Analysis
BDQM	: Big Data Quality Metrics
CNN	: Conventional Neural Network
CSC	: Cloud Service Consumer
CSP	: Cloud Service Provider
DL	: Deep Learning
GRU	: Gated Recurrent Unit
IaaS	: Infrastructure as a Service
i.e.	: That is
LSTM	: Long short term memory
ML	: Machine Learning
NLP	: Natural Language Processing
OM	: Opinion Mining
PaaS	: Platform as a Service
RNN	: Recurrent Neural Network
SA	: Sentiment Analysis
SaaS	: Software as a Service
SQ	: Service Quality
w.r.t	: With respect to

1. INTRODUCTION

The Internet and Rapid technological advancements have ushered in the E-commerce age. E-commerce refers to the online sale and purchase of products and services. Nowadays, every client wants to express their thoughts and views regarding the items or services they've purchased on the internet. These points of view and opinions aid in the decision-making process and significantly impact the company model. The reviews of an item will also help establish client trust, which will aid in business expansion. As a result, consumer feedback must be evaluated and analyzed. The majority of reviews are written in text format. On the internet, each product has received several favourable and inadequate evaluations. It is, however, difficult to read each review in its entirety. Recent studies have found that a graphic depiction is more effective and easier to learn and understand than a written one. As a result, it becomes necessary to translate written assessments into a visual format, increasing the trustworthiness of decision-making processes.

1.1. Related Works

1.1.1. Sentiment Analysis using Deep Learning Methods

SentiVec [1], a kernel optimization function system, is proposed to have 2 phases. The 1st phase uses a supervised learning algorithm that performs sentiment word embedding to vectors. These vectors are fed into the 2nd phase. Two unsupervised learning models, i.e. object-word-to-surrounding-words reward model (O2SR) and the context-to-object-word reward model (C2OR), identify sentiment similarities and learn sentiments vectors for both sentiment and non-sentiment words. The approach performs NLP evaluation methods like word analogy, word similarity and sentiment analysis and updates all the semantic information of text words within a corpus. Paper [2] uses Conventional Machine Learning in combination with the Convolutional Neural Network (CNN) as a deep learning model for sentiment analysis of a digital text. The technique aimed to extract the emotional state encapsulated within the digital textual data. Document Term Matrix and Support Vector Machines (DTM-SVM) approaches compute the statistical evaluation of data. CNN performs the Doc2Vec, an extension of word2vec, converting documents into vectors.

The thesis [3] performs categorization of Turkish news text using Machine Learning and Deep Learning methods. The purpose is to identify the type of news as per the content criteria. Document form data is classified using Support Vector Machine (SVM) while label form data is fed to Recurrent Neural Network (RNN). Classifications prove that RNN has a higher accuracy rate than that of SVM. However, the speed ratio of SVM was faster than RNN. Deep learning was used for sentiment analysis of Big data of a financial social platform in the paper [4]. StockTwits, an online social platform where investors and traders from all over the world share their investing ideas, experiences and views about stock market fluctuations, was evaluated using data mining methods, logistic regression and deep learning approaches. Deep learning methods achieved a higher accuracy rate. Out of which, Convolutional Neural Network (CNN) outperforms than the rest extracting accurate sentiments from author's words regarding stocks.

Authors [5] have created a German SentiWordNet lexicon containing German words translated from English lexicons. The derived lexicon and two other lexicons are fed to a context-based sentiment analysis system using Recurrent Neural Network (RNN) that learns contextual sentiment weights changing the lexicon polarity of words w.r.t context. Evaluations show that better performance is achieved using deep learning approaches. Using deep learning-based sentiment analysis, the paper [6] performs cross-cultural polarity and emotion detection regarding COVID-19 pandemic tweets. The process first identifies the polarity of tweets as positive or negative, and then emotion analysis is performed using Deep Long Short-Term Memory (LSTM) models. The study's purpose was to analyse citizens' reactions from neighbouring countries, i.e. Denmark and Sweden.

Deep Learning for Affective Computing [7] finds Recurrent Neural Network as the best method to recognize text-based emotions of individuals in affecting computing. Sentiment analysis in the form of emotion recognition helps better decision support. For performance improvements of the deep neural network, bidirectional LSTM processing and dropout regularization and weight-loss functions are performed. Transfer learning, i.e. senti2affect approach, transfers knowledge from source to a different but related problem and overcomes overfitting problem in DNN. ConvLSTMConv network [8] is a proposed deep learning approach for sentiment analysis of social media reviews. The system integrates Convolutional Neural Network (CNN) and LSTM (Long Short-Term Memory) on Google cloud for storage purposes and Google Colaboratory for computing. The deep learning algorithm is implemented with word embedding technique, features are learned and extracted through CNN layer, the extracted features are then sent to a bidirectional LSTM layer for long-term feature dependencies. Another CNN layer is used to again learn the extracted features for better learning and machine understanding performance. The approach performs a binary classification, i.e. positive or negative sentiment categorization.

Paper [9] addresses the problem of sentence-level sentiment classification on IMDB dataset and the Restaurant reviews dataset. Two deep learning models are used, i.e. Deep feedforward Neural Network model with global average pooling and the LSTM model with dense layers. Deep Feedforward NN model achieved 88% accuracy rate while 78% for LSTM model. The authors have presented a novel neural network model, i.e. NgramCNN [10]. The new architectural model performs sentiment analysis of long text documents with diverse contents. It follows a

Convolutional Neural Network representing text features out of the pre-trained n-grams, convolutional and max-pooling layers for feature extraction and a single layer simple classifier to classify each document per their polarity. The model is experimented on three different review datasets to perform better and have high prediction and accuracy rates than CNN and Deep RNN approaches.

The thesis [11] works to mitigate the effects of Cyberbullying thus, highlighting the importance of sentiment analysis of abusive comments on social media. The work uses graph structures and deep learning approaches for this purpose. The text tweets are represented with sentence embeddings generated using two models, namely DistilBERT and ELECTRA whereas node embeddings are formed with the help of NBNE and Graph-Wave approach. A 3-headed CNN model, i.e. Hydra model, is used to do sentiment classification of the text. ConvLSTM model [12] is proposed combining CNN and LSTM models that perform classification of short texts. The model is supposed to overcome the defects faced in traditional deep learning approaches. As traditional approaches used CNN which needed more layers for long term dependency capturing, thus in this model, LSTM layer is used as a single pooling layer to CNN that helps to capture long-term dependencies in short texts and to minimize the loss of detailed information. The model proves to perform efficiently with minimum number of parameters.

Paper [13] highlights the importance of text quality from sentiment analysis of online reviews and how it affects classification performance. Two basic textual quality features that affect classification performance i.e. word count and review readability, are considered. CNN and two RNN models, i.e. Simple Recurrent Network (SRN) and LSTM are used for sentiment analysis. According to the study performed by a text evaluator, small reviews, i.e. less word count with higher readability, achieve higher accuracy rates. Thus, the high-quality text is helpful to providers and consumers for business purposes and the decision-making process. Binary and multi-class sentiment classification of reviews using deep learning [14] is performed on movie reviews 45dataset. Reviews are transformed into vector representations using word2vec and pre-trained Global vector (GloVe) embedding. Based on the content, the star ratings of reviews are also predicted. Performances of RNN-LSTM, RNN-Bidirectional LSTM and CNN, are evaluated and compared. Word embedding, feature extraction and class nature have an influence on classification outcomes. Results of evaluation metrics show that RNN-Bidirectional LSTM achieves a higher accuracy rate than the others.

Multimodal Sentiment Analysis (MSA) [15] is performed using deep learning approaches. The study involves polarity classification not only of the textual data but also of visual and audio data. Machine Learning approaches such as SVM, Naïve Bayes, Decision Tree and Random Forest etc., have achieved less accuracy than the Deep Learning models for both one-modality and multi-modality data to analyze the polarity. Author [16] has used an Attention model to carry out sentiment prediction of news text. News articles have an enormous amount of mixed and biased information, and to extract the real ones, its sentiment analysis is quite important. Glove embedding technique performs word2vector representation task. A sentiment analyzer performs Sentence-level and document-level sentiment prediction into positive, negative and neutral categories. The proposed sentiment analyzer can further be used for detecting COVID-19 people's sentiments as described.

Aspect-based Financial Sentiment Analysis [17] uses financial headlines and microblogs having financial tweets and performs aspect-based sentiment analysis using Deep learning. A sentence is divided into snippet and target. Aspects in the snippet is searched out, and the sentiment score of the sentence is calculated. RNN model with Bi-LSTM extracts aspects, i.e. features from the text, and a multi-channel CNN model for its sentiment analysis. The framework also calculates the sentiment intensity score of a sentence. Paper [18] conducted a Deep learning-based sentiment classification on Hindi movie reviews from online websites and newspapers. CNN model performs sentence-level analysis into sentiments as positive, negative or neutral.

In Hybrid CNN-LSTM model [19], CNN model extracts high-level features with its convolutional layers and also requires more layers to capture long-term dependencies. Thus, the LSTM model captures long-term dependencies between word sequences and learns sentiment polarity from texts. The proposed model captures both models' features and has achieved a 91% accuracy rate than machine learning and deep learning approaches. The authors [20] conducted a Financial sentiment analysis of social comments to predict the direction of stocks in the Turkish stock market using CNN, RNN, LSTM and a new generation word embedding model, i.e. Bidirectional Encoder Representations from Transformers (BERT). According to the authors, the combination of deep learning models and a new NLP model has improved the accuracy performance, helping investors predict stock directions correctly and guide their investments to gain maximum profits.

Paper [21] describes image sentiment analysis performed by the authors to identify the sentiments of unlabeled images. DNN, CNN, Region-based CNN (R-CNN), and Fast R-CNN techniques are utilized to achieve the purpose. Similarly, the paper [22] presents Residual Attention-based Deep learning Network (RA-DLNet), a deep learning architecture for sentiment analysis of visual data. The model uses CNN-based architectures to learn all the hierarchies of image features. Sentiment classification of texts by Parts of Speech (POS) tags and joint sentiment topic measures [23] is conducted using SVM and ANN models. POS tags and joint sentiment topic measures are extracted as unigrams, bigrams and bi-tagged from the text as features. Maximum relevancy and minimum redundancy mutual information of the features are the key points to feature selection for sentiment analysis. ANN is applied for large datasets, while

SVM for the non-linear ones are thus applied for balanced and unbalanced data. The method improves the accuracy performance of the analysis.

BowTie [24] is a novel deep learning feedforward neural network method for sentiment analysis. The model aims to produce sentiment classification accurately, maintaining low losses. The network is capable of being transferred to other datasets. The method is used with 2-word encoding types and ensures low loss estimates for trained data to enable accurate transfer to other datasets. The paper [25] has performed sentiment analysis of Amazon reviews and ratings using deep learning methods. Using paragraph vector, the reviews are transformed into vectors and is fed to train the RNN model. The study solves the problem between reviews and mismatched ratings by creating a web service application. The application predicts the rating score of the review and compares it with the submitted rating score. If different, it informs the user via feedback.

Deep learning is used to conduct sentiment analysis of COVID-19 Twitter tweets [26]. By dividing the tweets into 2-time span categories, studies showed that while most internet users share positive or neutral tweets, some netizens use to re-tweet negative tweets that misguided the people and spread negativity. Deep learning and a Fuzzy rule-based technique performed the sentiment classification into positive, negative and neutral categories to accurately identify sentiments from tweets. Authors [27] discussed a Convolutional Fuzzy Sentiment Classifier (CFSC) model that combines a Deep Convolutional neural network and Fuzzy logic to predict sentiments in text sequences and videos. The study performs fuzzy commonsense reasoning for multiple modality sentiment analysis. Aspect-based sentiment analysis (ABSA) of target-dependent Twitter tweets [28] on product reviews performs aspect, i.e. feature extraction and then applies sentiment classification on aspects in the survey paper. Several deep learning approaches and their variants, including CNN, LSTM, RNN, GRU models, are compared and analyzed based on their performance and results.

Paper [29] discussed Arabic Subjective Sentiment Analysis (ASSA) with the help of deep learning methods applied on the social media domain. The study contains a review of all deep learning approaches used for sentiment analysis of ASSA. Arabic NLP problems are also being addressed. Research shows that CNN and RNN models are widely used for ASSA tasks. A Smart COVID-19 Chatbot is designed [30] that interacts with different users to make aware them of precautions and get their opinions and feelings about pre and post-quarantine conditions. The data is sentimentally analyzed using deep learning and NLP techniques to identify crucial information. Bengali Restaurant reviews are analyzed using SentiLSTM technique [31], a deep learning approach, i.e. BiLSTM performs sentiment analysis of 9000 corpora of customer opinions regarding the restaurant. Performances of other machine learning approaches are compared, out of which BiLSTM performs the best.

1.1.2. RNN, LSTM & GRU-based Sentiment Analysis

The novel Attention-based LSTM is used to perform Aspect-based Sentiment Analysis [32]. Because content polarity alone is insufficient to identify text sentiment, aspect polarity is assessed with content polarity. Using RNN-based Long Short Term Memory, an attention mechanism is employed to highlight the sentence segment determined by a specific aspect (LSTM). Various LSTM models were utilized, with the AE-LSTM and ATAE-LSTM outperforming the others. In the thesis [33], The sentiment categorization of the Turkish domain is done using Recurrent Neural Networks (RNN). Two sets of Turkish movie reviews datasets were categorised using various LSTM and Bidirectional LSTM type models. The best approach was LSTM with Word2Vec, which had the greatest F1-score of 95.1 per cent.

Sentence Representation- LSTM (SR-LSTM) and Sorted Sentence Representation- LSTM (SSR-LSTM) models [34] were developed for purpose of sentiment classification at the document level on three datasets. SR-LSTM employs two hidden layers, one for the generation of sentence vectors and another for the generation of document representations utilizing sentence vector inputs. By screening phrases with little sentiment polarity, SSR-LSTM enhances SR-LSTM. For the objective of attaining high accuracy, binary sentiment classification is conducted using the RNN LSTM model [35]. The thesis sorted tweet phrases into favourable and unfavourable categories using two Twitter datasets. Despite the time and other constraints, an accuracy rate of 75-80% was attained, higher than that of several neural network models.

The notion of undertaking sentiment classification using hybrid machine learning and deep learning techniques such as the Nave Bayes classifier and Neural Network classifier was described in the paper [36]. Both classifiers' strengths are utilized to overcome their flaws and increase performance and accuracy. On text reviews from TripAdvisor, Amazon, and IMDB movie reviews, the authors [37] used LSTM to classify emotion. Tokenization and embedding steps were undertaken before applying LSTM. The entire procedure was run through ten epochs, with computed accuracies at each stratum and epoch. It was possible to reach a net accuracy rate of 85 per cent.

Deep learning approaches are used to classify the sentiment of business and monetary news [38]. NLP uses GloVe to embed the news, determine polarity, and retrieve emotional reactions, then put into the RNN LSTM model to explore and anticipate potential stock values. The datasets TRNA and TRHA are utilized for this. For emotion categorization, the Bidirectional Encoder Representation from Transformers (BERT) model [39] is suggested. BERT is a deep-based language model that represents the context embedding characteristic with a larger text corpus and expanded features. Pre-training and fine-tuning tasks using the BERT model increase categorization accuracy.

The article [40] has used four alternative methodologies, including lexicon-based SentiWordNet, machine learning logistic regression-based, deep learning LSTM-based, and advanced deep learning BERT-based methods, to undertake a comparative analysis and sentiment categorization of IMDB reviews. The taxonomy and relevance of sentiment categorization were defined in a survey on Sentiment Analysis using Deep Learning Architectures [41]. It analyzed several datasets and emphasized various deep learning algorithms suited for sentiment analysis, such as CNN, RNN, LSTM, GRU, etc. As per the authors, deep learning techniques are the perfect options for sentiment classification challenges in the algorithm design and construction stages.

1.1.3. Sentiment Analysis of Cloud Reviews

Paper [42] performs sentiment analysis of customers' tweets on cloud service providers Twitter pages, i.e. Microsoft Azure and Amazon Web Services (AWS). The study aims to identify customers' opinions and reviews about cloud products of service providers. Naïve Bayes classification is used for sentiment analysis identifying two lexicons for polarity identification and emotion classification. Results show that AWS had 50% positive tweets and 45% negative tweets, while polarity classification for Microsoft Azure was 65% positive and 25% negative. However, the 'joy' category was better for Microsoft Azure while AWS fell in the 'sadness' category for emotion classification.

Authors [43] have developed a cloud-based tool for opinion mining and sentiment analysis of travellers' comments on TripAdvisor.com about visiting places and restaurants booking etc. The tool gathers data from social websites, analyses people and their comments, detects the polarity of comments, and allows all the operations like data collection, pre-processing, and visualizations to be carried out in the same application.

Cloud Provider Selection Framework (CPSEL) [44] is presented that ranks all the available service providers w.r.t their capabilities and helps select an optimal cloud service provider to fulfil the user requirements. The framework identifies specific indicators as services offered by cloud providers and weights them as per the organization requirements in the criteria set. Providers are then ranked using the Fuzzy TOPSIS method based on the indicators and user feedbacks that increases the trust between cloud providers and users. Finally, a cloud service provider matching user requirements is selected.

The paper [45] discusses the importance of sentiment analysis in human-agent interactions. Chatbots as agents need to accurately identify the speaker's emotions, cope with their emotional status, and give it a vocal expression. The study applies three approaches, i.e. detects human-agent sentiments from cloud-based sentiment analysis services, i.e. Google, Amazon and

Microsoft. It uses the same sentiments and applies machine learning-based approaches for document-level sentiment analysis. Finally, it compares both results with the human ratings.

Tweet Sentiment Analysis (TSA) [46] aims to understand the users' feedback and general opinions regarding cloud computing services and their attitudes towards existing cloud service providers, i.e. Microsoft Azure, Google Cloud Platform and Amazon Web Services (AWS). The study uses Twitter as a source of data tweets applying Bing, AFFIN and NRC lexicons to perform sentiment analysis and emotion classification. The overall polarity classification of all three service providers is 68.5% positive and 3.5% negative.

Deep learning-based Sentiment Analysis of data in Edge Computing [47] emphasizes sentiment classification of client-side data on mobile devices itself. Mobile applications, on behalf of edge computing devices, are supposed to interact with the cloud, take data from Netflix or Amazon etc. and perform sentiment analysis using Deep CNN on devices independent of the server-side. An Android app is designed to classify IMDB and Rotten Tomatoes movie review training and testing datasets without needing a cloud.

1.1.4. SLA Formation in Cloud Computing

A comparative analysis of different time-series prediction approaches using machine learning methods is performed in the paper [48]. These approaches aim for violation management by forming viable Service Level Agreements (SLAs) in the cloud. Cloud service providers mitigate the risks of service violations and violation penalties by adopting the appropriate methods and predicting future Quality of Service (QoS) parameters according to the conditions. Results revealed the level of accuracy achieved by different time-series machine learning prediction techniques. Authors [49] have evaluated some QoS prediction methods for resource management and risk mitigation to SLA violations in SMEs. All the approaches, including neural network techniques, stochastic methods and time-series methods, are performed over different time intervals and predict the resource utilization occurring in the post-interaction time phase when SLA is established and executed. Among all the methods, ARIMA method achieved the highest prediction accuracy results.

The centralized Cloud Service Repository (CCSR) framework [50] assists cloud service consumers to choose appropriate cloud services and efficiently discovers suitable services for users. Within the framework, the Harvesting-as-a-Service (HaaS) module helps to harvest web data without programming and arranges it in different formats. At the same time, the service repository module functions to locate different cloud services, serving as a source of available cloud services for service consumers to avail. The framework also provides a real-time cloud reviews dataset on cloud services. The paper [51] highlights one of key elements to identify SLA violation from a service consumer perspective. A cloud service consumer's previous resource utilization profile can predict the chances of service violations in the cloud. It determines the future possibilities for similar conditions, thus prompting service providers to take serious actions. Moreover, the profile history of the neighbouring consumers can also serve the purpose.

Authors [52] have emphasized the significance of trust maintenance in cloud computing. Trust proves to be an important parameter to viable SLA formation and a reliable service provider-consumer relationship. One of the methods to maintain trust is continuous SLA monitoring. Various SLA monitoring approaches are described, including the Self-manageable case-based reasoning approach, the SLA-Based Trust Model approach, the Broker-based approach and the Reputation-Based and workflow Composition Approach. A comparative analysis is also performed, determining the characteristics of approaches. Risk Management Framework for SLA violation (RMF-SLA) [53] conducts risk management by helping service providers to mitigate risks. A Fuzzy Inference System (FIS) takes various factors such as consumer reliability, provider's attitude towards risk, predicted consumer behaviour etc., into consideration and makes predictions. The results are pretty beneficial for SME providers for an optimal risk management paradigm.

Paper [54] emphasizes forming a trusted relationship between a service provider and consumer by recognizing various crucial factors. Establish a centralized QoS repository and a suitable assessment framework that deals with resource management and decides the optimal quantity of resources needed by consumers. QoS filtering is an essential step for decision-making to distinguish between real and biased QoS information during SLA formation. The paper performs a critical analysis of the literature to extract the QoS parameters. The profile-based SLA Violation Prediction model [55] is designed from the provider's perspective. Before completing the SLA, the approach starts monitoring it in the pre-interaction time phase. It smartly anticipates the consumer's anticipated resource use based on prior transaction history and calculates the number of necessary resources based on their reliability. The framework assists service providers in determining whether or not to create SLAs, maintaining profitability, and preventing service breaches in the post-interaction time phase.

A Decision Support Model for Cloud Marginal Resource Allocation [56] supports service providers to handle marginal resource management decisions. It manages cloud SLAs before implementation for cloud providers by considering all circumstances, such as whether the client is new or a regular one and the nature of marginal resources demanded by it. Decisions for marginal resource allocation are made using a User-based Collaborative filtering model, top K-NN algorithm and Fuzzy Logic Systems. In the paper [57], there are two time phases in the SLA Management framework; pre-interaction time phase and post-interaction time phase. In situations when a predicted QoS parameter is differentiated from an agreed one, the framework assists a service provider in forming an opinion on a consumer request, offering the appropriate quantity of resources to the consumer, predicting QoS parameters, evaluating run-time QoS metrics, and taking measures to reduce uncertainty.

An Automatic Derivation of Cloud Marketplace and Cloud Intelligence (ADCM&CI) [58] serves cloud consumers in finding efficient and appropriate cloud services. The method uses HaaS to extract real-time cloud data. Sentiment analysis extracts the polarity of reviews within the dataset and identifies the required services of consumers. Classification metrics are calculated using different machine learning algorithms. Optimized Personalized Viable SLA (OPV-SLA) framework [59] is proposed that aids an SME service provider form a feasible SLA. Before the SLA is developed and implemented, a framework begins to manage SLA violations. It also aids a service provider in making the best service formation decision and allocating the right number of marginal resources. Provider-Based Optimized Personalized Viable SLA (OPV-SLA) Framework [60] works from the provider's perspective by offering a unique optimum SLA creation architecture that allows the provider to assess a consumer's durability when getting to the SLA. A consumer's profile is an important factor in assessing their reliability. Therefore, a service provider monitors service consumers’ behaviour in the post-interaction time phase and avoids service penalties through this model. Fuzzy re-SchdNeg Decision Model [61] provides a solution for SaaS providers to use rescheduling strategies to distribute possible violated jobs among existing resources. When the system determines that the current services are inadequate to address the possible violated position, a renegotiation session is initiated between provider and consumer, creating new SLA with additional conditions. Table 1.1 shows a comparative analysis of literature.

1.2. Purpose of Thesis

The thesis aims to perform sentiment analysis using current deep learning approaches that can be used for cloud provider selection purposes. Cloud provider selection is one of the crucial steps for ensuring a successful cloud provider & consumer relationship and to avoid SLA (Service Level Agreement) violation in the cloud. To achieve the purposes, polarity classification or sentiment analysis of opinions of different service consumers regarding different service providers is to be performed using deep learning approaches.

1.3. Definition of Problem

Within the cloud paradigm, an ideal cloud relationship is formed if the customer is satisfied with the service provider in terms of resources, storage, time and cost requirements etc. This condition could only be achieved if one could get deep into the consumer’s emotions and understand the requirements that would help to choose a particular service provider to satisfy it.

To achieve the purpose, sentiment analysis of customer/consumer's ideas is performed using the deep learning method.

Table 1.1. Comparative study of research approaches performing Sentiment Analysis [65]

Author	Year	Dataset Used	Research Approaches Used	ML/DL	Cloud Paradigm (Yes/No)
Zhu, L., et al. [1]	2020	Stanford Sentiment Treebank dataset and IMDB dataset	SentiVec	Deep Learning	No
Kale, M., et al. [2]	2018	News dataset	Conventional ML and CNN	Both	No
Abbas, S.S.I. [3]	2019	Turkish news text dataset	SVM and RNN	Both	No
Sohangir, S., et al. [4]	2018	StockTwits dataset	CNN, LSTM	Deep Learning	No
Naderalvojoud, B., et al. [5]	2017	German SentiWordNet lexicon	RNN	Deep Learning	No
Imran, A.S., et al. [6]	2020	COVID-19 Twitter tweets dataset	LSTM	Deep Learning	No
Kratzwald, B., et al. [7]	2018	6 benchmark datasets	Bidirectional LSTM and RNN	Deep Learning	No
Ghorbani, M., et al. [8]	2020	Social media reviews dataset	ConvLSTMConv	Deep Learning	Yes
Pathak, A.R., et al. [9]	2019	IMDB dataset & Restaurant reviews	DNN and LSTM	Deep Learning	No
Çano, E. and M. Morisio. [10]	2018	Song lyrics, IMDB movie reviews and Amazon reviews	NgramCNN	Deep Learning	No
Markoski, Filip. [11]	2020	Twitter tweets dataset	CNN and Graph Structures	Deep Learning	No
Hassan, A. and A. Mahmood. [12]	2017	IMDB dataset and Stanford Sentiment Treebank dataset	ConvLSTM model	Deep Learning	No
Li, L., et al. [13]	2020	Online Website reviews dataset	CNN, RNN and LSTM	Deep Learning	No
Qaisi, L.M. and I. Aljarah. [42]	2016	Twitter tweets dataset	Naïve Bayes	Machine Learning	Yes
Agüero-Torales, M., et al. [43]	2019	TripAdvisor reviews dataset	VADER	None	Yes

1.4. Hypothesis

Consumer's sentiments can have a more significant impact on the potential consumer's decision to choose the service provider. The cloud consumer always needs an optimal decision system that guides a consumer in service selection. The right choice can depend on the consumer's preferences, service requirements, budget and other factors. This could only be possible to learn existing consumers' feedback. According to the thesis, the problem could be best handled with Deep learning algorithms. Deep learning, which works on the principle of ANN (Artificial Neural Network), can maximize unstructured data utilisation. It executes feature engineering by itself, which enable us to get optimal prediction results for a better decision making process. The proposed system can work to assist the service consumer in deciding on an optimal service provider.

1.5. Advantages and Disadvantages of Proposed Approach

Keeping in view the Table 1.1, authors have performed sentiment analysis of textual data from different social media websites, online reviews and text datasets. They carried it out using deep learning techniques. However, none of them performed aspect-based sentiment classification of online cloud reviews using a combination of RNN-based techniques of LSTM and GRU. None of the articles mentioned the best approach through a comparative analysis. Addition of multi-criteria decision model is also not observed in the literature. Some of the features of our proposed approach are as following:

- Our proposed approach highlights important aspects from reviews that are crucial for a good cloud service provider. This approach is not found yet in literature.
- It identifies the sentiments of aspects and calculates the sentiment polarities.
- It compares results achieved from three proposed deep approaches. This gives us the best approach to perform ABSA in cloud reviews. Moreover, it compares the values of classification metrics and derives the most accurate and efficient model.
- AHP model assists to find the most important aspect according to cloud service consumers. It is, for the first time, a combination of ABSA and an AHP decision model to make decisions.

1.6. Structure of Thesis Study

The thesis is composed of seven chapters. Chapter 1 is the introduction chapter describing the importance of the thesis, the purpose, problem statement and hypothesis. It also includes a literature review conducted for the study. Chapter 2 contains information regarding Sentiment Analysis, its importance and levels of sentiment classification. Chapter 3 describes Cloud Computing, cloud models including IaaS, PaaS and SaaS, Quality of Service (QoS) and Service Level Agreements (SLA). Chapter 4 explains the methodology used, i.e. Deep Learning methods of RNN, LSTM and GRU for sentiment classification purposes. The implementation of deep learning techniques and the results obtained are mentioned in Chapter 5 of the thesis. Chapter 6 highlights the AHP decision-making model used for the decision-making task of Cloud Provider Selection. However, Chapter 7 is the last chapter containing the conclusion.

2. SENTIMENT ANALYSIS

Sentiment Analysis is a critical task in Natural Language Processing and Information Extraction that seeks to evaluate many documents and extract the writers' sentiments represented in the form of positive or negative remarks, queries, and experiences. “Hello everyone, we are much too happy and expect to be happier tomorrow,” for example, is a positive text. In reality, the goal of sentiment analysis is to determine what a speaker or writer thinks or feels about a topic or a document's general functioning. Opinion mining is another term for sentiment analysis in the literature. Sentiment Analysis's primary goal is to determine the polarity of a stated viewpoint in a text, whether positive, neutral, or negative.

Technorati, a social media tracking business, has released statistics on internet usage by internet users. According to their findings, four out of every five Web users are a potential source of social media usage in some way. This might include social networking sites and applications such as Facebook, Ablo, blogging and micro-blog sites, and content and video sharing apps such as Twitter. The evolution of the WWW into a socio-participative and co-creative Web is fascinating to see. Users can contribute in a variety of ways to the source. Even those who are new to the mechanics of Web publishing can create Web content. In reality, a Website's value is now determined by its number of users, representing the amount of data.

2.1. Text Classification Sentiment Analysis

As one means that users contribute to internet services and commodities are by sharing their opinions and experiences. There are a lot of websites that allow consumers to leave a review for a product or service. User evaluations can be found all over the internet for various items ranging from technology such as cell phones to luxury such as vacations and hotel accommodations. It should be emphasized that these evaluations are a possible source for harnessing collective intelligence and a source of user opinions. A traveller looking for a place to stay, such as a hotel, in his or her visiting city should go to the hotel's website and read the reviews of the available hotels in the city before making a reservation at that hotel. Alternatively, a user interested in purchasing a specific house item may want to read the evaluations left by many other users before making a purchasing decision. This procedure not only assists customers in receiving additional and relevant information about numerous services with a single click, but it also assists them in making a definite and trustworthy selection. Some people pretend to have their experiences written about the findings in a blog post instead of a detailed review. However, reviews are conveyed in texts in both situations. A general process of text sentiment analysis is given in Figure 2.1.

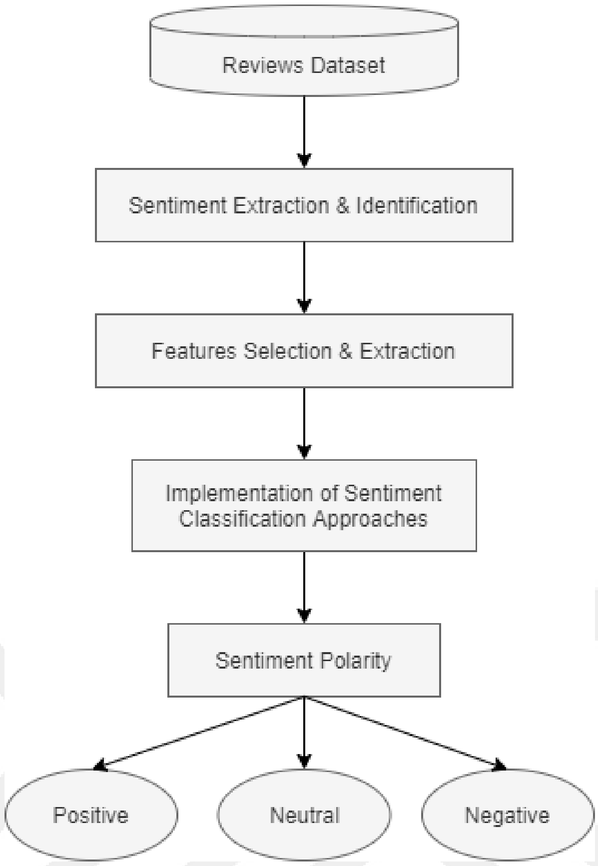


Figure 2.1. An overview of Sentiment Analysis process of a review dataset [10]

Even though experiences and reviews shared are incredibly beneficial and informative, a new user may find it challenging to read all of the reviews/posts in a small amount of time. A remedy is derived in the shape of a consequent comprehensive summary as the kernel of a volume of reviews to cope with such information overload situations. In today's world, the new Information Retrieval formulations, also known as sentiment classifiers, have two tasks: labelling or declaring a text or a review as positive or negative and removing and emphasizing the good and negative aspects of the offering. In its numerous connected disciplines, sentiment analysis finds its relevance in Information Retrieval based formulations. It systematically collects various comments about an item and emphasizes the essential aspects of those reviews as good or bad qualities.

2.2. Levels of Sentiment Classification

As stated recently, Sentiment Analysis examines a text or document and categorizes it into positive, negative, or neutral categories based on the thoughts conveyed in the text. Sentiment analysis has been studied on various language levels, the most common of which are Document-level, Sentence-level and Entity/Feature-level or Aspect-based Sentiment Analysis.

2.2.1. Document-level Sentiment Analysis

At this level, the objective is to determine if a whole document reflects either a positive or negative image. The system, for example, assesses if a product review reflects a generally favourable or unfavourable impression about the product. This level of analysis implies that each document reflects a single entity's viewpoint. As a result, it cannot be used in papers that assess or compare several entities.

According to the author [33], this sentiment analysis implies that the documents convey feelings about a particular item, such as product reviews, movie reviews, and hotel reviews. The typical length of a document varies on the domains, and the findings of document-level sentiment analysis typically contain two or possible three outcomes. Besides, in certain situations, such as news domains, a single publication may include many viewpoints and criticize multiple targets.

2.2.2. Sentence-level Sentiment Analysis

Sentiment Analysis at the sentence level is more fine-grained than Sentiment Analysis at the document level, and each sentence has conveyed a neutral, positive, or negative view. Because the sentences are short texts, there is no significant difference between the two degrees of sentiment analysis presented. Subjectivity classification is a crucial element of this level, in which objective phrases expressing accurate information are separated from subjective words expressing feelings. Furthermore, each sentence is structurally and emotionally linked to the rest of the text [33]. As a result, both domestic and international relevant information is needed for this level assignment. However, various methods are used to complete the assignment, including conditional, inquiry, and even complex sarcastic sentences.

Sentiment categorization at the document level is frequently used to determine the prevalent sentiment towards the whole topic. It considers the object's various angles. For example, when a document is favourable in a movie review, it implies that the writer enjoys the film in general. However, if he focuses on various things, such as the plot or the performers' acting, his comments may be ignored even if they are negative. The outcome of extracting each characteristic and related opinion from each subjective sentence in this example is a Sentence-Level sentiment categorization. To achieve an accurate analysis of an item, sentence-level sentiment categorization is necessary [62]. For example, in the case of products, such a detailed study is required to enhance products by determining which products' characteristics (components or qualities) are loved and hated by customers. Document-Level Sentiment Analysis does not provide this information.

2.2.3. Aspect-based Sentiment Analysis

Aspect-based Sentiment Analysis, also termed Entity-based/ Feature-based Sentiment Analysis, is the most fine-grained sentiment classification method. Except for the previously mentioned two levels, aspect level sentiment analysis investigates how the holder feels about the target, including whether they love or dislike it. Producing a forecast gathers information on a given entity, typically with numerous features, qualities, and viewpoints. It emerges on product review websites or discussion forums in the products reviews. It is a more complicated work than previous levels of analysis because it involves three steps for sentiment classification: target characteristics identification, evaluating feature-wise polarity, and overall assessment summarization [33]. Furthermore, entity-based sentiment analysis necessitates more advanced machine learning techniques, as simple algorithms cannot deal with complex phrases.

For example, let us consider a hotel review from a tourist regarding the services and location of the hotel; “Although the services offered were great, the location is not good”. Since the first part of the sentence gives a favourable opinion while the second portion is negative, the average sentiment classifier classifies the sentience to be neural, thus losing the primary information. Therefore, it is necessary to anticipate sentiments on aspects or entities. For the above example, services and location are aspects while their sentiments are positive and negative, respectively [63].

3. CLOUD COMPUTING

Cloud computing, commonly known as "The Cloud," is a new technology that has emerged in today's society. It is indeed a type of internet-based computing emphasizing data management and processing. Cloud computing is a computational paradigm that allows users to store, analyze, and retrieve their sensitive data securely. The term "cloud" was used to develop the concept that continues to move and acts as a common medium. These clouds are essentially datacenters, and they are in charge of all data-related duties. As a result, cloud computing encourages us to outsource the processing of vital data. A client from a distant area may quickly access files in the cloud via cloud computing. There is no need to employ any infrastructure for information storage or computation purposes.

The introduction of cloud computing is nothing short of technological development and miracle. It has enabled the internet to provide on-demand access to a variety of scalable and flexible resources. Computing and storage resources are being reduced as much as possible. Cloud computing improves user agility, flexibility, and data availability at a much more reasonable and affordable cost. As per the definition of Cloud Computing defined by the National Institute of Standards and Technology (NIST),

"Cloud computing is a paradigm for providing easy, on-demand access to a common stock of customizable computing resources, such as hosts, connections, memory, processes, and applications, which can be instantly provided and released with minimal administrative effort or cloud provider's involvement".

Cloud computing has caused a revolution in today's world due to its remarkable cost-effective, diversified and scalable features. Based on its different services, cloud computing can be classified into three main service model categories. They are usually named on their services as Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

3.1. Cloud Service Models

3.1.1. Infrastructure-as-a-Service (IaaS)

IaaS (Infrastructure as a service) gives charged access to basic computational resources such as online and offline servers, networking, and memory through the web. End-users may expand and decrease required resources as required using IaaS, which eliminates the necessity of large, ahead of operating expenses, unneeded on-premises or 'owned' infrastructure and overbuying facilities to meet periodic spikes in demand.

3.1.2. Platform-as-a-Service (PaaS)

Within Platform as a service (PaaS), users are supplied with a platform equipped with hardware, a software package stack, development frameworks, and different software tools to create operate, manage, and test applications without the expense, complexities, or rigidity of doing so on-premises.

3.1.3. Software-as-a-Service (SaaS)

Software as a service (SaaS) distributes software programs online, on request, and usually for a fee. Cloud service providers host and maintain the software application as well as the framework using SaaS. These companies also take care of routine maintenance, such as software upgrades and security updates. Users use the app through the internet, generally using a web browser on their phone, tablet, or computer. A general hierarchy for cloud service models of IaaS, PaaS and SaaS is given in Figure 3.1.

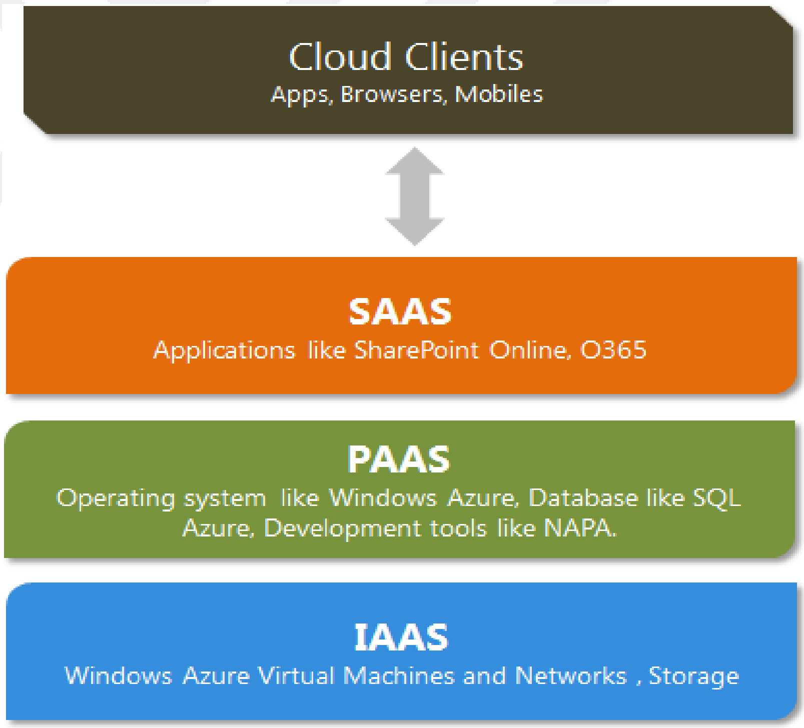


Figure 3.1. Examples of Cloud Services Models [64]

3.2. Service Level Agreement (SLA)

Service Level Agreement (SLA) is the key document between the service provider and a service consumer that binds them for a certain period and the commitment of defined services. In the cloud environment, a provider-consumer relationship is essential to trust and be reliable following a set of SLA. A contractual association between provider and consumer lays out the contract terms for providing cloud services such as memory, software, and computing. The measures must be taken in the event of an SLA infringement. The cloud service provider may manage its resources with the help of SLA. As a result, SLA ensures QoS.

For a viable SLA and to ensure a successful and trustful relationship between the two parties, one of the case-sensitive factors is selecting a cloud service provider. For that purpose, the selection of a cloud service provider may affect the relationship. A cloud service provider must be reliable, scalable, ease administration, provide backup and storage facilities, data management and security, etc.

3.2.1. Quality of Service (QoS)

One of the essential elements of an SLA is Quality of Service (QoS), which allows for developing a reliable connection. Its characteristics, such as trust, security, confidentiality, resource management, risk mitigation and assessment etc., come together to establish a sustainable SLA between the service provider and the customer. A trusted cloud connection is built on a QoS repository and evaluation system and a committed provider for a customer who reserves resources whether the consumer utilizes them or not. Any uncertainty in agreement terms and wording results in consumers having difficulty meeting their demands. As a result, SLA modelling and monitoring techniques are used to analyze the SLA idea comprehensively.

4. METHODOLOGY

Text Sentiment Analysis is performed using differently supervised and unsupervised learning approaches, including Lexicon-based methods, Machine learning and Deep Learning algorithms. Deep Learning approaches such as RNN, LSTM and GRU have distinct architectures and perform sentiment analysis with accuracy. With the brief introduction of deep learning, its various types are described in this chapter.

4.1. Deep Learning

Deep learning is a subclass of machine learning that is generally a multi-layered neural network. These neural networks are designed to mimic the human brain's function by learning from huge volumes of data. Predictions made by a single-layer neural network improve accuracy and enhance using multiple hidden layers in the architecture. Many artificial intelligence (AI) apps and services rely on deep learning to increase automation by completing cognitive and physical activities without human interaction. Every day goods and services like chatbots, audio TV remotes, fraudulent credit card identification, and innovations use deep learning technology.

In today's era, Deep Learning algorithms are used to depict the functionality and structure of a brain, known as Artificial Neural Networks (ANNs). Deep Learning enables prediction output for the given input parameters and creates patterns that become helpful in decision-making processes. It is used in different technologies like extracting and resolving Big Data from heterogeneous platforms. Artificial Intelligence (AI) algorithms are used to identify patterns and make decisions according to user requirements. Deep Learning is also used for text analysis and data mining purposes, particularly the LSTM (Long Short Term Memory) method that enhances sentiment analysis accuracy.

In deep learning, a simulation executes classification tasks based on various pictures, text, or audio or video inputs. Advancements in the results obtained through deep learning models continue to attain accuracy, even surpassing human performance in some cases. Vast volumes of labelled data are used to train the multilayer neural network model and is tested over any chunk of the dataset to verify the learning obtained.

In today’s world, a vast amount of data and information are generated and transferred from one end to another. This data, termed Big Data, is extracted from different sources and is unstructured. Thus, extracting relevant, meaningful information from Big Data is quite a difficult task for humans. Deep learning resolves it by feeding data to the neural network and then making decisions on its own similar to human decision makings.

Deep learning can be differentiated from machine learning so that Machine learning takes a set of structured/labelled data, trains itself, learns from the data, and makes decisions based on what it has learned from the results of trained data. However, Deep learning does not require any structured/labelled data to learn; it feds the data into different layers of ANN that learns it and makes intelligent human-like decisions. On the other hand, deep learning requires a huge amount of data for training since decisions are made when millions of data points in the ANN interact and consume more time and hardware resources (GPUs). Machine learning performs training on less data as compared to DL in less time and hardware.

4.2. Neural Network

A deep learning architecture consists of a neural network model known as Artificial Neural Networks (ANNs). Deep neural networks are made up of several layers of linked nodes, each of which improves and refines the forecasting or classifications. The structure of the human brain inspires the Neural Network architecture. Neural networks can be trained to do the same operations on data that our brains do when identifying patterns and classifying various sorts of information. Like a human brain, each neural network layer filters the operations from coarse to fine, enhancing detection and output accuracy. They are helpful to accomplish a variety of tasks, such as classification, clustering and regression. Neural Networks can be used for unlabeled data to label and group them based on similarities between the samples. Alternatively, in classification, the network can be trained on a labelled dataset for categorizing the dataset samples.

Forward propagation refers to the flow of calculations via the network starting from the input layer and proceeding towards the output layer through multiple hidden layers. The visible layers are the input and output layers of a deep neural network. The input layer accepts the incoming nodes for processing stages, while the final classification results are obtained through the output layer. Hidden layers perform all the internal processing using different AI algorithms. Within the network, input signals are transmitted layer to layer, having weighted connections until they approach the output layer. A nonlinear function could be activated in certain neurons. The objective of a learning process is to identify weights that will cause the neural network to behave in a specific way. Backpropagation is a method of training a model that employs techniques such as gradient descent to determine prediction errors and then changes the weight values of the function by going backwards in the network. Both types of propagation work together to enable a neural network in generating predictions and rectify any mistakes. The general architecture of a 3-layered Artificial Neural Network (ANN) model is described in Figure 4.1.

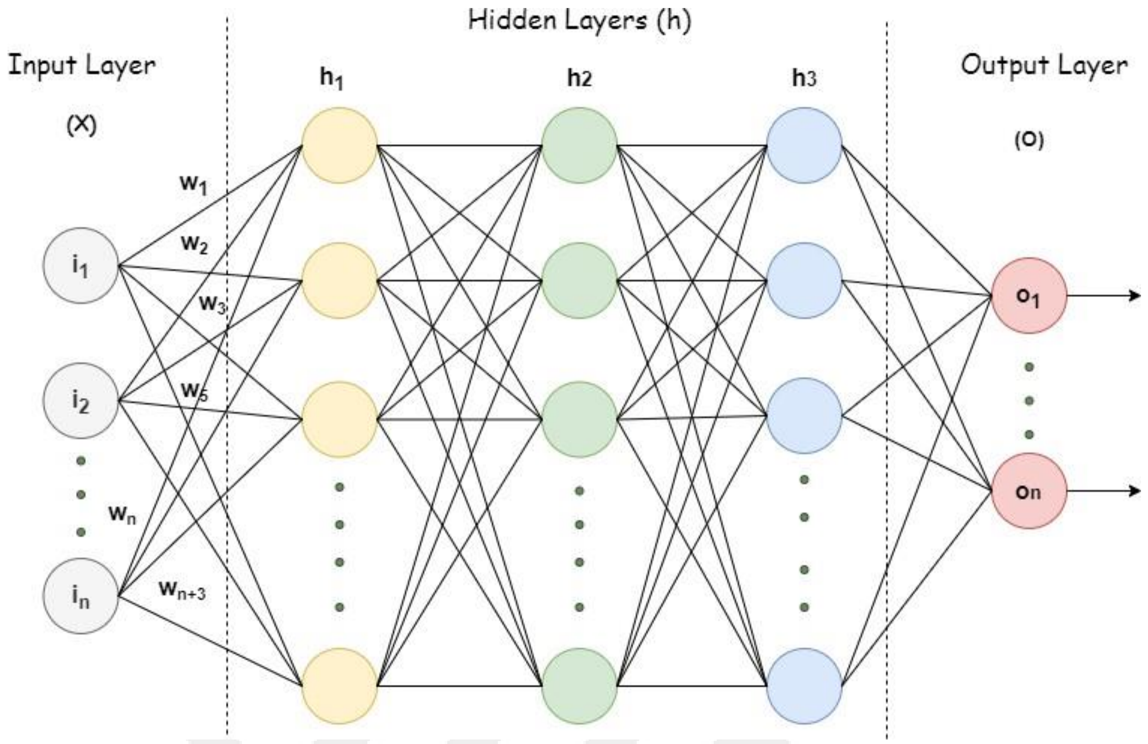


Figure 4.1. An architecture of three-layered Artificial Neural Network (ANN) [22]

Each input node in an ANN network is given a value weight, denoting its relevance for a scenario. A cost function is determined whenever a collection of input nodes is oriented towards an output node or perceptron. The total of all nodal inputs with their corresponding weights is the cost function. The total outcome is added to a threshold value, which is referred to as bias. It consists of a constant value that adjusts the model to be suitable for the given dataset. In an algebraic expression, a cost function of a neural network is calculated in equation 4.1 below. Node inputs are denoted as 'i' each with individual weight 'w', and the suggested threshold value is '$bias$'.

$$\sum_{a=1}^{n}(i_a w_a) + bias = i_1 w_1 + i_2 w_2 + \cdots + i_n w_n + bias \tag{4.1}$$

The expression represents the output function '$f(o)$' having value either '0' or '1' needed for the other decision-making process.

$$output = f(o) = \begin{cases} 0, if \ \sum_{a=1}^{n}(i_a w_a) + bias \leq 0 \\ 1, if \ \sum_{a=1}^{n}(i_a w_a) + bias > 0 \end{cases} \tag{4.2}$$

In ANN model architecture, the activation function is vital. The activation functions perform a nonlinear change on the input signal to learn and do more complex nonlinear tasks. Nonlinear activation functions are often utilized in neural networks. Sigmoid, Rectified Linear Unit (ReLU), and Hyperbolic tangent (tanh) are popular activation functions used in neural networks. A loss function, also known as a cost or objective function, is used to determine the potential of a statistical model's prediction. The loss function's basic premise is to calculate the error rate across anticipated and accurate target values. As a result, the output of the loss function should be reduced to obtain the most refined algorithmic model.

4.3. Types of Neural Network

Neural Networks can be classified into many different types based upon the structure of the network and the number of neurons it contains, neuron density, the number of layers and the flow of data through a network. Within the scope of the study, the three main types of NN are discussed as follows:

4.3.1. Recurrent Neural Network (RNN)

A Recurrent Neural Network is an artificial neural network in which the output of one layer is stored and recirculated to the input. This contributes to predicting the layer's outcome. The feedforward network's initial layer is created similarly. Such that, the total of the weights and characteristics is multiplied. The recurrent neural network process, on the other hand, begins at later levels.

RNN is a sort of Feedforward Neural Network that recognizes patterns in sequential input to make predictions. Feedforward NN has several disadvantages, including the fact that it does not utilize sequential data, has no internal memory, and only considers the current input [65]. As a result, it uses its internal memory by storing the output data through one operation and transferring it into another for analysis and calculating the total output. An overview of the Recurrent Neural Network structure is shown in Figure 4.2.

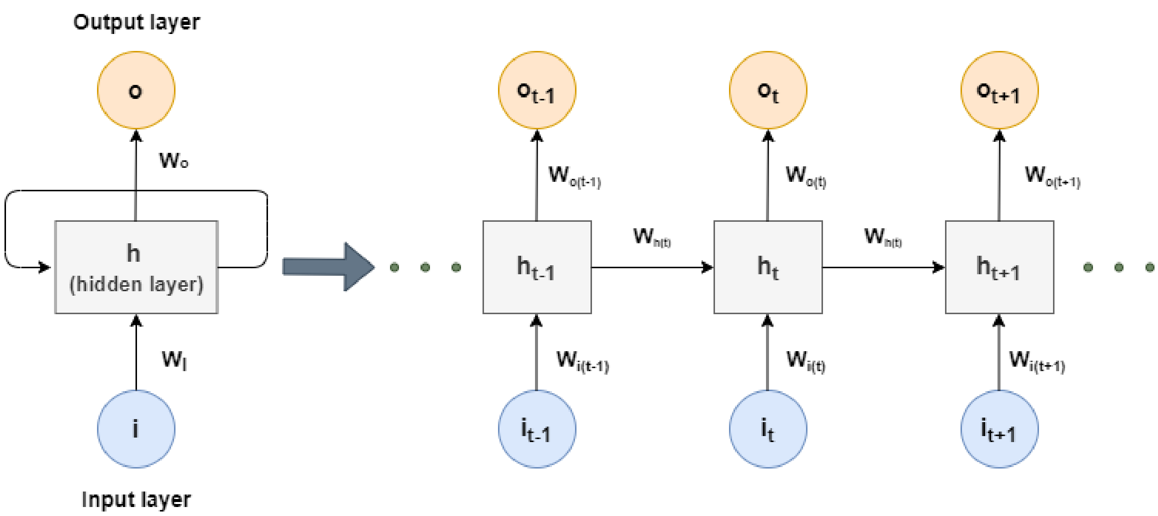


Figure 4.2. An overview of RNN structure [30]

In the figure above, 'i' is the network's input layer, hidden layer is denoted by 'h' while 'o' is the output layer. For a time instance 't', two inputs, i.e., i_t, the current input state and h_{t-1}, the output from the previous state is fed into a new RNN cell and generates a new current hidden state h_t. Thus, the below formula is the network's current state.

$$h_t = f(h_{t-1}, i_t) \tag{4.3}$$

By applying the Hyperbolic tangent activation function, the equation transforms into the below form where weight from the last hidden state is W_h and weight of current input is W_i.

$$h_t = tanh(W_h * h_{t-1} + W_i * i_t) \tag{4.4}$$

The net formula for the output layer o_t where W_o is the weight of output state and h_t is the current hidden state will be

$$o_t = W_o * h_t \tag{4.5}$$

Since RNN processes time-series data (sequential data), it finds applications in language translations, speech and image recognition, sentiment analysis, natural language processing and time-series prediction purposes. Depending upon the number of inputs and outputs in a network, RNN is classified into four types. Figure 4.3 shows types of RNN.

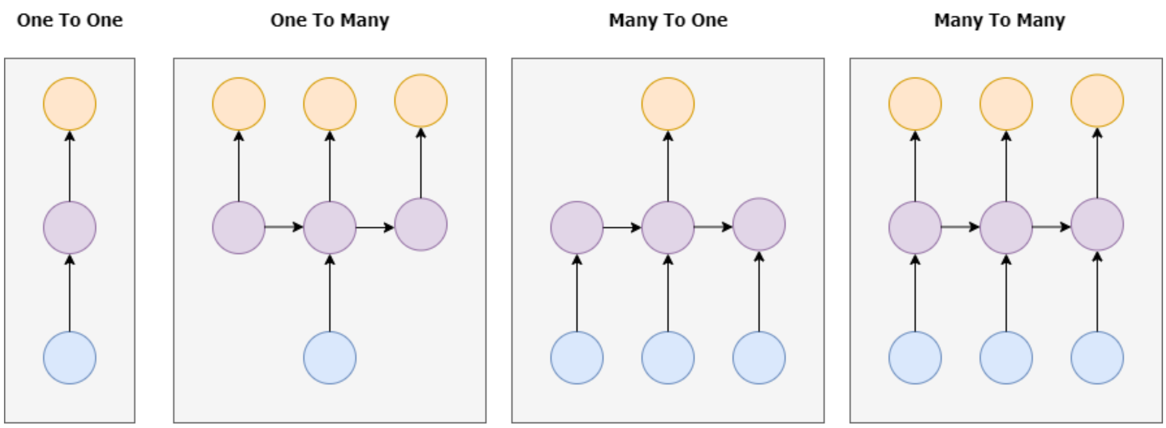


Figure 4.3. Types of Recurrent Neural Network [46]

RNN models must be given training for backward propagation once forward propagation is done. Backpropagation through Time (BPTT) is an augmentation of backpropagation and is used for this objective. Time is a crucial factor by linking one time step to the next, calculating a sequence of weight gradients, and guaranteeing a backpropagation process. The backpropagation technique may be costly in terms of operating time and computing cost for too lengthy sequences. As a solution to the gradient vanishing and expanding gradient difficulties, a

Truncated BPTT is employed, which truncates the feedback link and causes the gradient to circulate, preventing the network from learning complete dependencies and solving the exploding gradient problem. The choice for an appropriate activation function is the solution of the vanishing gradient problem. Alternatively, another ANN architecture, known as LSTM, is used to handle the issues efficiently.

4.3.2. Long Short-Term Memory (LSTM)

The LSTM is a sequential network, a more sophisticated RNN that allows data to survive long in the network. It can deal with the vanishing gradient problem since it was a problem in RNN. It addresses the issue of RNN's long-term dependency and backpropagation for model training, where the RNN cannot anticipate words stored in memory but can make more accurate predictions based on current data. In the cases when the gap length is extended, LSTM outperforms RNN. The control flow of an LSTM is similar to that of a recurrent neural network. It processes data and passes the information on as it moves along. The processes within the cells of the LSTM only change. The LSTM uses these processes to remember or forget information. Using time-series data, it is utilized for processing, forecasting and categorizing.

The cell state and its numerous gates are at the foundation of LSTMs. The cell state serves as a transportation route for relative information as it travels down the series chain. In principle, the cell state can hold meaningful information across the sequence's processing. Information is added or deleted from the cell state through gates as the cell state flows. The gates are several neural networks that determine whether information about the cell state is permitted. During training, the gates might learn what information is important to preserve or ignore. The structure of LSTM is described in Figure 4.4.

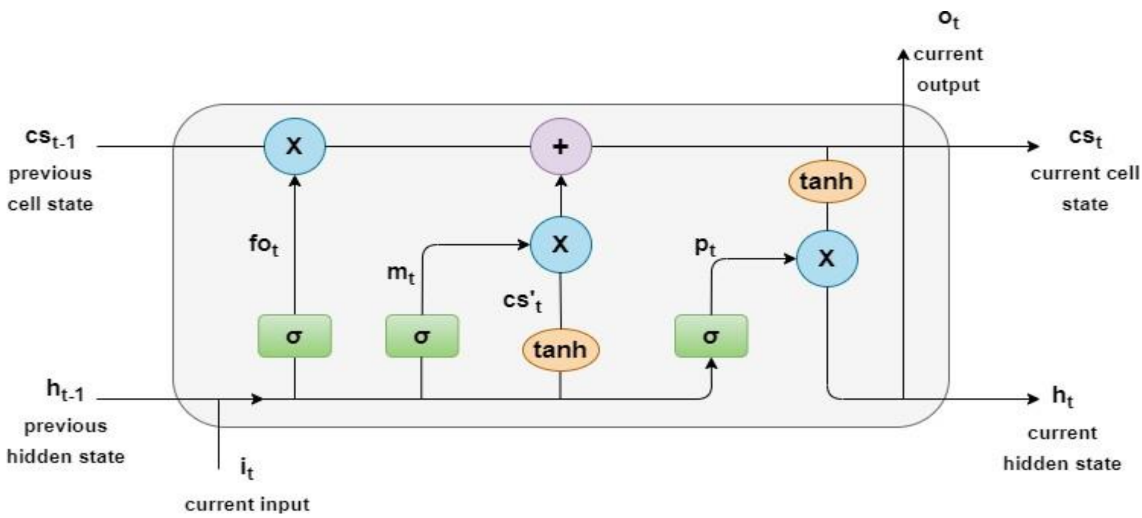


Figure 4.4. The structure of LSTM Neural Network [55]

The decision of values depends upon the results from activation functions such as Sigmoid, which squeezes values between 0 and 1, while tanh activation function squeezes values between -1 and 1. Proceeding towards the gates, there are three gates in the LSTM cell. Forget gate determines whether information should be discarded or saved. The sigmoid function passes previously hidden state data as well as current input information. The values are either 0 to forget or 1 to save. The equation for forget gate is given where W_{fo} represent the weights of current input and previous hidden state.

$$fo_t = \sigma\ (W_{fo} * [h_{t-1}, i_t] + bias_{fo}) \tag{4.6}$$

It is divided into two sections to decide the amount of data that should be kept in the cell state. A sigmoid layer is used to send the conjunction of i_t and h_{t-1}, to return 0 or 1. The very same data is then sent into the tanh function, which yields $cs_t{}'$, which is subsequently inserted into the cell state. The following are the input gate formulas:

$$m_t = \sigma\ (W_m * [h_{t-1}, i_t] + bias_m) \tag{4.7}$$

$$cs_t{}' = \sigma\ (W_{cs'} * [h_{t-1}, i_t] + bias_{cs'}) \tag{4.8}$$

To bring the prior cell state cs_t' up to date, the product of fo_t and cs_{t-1} forgets useless data while the product of $cs_t{}'$ and m_t achieves new data. The resultants from both operations are summed as:

$$cs_t = (fo_t * cs_{t-1}) + (m_t * cs_t{}') \tag{4.9}$$

Finally, the output gate shows whatever the input and cell have decided to output. The input is sent via a sigmoid layer, which returns a value of 0 or 1. Besides, the cell state is processed through a tanh layer, which assigns weight to scores ranging from -1 to 1 based on their significance. The hidden layer data is then multiplied by the sigmoid layer outcome.

$$p_t = \sigma\ (W_p * [h_{t-1}, i_t] + bias_p) \tag{4.10}$$

$$h_t = p_t * tanh(cs_t) \tag{4.11}$$

4.3.3. Gated Recurrent Unit (GRU)

GRU, which stands for Gated Recurrent Units, is an extended version of LSTM. Its main functionality includes handling vanishing gradients and exploding gradients issues encountered in the RNN structure. GRUs are LSTM that uses a gating mechanism to control the flow of information inside and between cells. Instead of forgetting, input and output gate, GRU contains the update gate and the reset gate. The vanishing gradient problem may be controlled with the help of these gates. Unlike LSTM, GRU preserves a hidden state rather than the cell state. As a result, small datasets can be trained easier and quicker, and RNN's memory capacity can be increased using GRU. The structure of GRU is described in Figure 4.5.

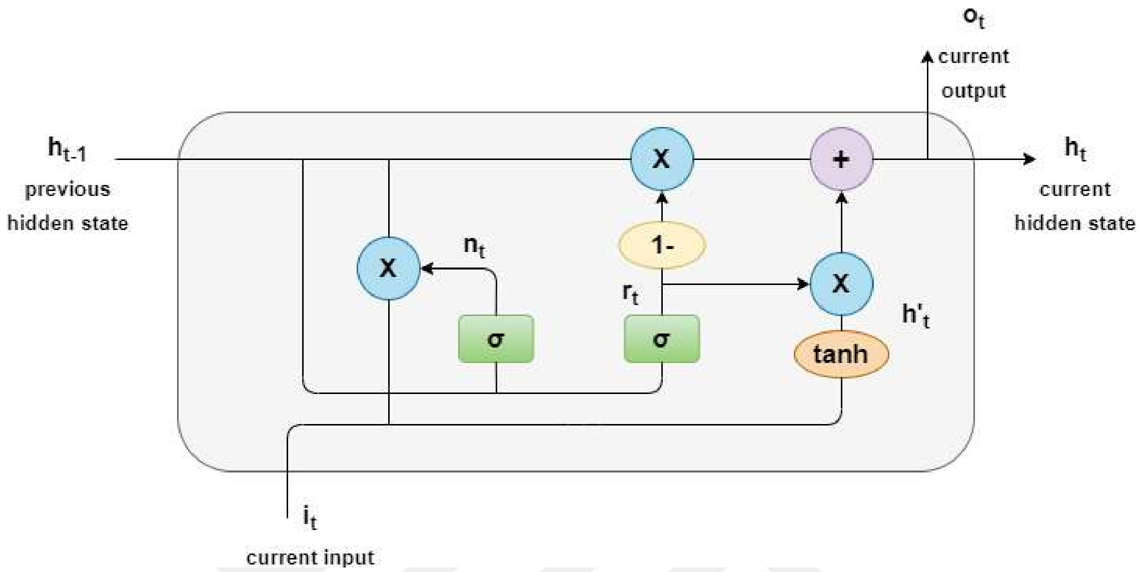


Figure 4.5. The structure of GRU Neural Network [59]

The Reset Gate defines the amount of prior information to be forgotten or ignored within the cell. It is similar to how an LSTM recurrent unit's Forget Gate works. It manages the short-term memory of GRU and represents its hidden state. It categorizes irrelevant data and informs the user to proceed without using that data. The equation for reset gate is given as follows:

$$n_t = \sigma\,(W_n * [h_{t-1}, i_t] + bias_n) \tag{4.12}$$

The Update Gate assists the model to decide how much previous information must be handed on to future time steps. This is quite powerful since the model may choose to copy all of the features from the past, thereby eliminating the risk of fading gradients. It handles GRU's long-term memory. The equation for the update gate is given below:

$$r_t = \sigma\,(W_r * [h_{t-1}, i_t] + bias_r) \tag{4.13}$$

A temporary candidate hidden state h'_t is calculated, which helps in determining the current hidden state h_t. The most crucial aspect of this equation is that the reset gate's value regulates how far the earlier hidden state may impact the candidate state. If n_t value is 1; it implies that the whole information from the preceding hidden state h_{t-1} is taken into account. Similarly, if the value of n_t is 0, the preceding hidden state information is totally ignored.

$$h'_t = tan\,h\ \ (W\ [\, n_t * h_{t-1},\ i_t\ ]\,) \tag{4.14}$$

So using the value of candidate hidden state, the value of update gate is taken into account and the equation for current hidden state h_t is derived as follows:

$$h_t = (1 - r_t * h_{t-1}) + (r_t * h_t') \tag{4.15}$$

GRU is superior to LSTM since it is simple to alter and does not require memory units. As a result, it is easier to train than LSTM and to provide according to performance. If the data series is tiny, GRU is used; when the data series is more significant, LSTM is used. LSTM does not expose the whole memory or veiled layers, while GRU does. In complicated domain areas like machine translation, stock market prediction, sentiment classification and speech recognition, LSTM and GRU are necessary.

5. IMPLEMENTATION AND EXPERIMENTAL RESULTS

This chapter includes the description of all experiments performed and results of sentiment analysis of cloud consumer reviews dataset. Its sections include a Flowchart of ABSA methodology, a description of the Cloud dataset, including the HaaS system. Statistics regarding the cloud reviews dataset is also presented. The third section elaborates on the aspect extraction process of the cloud reviews. Classification metrics and their values for this study are also calculated in this section. Figure 5.1 represents a general flowchart of the proposed Aspect-Based Sentiment Analysis process.

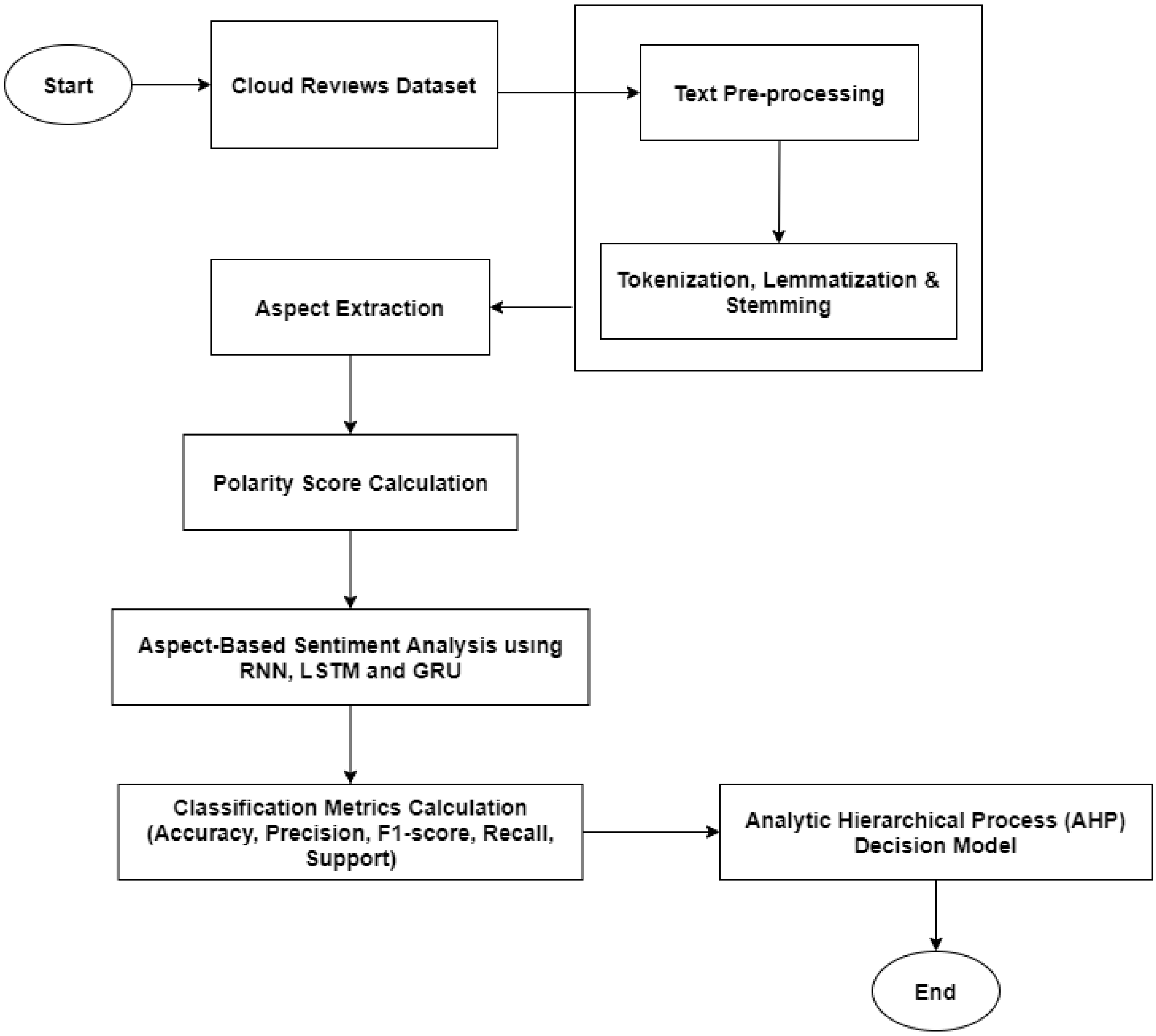


Figure 5.1. Flowchart diagram of proposed ABSA process

5.1. Data Collection

Data collection for the thesis is carried out using Harvesting-as-a-Service (HaaS) Framework [66]. HaaS is an intelligent harvester that collects cloud services (IaaS, PaaS, SaaS) data from disparate Web portals across the internet. It can collect web pages with a variety of structures without the requirement for coding. HaaS is a service-based harvester with a user interface that is simple to use. Individuals who want to collect cloud services portals using the harvester don't need any programming knowledge. Due to this purpose, it is termed as "intelligent". Within the harvester, the Cloud Services Repository is used to hold the information on cloud services that have been obtained and downloaded. It covers all of the information on cloud services, including data about cloud service offers. Users can get this data as a CSV, PDF, or SQL file. For the thesis, the cloud reviews dataset is obtained through the data harvested from three different web platforms: The Getapp, CloudReviews and Serchen web portals.

5.2. Dataset Description

A Cloud Reviews dataset is used for this thesis. The dataset contains a collection of online cloud service consumer reviews about different topics, including cloud services and various cloud service providers. These reviews describe and help service consumers to identify the nature of cloud service providers. Service consumers tend to elaborate their experiences regarding specific providers from whom they were availing cloud services. Services may include storage services, infrastructure domains and application terms. All the reviews are in textual format and in user-understandable English language. Each post in the dataset shows different attributes, e.g. the review, reviewer name, date of online review posting, the category and service name, review rating, number of reviews etc. All these reviews are extracted from heterogeneous web sources through the HaaS system [66]. The reviews dataset contains 6259 cloud reviews from different web sources. The whole dataset is divided into three classes with five detailed review ratings. Reviews having ratings from 4 to 5 are considered positive reviews. Reviews of 3 as review ratings are neutral. However, negative reviews have ratings ranging from 1 to 2.

Upon classification by sentiment analysis, the cloud reviews dataset has 5855 positive, 139 neutral and 264 negative reviews. K-fold cross validation technique is used for training and testing datasets where k=5 folds. Dataset is divided into 5 equal sets, each containing 4 sets of training and 1 set for testing purposes. However, a general tabular description of a few cloud consumer reviews with their review ratings and categories is given in Table 5.1. A visualization describing different Cloud Service Providers and the number of reviews per CSP is given in Figure 5.2. However, Table 5.2 expresses the statistics of reviews in tabular format.

Table 5.1. An overview of cloud reviews dataset

Review	Category/Service Name	Review Rating
SiSense has consistently impressed our developers and users with everything you look for in a robust BI tool. From setup to dashboard use, SiSense provides us outstanding results for all of our use cases, inside and out. With its ability to pull large amounts of data from multiple data sources at impressive speeds, quickly create powerful dashboards that are accessible on any device, and provide excellent customer support, SiSense truly exemplifies itself as leader in Big Data industry.	Business Intelligence & Analytics/Sisense	5
We switched from SalesForce cases to Zendesk 16 months ago. The thing we were giving up was the ability to easily look up customer account information. Now with the new integration, we can push whatever fields we want (even custom ones) to Zendesk so our agents ever have to leave our Zendesk system. My only gripe is that the initial search lookup can be pretty slow (up to 10 seconds to load data), but I suspect this is more on the SalesForce side as opposed to something Zendesk could optimize.	Customer Service & Support/Zendesk	4
The more research you do, the more you will find similarities between email providers versus differences. HonestMail.net differentiates itself behind the scenes. As it is a newer service not beholden to legacy infrastructure, it is priced over 50% less than Constant Contact, iContact, Mail Chimp, and AWeber with nearly all of the same features.	Communications Applications/Constant Contact	3
We've been evaluating Podio for several months, and while they have a truly groundbreaking concept (create your own apps without programming), somebody really thinking it all through while they created it. The most appalling oversight is the lack of support for real (customers). While Podio makes it super-easy (perhaps too easy) to link collaborators with the apps you create to work with clients, clients themselves are a nightmare. BE CAREFUL. Those are being treated as collaborators, staff, not as real customers, and you'll rip your hair out trying various workarounds before realizing they just blew it, and there is no workaround. To their credit, Podio seems to be taking an interest in the gripes about this fatal design flaw in their user forums, and hopefully, will step up to the plate to get it fixed, but until then, it's just a nifty idea.	Project Management/Podio	2
This software has some flash features like sorting jobs on a map for the closest field worker. It's very clear that the basic functionality is full of bugs and un reliable. At present the system has un rounded invoices which randomly ad or remove 1cent from a rounded figure when gst is added. The mobile app crashes at random, work offline. Losses information even when it says saved. This app seems more like an impressive school project than a business grade product. Un reliable, unstable, not a complete app, un usable with incorrect tax invoices.	Operations Management/GeoOp	1

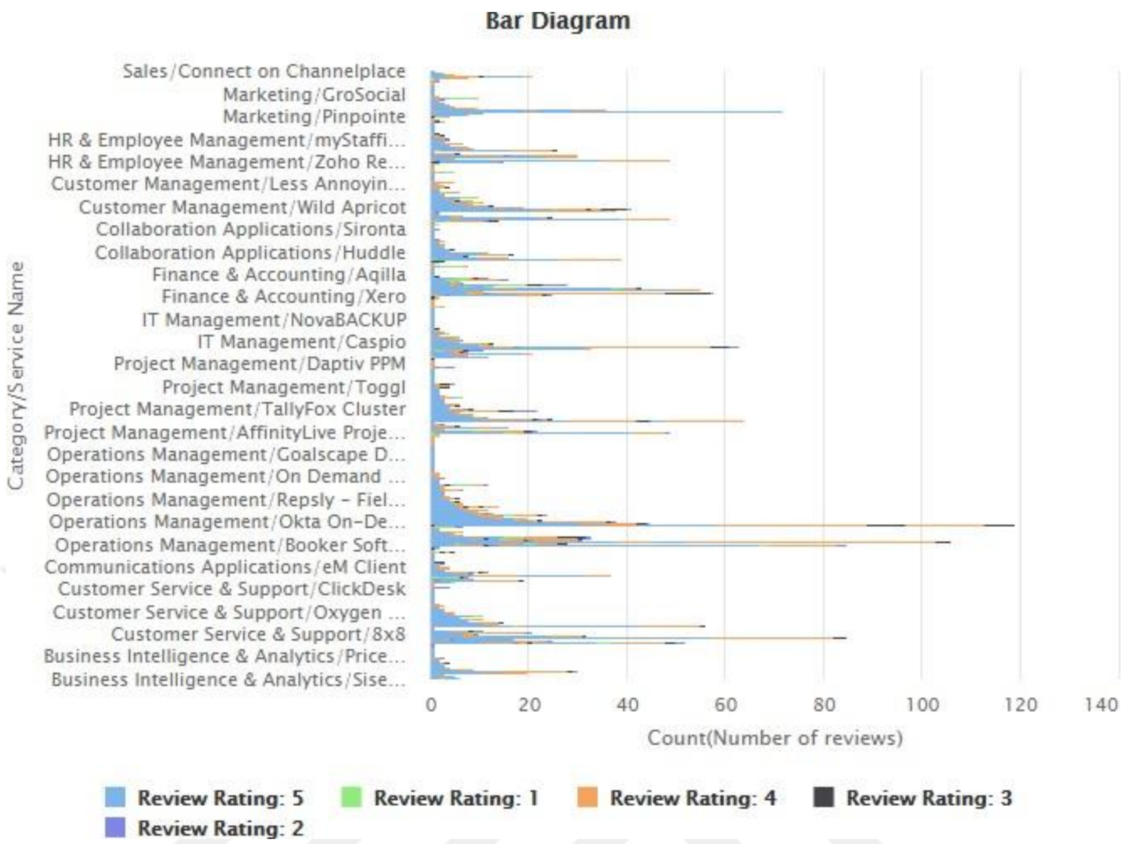


Figure 5.2. A Bar graph describing Cloud Service Providers w.r.t number of reviews

Table 5.2. Statistics of sentiment analysis of cloud reviews dataset

Statistics	Positive		Neutral	Negative	
	5	4	3	2	1
number of reviews	4617	1238	139	97	167

5.3. Pre-processing Steps

The pre-processing steps include tokenization, removal of stopwords and punctuation marks. Tokenization involves the conversion of text into tokens or words corresponding to their frequencies in the text. For that purpose, one of the vectorization methods of Sklearn text processing is used. CountVectorizer function converts the text into matrix of token counts. Each review text is taken into account and a matrix of token is generated based on the frequencies of words in the text. For instance, a review text about a Cloud Service Provider i.e. Cleanbill Litebooks is given as:

“LiteBooks has simplified the way we do business, easy to add clients, do quotation and invoicing via email. Best of it all its online”.

The above review has four sentences listed as:

[*“LiteBooks has simplified the way we do business,*

easy to add clients,

do quotation and invoicing via email.

Best of it all its online”.]

The first step, after applying the CountVectorizer, involves conversion of text into small chunks called tokens. It makes the text into the following format:

[*`LiteBooks`, `has`, `simplified`, `the`, `way`, `we`, `do`, `business`, `,`, `easy`, `to`, `add`, `clients`, `,`, `do`, `quotation`, `and`, `invoicing`, `via`, `email`, `.`, ` Best`, `of`, `it`, `all`, `its`, `online`*]

After tokenization, it’s time for removal of stopwords from the text. Stopwords are the words that add less or no meaning to the text. These may include conjunctions, articles, prepositions and pronouns etc. It also removes the punctuation marks in the text. An identified list of stop words in NLP helps to remove those words from text. As a result, the text is transformed into the following form:

[*`LiteBooks`, `simplified`, `way`, `business`, `easy`, `add`, `clients`, `quotation`, `invoicing`, `email`, ` Best`, `online`*]

All the words achieved are now unique and are arranged in alphabetical order to calculate their occurrence. All the letters of tokens are converted into lowercase. Table 5.3 shows the matrix formation of tokens identified in the text and their occurrence for each of the individual text sentence.

Table 5.3. Tokenization of a cloud consumer review

	add	**best**	**business**	**clients**	**easy**	**email**	**invoicing**	**litebooks**	**online**	**quotation**	**simplified**	**way**
[1]	0	0	1	0	0	0	0	1	0	0	1	1
[2]	1	0	0	1	1	0	0	0	0	0	0	0
[3]	0	0	0	0	0	1	1	0	0	1	0	0
[4]	0	1	0	0	0	0	0	0	1	0	0	0

Each token is given a particular index in the matrix. For instance, in Table 5.4, `add` is given as index 0, `best` is given as index 1, `business` is given as index 2 and so on. The matrix formed is called Sparse Matrix.

Table 5.4. Tokenization w.r.t indexes of tokens

	0	1	2	3	4	5	6	7	8	9	10	11
[1]	0	0	1	0	0	0	0	1	0	0	1	1
[2]	1	0	0	1	1	0	0	0	0	0	0	0
[3]	0	0	0	0	0	1	1	0	0	1	0	0
[4]	0	1	0	0	0	0	0	0	1	0	0	0

The vectors formed for each of the text sentence is as follows:

For 1st sentence, [0 0 1 0 0 0 0 1 0 0 1 1]

For 2nd sentence, [1 0 0 1 1 0 0 0 0 0 0 0]

For 3rd sentence, [0 0 0 0 0 1 1 0 0 1 0 0]

For 4th sentence, [0 1 0 0 0 0 0 0 1 0 0 0]

For our analysis, we have used CountVectorizer for tokenization process which achieved better results.

5.4. Aspects Extraction

Aspects Extraction includes the distribution of cloud reviews concerning specific aspects or entities. The process is termed as Aspect-Based Sentiment Analysis. Within the proposed study, three aspects, i.e. Performance, Price and Service, are considered, and the cloud consumer reviews are sentimentally classified using the mentioned aspects. Some reviews may mention one aspect in its text, while more than one or all three aspects can also be found within the individual reviews. So, the number of reviews per aspect may vary. The division of reviews based on aspects shows a rational difference and are shown in the Table 5.5.

Table 5.5. Division of cloud consumer reviews w.r.t extracted aspects

Aspects	Positive		Neutral	Negative	
	5	4	3	2	1
Performance	169	56	1	2	7
Price	455	146	24	29	41
Service	1418	323	56	52	175

5.5. Polarity Scores Calculation

After the extraction of aspects from reviews, sentiment scores for the aspects are calculated using NLTK Library. The Vader Sentiment Analyzer is used to perform the sentiment score calculations of the aspects in each cloud consumer review. VADER calculates the polarity as well as the intensity of sentiments. It maps the words to those in the built-in dictionary and results in the sentiment score by adding the intensities of all words in the reviews that tend to elaborate on the performances of CSPs. Sentiment Intensity Analyzer classifiers for positive, negative and neutral texts act for the purpose. For reviews containing both positive and negative parts, the compound intensity is calculated which gives the resultant intensity for a CSP. The algorithm for sentiment classification in our proposed method is described in Algorithm 1.

Algorithm 1 Pseudocode used for Sentiment Classification of Cloud Reviews Dataset

```
Input:      Review= Cloud Consumer Text Reviews
Output:     Sentiment= sentiment polarity of text reviews
procedure Sentiment Classification(Cloud_Reviews_Dataset)
    for Each Reviews in Cloud_Reviews_Dataset do
        Reviews = Preprocessing(Reviews)
            Review_Text = CountVectorizer(Reviews)
            Tokenized_Review_Text = Tokenization(Review_Text)
            Tokenized_Review_Text = StopwordsRemove(Tokenized_Review_Text)
        Aspect_Text = ExtractAspects(Tokenized_Review_Text)
        Text_PolarityScore =VADER(Tokenized_Review_Text, Aspect_Text)
        Sentiment = Apply_RNN/LSTM/GRU-Classifiers(Tokenized_Review_Text,
    Aspect_Text)
        SentimentClassify = Apply_Check(ReviewRating, Sentiment)
        ClassificationScores = Apply_EvaluationMetrics(SentimentClassify)
    return ClassificationScores(Accuracy, Precision, F1-score, Recall)
    return List_of_Top_CSPs = AHP_Model (Text_PolarityScore)
```

Algorithm 1. Pseudocode for sentiment analysis of cloud reviews

The compound score is calculated by adding the valence values of each individual term in the corpus, adjusting them according to the criteria, and afterwards normalizing them to a range of -1 to +1. It is best method to employ if only one-dimensional sentiment calculation is the purpose. The formula for compound score is given as:

$$x = {}^{x}\!\big/\!\sqrt{x^2 + \alpha} \quad (5.1)$$

Here, x is the summation of valence values of constituent words while α is the Normalization constant. For instance, if we take a review to calculate the polarity score for Freshdesk CSP i.e.

"Good pricing, easy to use, a lot of features. Freshdesk is quite slow to load and the search bar works in a really weird way".

Here, it contains both negative and positive sentiments and sentiment scores. So a compound sentiment score is calculated as:

[Negative score: 0.239, Positive score: 0.307, Compound score: 0.1531]

Table 5.6 shows a list of Cloud Service Providers with the aspects identified in the reviews mentioning them. At the same time, a column for sentiment scores is also noted to determine the intensity of aspect in the view.

Table 5.6. Sentiment scores of different aspects in reviews

Cloud Service Providers	Aspect	Sentiment Scores
Fathom Voice	Service	0.317019
Grasshopper	Service	0.275561
Postscan Mail	Service	0.019945
Agorapulse	Service	0.541242
Surveygizmo	Service	0.523079
Samanage	Service	0.498245
Teamsupport.Com	Service	0.408167
Zendesk	Service	0.231123
Stitch Labs	Service	0.504187
Tradegecko	Service	0.364139
Wintac	Service	0.321626
Insightsquared	Performance	0.343113
Close.Io	Performance	0.146800
Samanage	Performance	0.471033
Namely	Performance	0.240600
Stackify	Performance	-0.020925
Brightpearl	Performance	-0.175650
Copperegg	Performance	0.142029
Incapsula	Performance	0.454700
Optmyzr	Performance	0.287000
Wethrive	Performance	0.056575
Atum	Performance	0.699725
Optmyzr	Price	0.145625
Instapage	Price	0.150171
Creately	Price	0.184814
Jobber	Price	0.123686
Hubstaff	Price	0.134614

Intuit Quickbase	Price	0.505271
Paymo	Price	0.094557
Intouch	Price	0.190633
Nimble	Price	0.048444

5.6. Hyper-parameter Configuration

The hyper-parameters tuning process takes place to achieve the desired outcomes from deep learning approaches of RNN, LSTM and GRU, setting parameters to their average values. For that purpose, Table 5.7 describes the configuration of various hyper-parameters of deep learning architectures such as RNN, LSTM and GRU used in sentiment classification.

Table 5.7. Hyper-parameter configuration of Deep Learning approaches

Hyper-parameters	**Values**
Model	Sequential
Activation Function	Softmax
Batch Size	32
Optimizer	Adam
Hidden Layer	3
Number of Neurons	100
Learning Rate	0.001
Epochs	10
Dropout Value	0.5

5.7. Evaluation Metrics

Deep Learning approaches classified the cloud reviews based on aspects extracted. Each aspect has its implementation results. The implementation of the framework has been evaluated using various classification metrics. Performance of each aspect is measured using Precision, Recall, F1-Score, Support and Accuracy. These metrics determine the overall achievement of processes to confirm their validity for the framework.

5.7.1. Precision

Precision is the measure of actual positive predictions out of all predicted positives. The True Positive (TP) ratio is the sum of positive predictions (*TP and FP*). The equation for Precision is given as:

$$Precision = {}^{TP}/_{(TP + FP)} \tag{5.2}$$

5.7.2. Recall

Recall, also known as Sensitivity, is the measure of actual positive predictions out of all the positives. It is the proportion of True Positive (*TP*) to the sum of all positives (*TP and FN*). The equation for Recall is given as:

$$Recall = {}^{TP}/_{(TP + FN)} \tag{5.3}$$

5.7.3. F1-Score

F1-Score is the Harmonic Mean (*HM*) of Precision and Recall. Its value varies between 0 and 1. The higher the values of Precision and Recall, the more F1-Score value moves near to 1 and vice versa. It maintains an equilibrium between both metrics for classifiers. The equation for F1-Score is given as:

$$F1 - Score = 2 * ({}^{Precision * \ Recall}/_{Precision + \ Recall}) \tag{5.4}$$

5.7.4. Accuracy

Accuracy is the degree of correctness of a framework with respect to its performance. It determines the number of accurate predictions out of the total number of predictions. Mathematically, it is represented by the sum of True Positive (*TP*) and True Negative divided to the sum of all predictions *(TP, TN, FN and FP*). The equation for Accuracy is given as:

$$Accuracy = {}^{(TP + TN)}/_{(TP + TN + FP + FN)} \tag{5.5}$$

5.7.5. Support

Support is the measure of actual responses of a particular class in a dataset. It indicates the true occurrences of samples in a class.

Tables 5.8, 5.9 and 5.10 highlight the actual and predicted values and Accuracy (Acc) according to the deep learning models of RNN, LSTM and GRU for the given aspects of Price, Performance and Service in the testing dataset. The cloud dataset is classified according to the aspects each time, and so the splitting of training and testing datasets. Each model shows its accuracy rates. Table 5.11 shows the values of classification metrics achieved by each aspect after implementation. However, a visual comparison of RNN, LSTM and GRU accuracy rates per epoch is shown in line graphs in Figure 5.3. RNN model achieves the highest accuracy in all cases.

Table 5.8. Actual and predicted values for Price aspect in test dataset

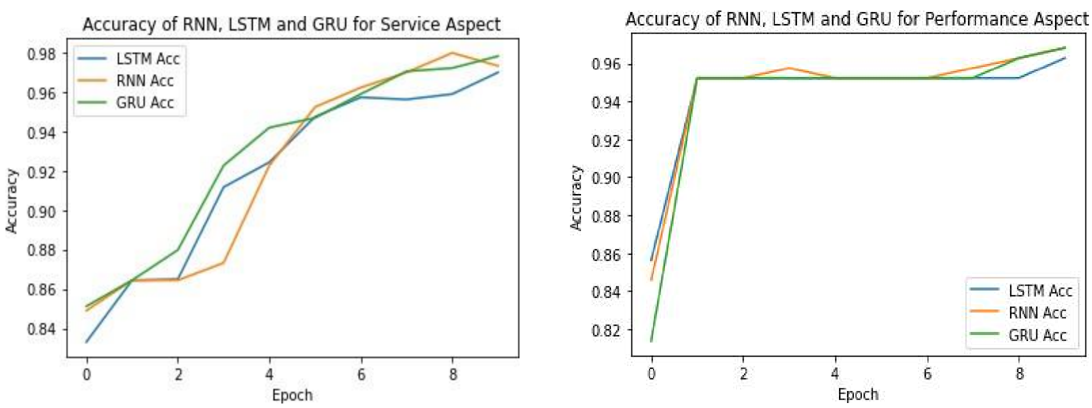
Model	Train	Test	Actual Positive	Predicted Positive	Actual Neutral	Predicted Neutral	Actual Negative	Predicted Negative	Acc
GRU	556	139	124	123	3	3	12	10	93%
RNN	556	139	124	124	3	3	12	10	**95%**
LSTM	556	139	124	121	3	3	12	9	89%

Table 5.9. Actual and predicted values for Performance aspect in test dataset

Model	Train	Test	Actual Positive	Predicted Positive	Actual Neutral	Predicted Neutral	Actual Negative	Predicted Negative	Acc
GRU	188	47	46	44	0	0	1	1	95%
RNN	188	47	46	45	0	0	1	1	**98%**
LSTM	188	47	46	43	0	0	1	0	91%

Table 5.10. Actual and predicted values for Service aspect in test dataset

Model	Train	Test	Actual Positive	Predicted Positive	Actual Neutral	Predicted Neutral	Actual Negative	Predicted Negative	Acc
GRU	1821	364	335	331	10	9	19	18	82%
RNN	1821	364	335	334	10	9	19	18	**91%**
LSTM	1821	364	335	329	10	8	19	17	80%



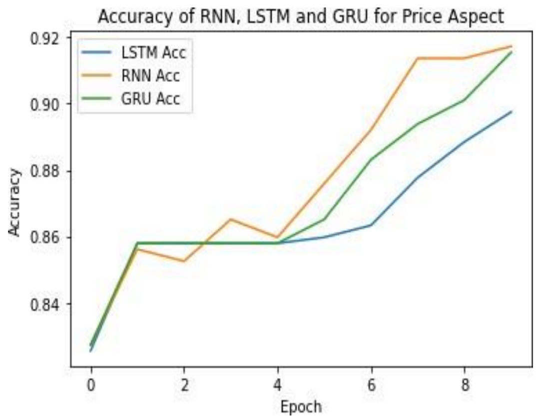


Figure 5.3. Comparison of accuracy rates per epoch for Performance, Price and Service attributes

Table 5.11. Classification metrics average values for each aspect

Aspects	Precision	Recall	F1-Score	Support	Accuracy
Performance	0.96	0.98	0.97	47	98.00%
Price	0.92	0.90	0.91	139	95.27%
Service	0.87	0.84	0.85	364	91.02%

It is evident from the above tables that among the three deep learning approaches, RNN proves to achieve the highest accuracy rate for three aspects. The results represent that aspect of Performance achieves the highest accuracy rate of 98%, followed by Price with 95.27% accuracy. In comparison, the Service aspect has 91.02% forming the least accuracy rate among all the aspects.

A decision-making model is necessary to decide which aspect of prioritising and calculating their actual criteria scores. A Multi-Criteria Decision Making (MCDM) method can help organizations make critical decisions based on the generated data. Analytic Hierarchical Process (AHP) is an MCDM method used in the proposed study to decide which aspect has the highest priority score that affects the decision-making of cloud service consumers in selecting a suitable cloud service provider. The next chapter explains AHP Model.

6. ANALYTIC HIERARCHY PROCESS (AHP) DECISION MAKING MODEL

The Cloud Service Consumers (CSCs) mention their experiences of requiring services from a Cloud Provider in reviews. The reviews contain various aspects indicating the overall performance of a provider [51]. The Deep Learning methods are utilized to extract the sentiment scores of mentioned aspects within the cloud consumer reviews. Through the calculated sentiment scores, the intensity of the aspects in a review can be highlighted. Each review gets different sentiment scores concerning either aspect, which indicates the evaluation of the cloud services offered by a particular cloud service provider. This information is quite helpful for other new cloud service consumers to know before consulting a specific cloud service provider for cloud services [56, 67]. For that purpose, a decision model is required based on which cloud service providers are prioritized and provides ease to cloud service consumers in selecting an appropriate service provider.

One of the most famous and commonly used multi-criteria procedures is the Analytic Hierarchy Process (AHP). The methods of evaluating alternatives and aggregating to discover the most appropriate choices are combined in this methodology. The Analytic Hierarchy Process (AHP) is a math and psychology-based system for organizing and evaluating complicated choices. It was created in the 1970s by Thomas L. Saaty [68] and has subsequently been enhanced. The approach is used to rank a collection of options or to choose the best from a group of choices. The selection is based on a broad objective that is broken down into a pool of criteria. The model consists of three components: the ultimate aim or problem you're attempting to address, all feasible solutions (referred to as alternatives), and the factors you'll use to evaluate the alternatives [69]. By defining its parameters and alternative possibilities and tying those parts to the broader purpose, AHP gives a coherent foundation for the desired conclusion.

The application of the AHP Decision model within a particular situation is carried out in three respective steps. First, a hierarchical framework model is designed for the problem where the goal is mentioned at the top of the hierarchy, followed by different criteria grouped in the following levels. Secondly, Pair-Wise Comparisons (PWC) based on subjective preferences are collected to calculate relative priorities in the form of local numerical weights. Lastly, the combination of relative priorities to get global priorities as global weights reaches a final destination. To consider the above AHP process for the proposed framework, an ultimate goal as Cloud Provider Selection is identified in Level 1 of the hierarchy. Level 2 contains the identified aspects in the reviews, i.e. Performance, Price and Services that serve as the criteria factors, while the Cloud Service Providers (CSPs) in the last level is the conclusion after various complex calculations. Figure 6.1 illustrates an AHP hierarchical model for the proposed framework.

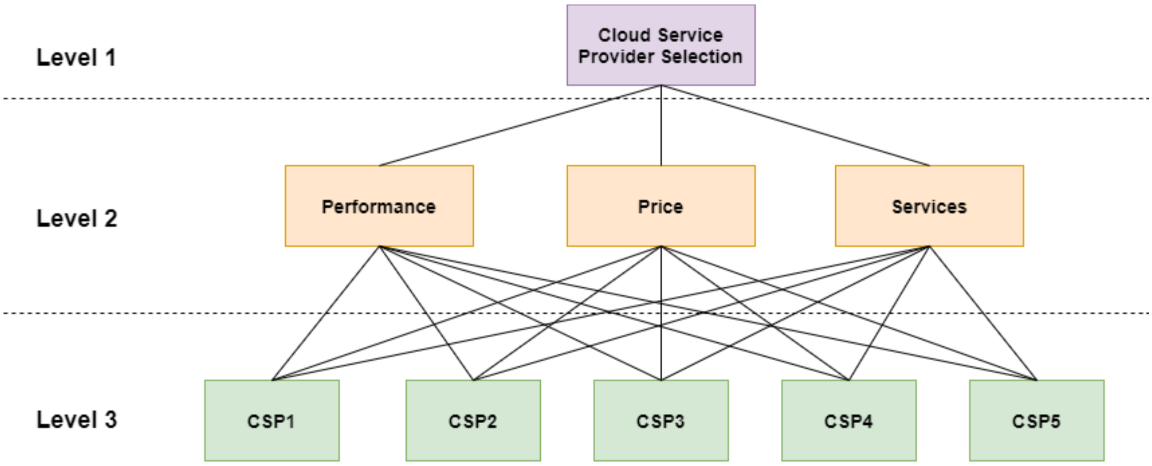


Figure 6.1. AHP Hierarchical model

6.1. The Comparison Scale (SAATY Scale)

The Comparison Scale, also known as Scale of Relative Importance or a SAATY Scale [68], compares pairs of considered parameters in the form of a matrix. The preferences of decision-makers are then listed using SAATY Scale as a standard. The scale has five levels and four sub-levels that describe the intensity of importance and have integer values ranging from 1 to 9. Using a pair-wise comparison mechanism of the hierarchy elements, a ranking of all criteria is accomplished. These pair-wise comparisons are used to transform verbal judgments into numeric data ranging from 1 to 9, with one indicating that two parameters will be of equal importance and nine meaning that one parameter is more important than the other parameter. Table 6.1 shows the SAATY Scale intensity levels with their definition.

Table 6.1. Levels of SAATY Scale [68]

Intensity of Importance	Definition
1	Equal Importance
2	Intermediate Importance between 1 and 3
3	Moderate Importance
4	Intermediate Importance between 3 and 5
5	Strong Importance
6	Intermediate Importance between 5 and 7
7	Very Strong Importance
8	Intermediate Importance between 7 and 9
9	Extremely Strong Importance
1/3, 1/5, 1/7, 1/9	Values of Inverse Comparison

6.2. Pair-Wise Comparison Matrix

A Pair-Wise Comparison Matrix establishes the relative importance of several attributes/criteria about the goal. In the perspective of this thesis, the aspects identified during sentiment classification serve as attributes/ criteria for performing pair-wise comparisons that relate to the purpose of Cloud Provider Selection. Here, we have three attributes for the proposed study, i.e. Performance, Price and Service, so that the Pair-Wise Calculation Matrix will be 3x3 dimensions.

Let us consider [67, 70] a Cloud Service Consumer (CSC) prefers the Performance to be the utmost important factor in a CSP to fulfil its requirements followed by the Price attribute while the Service attribute has the least importance for the particular CSC. It is clear from Table 5.2 where a population of 96.15% preferred Performance to be the most important attribute. On the other hand, 89.56% chose Price to be important while Service attribute seems important to 88.45% of CSCs. The statistics are calculated by only considering the quantities for positive and negative reviews. We achieve the percentage of positive reviews by dividing them by the sum of positive and negative reviews. These statistics are listed in Table 6.2. However, Table 6.3 calculates the values of paired attributes for the given scenario in a matrix form.

Table 6.2. Percentage of positive reviews for each aspect

Attributes	No. of Positive	No. of Negative	Percentage of Positive Reviews
Performance	225	9	0.9615
Service	601	70	0.8956
Price	1741	227	0.8846

Table 6.3. Pair-Wise Comparison 3x3 matrix of cloud attributes

Attributes	Performance	Price	Service
Performance	1	3	4
Price	0.33	1	2
Service	0.25	0.5	1
Net Sum	**1.58**	**4.5**	**7**

The attributes above are evaluated and prioritized based on their importance to CSCs in the pair-wise comparison matrix. To assume, if ‘Price’ is given an ‘x’ value, then ‘Performance’ will be given a ‘3x’ value since Performance is of moderate importance than Price. Similarly, if

‘Service’ is given an ‘x’ value, ‘Performance’ will be given a ‘4x’ value since Performance is of intermediate importance between 3 and 5 than Service. Similarly, comparing the Price and Service attributes, if ‘Service’ is given an ‘x’ value, ‘Price’ will be given a ‘2x’ value since Price is of intermediate importance between 1 and 3 than Service. Their inverse values will be applicable in the matrix for inverse relations of aspects. At the same time, the diagonal elements will be one since the same paired attributes will have equal importance.

6.3. Normalized Pair-Wise Comparison Matrix

Normalized Pair-Wise Comparison Matrix is calculated by dividing all the column elements by the sum of the column in a matrix, as shown in Table 6.4. The average of row elements is calculated in each row and is mentioned as values of Criteria Weights of attributes. Criteria Weights are the importance weights of attributes. The higher the criteria weight for an attribute, the more it will be important for cloud service consumers.

Table 6.4. Normalized Pair-Wise Comparison 3x3 matrix of cloud attributes

Attributes	Performance	Price	Service	Criteria Weights
Performance	0.632	0.666	0.571	**0.623**
Price	0.208	0.222	0.285	**0.238**
Service	0.158	0.111	0.142	**0.137**

6.4. Consistency Calculation

After calculating the normalized pair-wise matrix calculation, a consistency calculation step needs to be implemented. It is performed to check the accuracy of calculated values in a normalized pair-wise matrix. For this purpose, the Pair-Wise Calculation Matrix is considered and takes each column element's product with the criteria weights mentioned at the top of columns. Table 6.5 calculates the products of each column with criteria weights.

Table 6.5. Product of Pair-Wise Comparison matrix with Criteria Weights

Criteria Weights	**0.623**	**0.238**	**0.137**
Attributes	**Performance**	**Price**	**Service**
Performance	0.623	0.714	0.548
Price	0.205	0.238	0.274
Service	0.155	0.119	0.137

A Weighted Sum Value column is created in the matrix, which takes all the elements in a row. Table 6.6 describes the sum of updated values in Pair-Wise Calculation Matrix and sums it up into the Weighted Sum Value column. It calculates the addition of all row values for future calculations.

Table 6.6. Calculation of Weighted Sum Values

Attributes	Performance	Price	Service	Weighted Sum Value
Performance	0.623	0.714	0.548	**1.885**
Price	0.205	0.238	0.274	**0.717**
Service	0.155	0.119	0.137	**0.411**

Ratio Values are obtained by taking the ratio of Weighted Sum Values by their Criteria Weights for each attribute. The Ratio Values are denoted by (RV) and are then used for λ_{max} calculation. Table 6.7 calculates the Ratio Values (RV) for each category.

Table 6.7. Calculation of Ratio Values per attribute

Attributes	Performance	Price	Service	Weighted Sum Value	Criteria Weights	Ratio Values
Performance	0.623	0.714	0.548	1.885	0.623	**3.025**
Price	0.205	0.238	0.274	0.717	0.238	**3.012**
Service	0.155	0.119	0.137	0.411	0.137	**3.000**

By calculating ratio values, their average is taken by adding all the values and dividing them by the total number of attributes. It is termed as Principal Eigenvalue and is denoted by λ_{max}. It derives the priority vectors for a given matrix and indicates the order of preference among them. Thus, λ_{max} for the above values is shown in Equation 6.1 as:

$$\lambda_{max} = (RV_1 + RV_2 + RV_3)/3 \tag{6.1}$$

To determine whether the matrix is consistent or contradictory, the Consistency Index (CI) is calculated, which is given in Equation 6.2 as follows:

$$CI = {(\lambda_{max} - a)}/{(a-1)} \tag{6.2}$$

Here, 'a' represents the number of compared attributes. In the proposed study, the value of λ_{max} is calculated to be 3.012, the number of attributes 'a' is 3, so the Consistency Index is calculated to be 0.0061. To find out the value of Consistency Ratio (CR), the ratio of Consistency

Index (*CI*) and Random Index (*RI*) give the results. It determines the level of consistency of calculated values relative to the consistent values of large samples. Its formula is mentioned in Equation 6.3 as

$$CR = {}^{CI}/_{RI} \tag{6.3}$$

Random Index is the Consistency Index of the random generated pair-wise matrix. Table 6.8 shows the standard random indexes for 10 attributes. Since there are three attributes in the proposed study, therefore, we will consider a=3.

Table 6.8. Random Consistency Index

a	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

According to the Random Consistency Index table, the value of Random Index at a=3 is 0.58. So, the Consistency Ratio achieved will be 0.0106. Comparing this value with the standard *CI* value, i.e. 0.10, the proportion of inconsistency *(CR)* is less than the Standard Consistency Ratio. In other words, 0.0106 < 0.10.

Thus, it becomes valid that the proposed matrix is considered reasonably consistent, and AHP may be used as a decision-making model for the proposed thesis. Keeping in view the requirements of a CSC, the criteria weights calculated for the attributes can be used for decision-making purposes. Exploring Table 6.9, the 'Performance' attribute achieves the highest weightage followed by the 'Service' attribute while 'Price' has the least priority among all the categories.

Table 6.9. Priority Weights and percentages of attributes

Attributes	Criteria Weights	Percentage
Performance	0.623	62.3%
Price	0.238	23.8%
Service	0.137	13.7%

Based on the above criteria weights, different cloud service providers are prioritized and listed based on their performance, customer service and cost of the offered services. For a particular scenario, Performance achieved the maximum importance of 62.3% for specific CSC. In order to select the best CSP for specific aspects, we have considered three CSPs i.e. for instance, Wintac, Atum and Optmyzr. Performing the same above procedures to the CSPs w.r.t aspects, Tables 6.10, 6.11 and 6.12 represent the Pair-wise comparison matrixes. Keeping in view, Wintac is better than Atum and Optmyzer in terms of Service, Atum is best in Performance as compared to Wintac and Optmyzer. While Optmyzer has best prices than Wintac and Atum.

Table 6.10. Pair-Wise Comparison 3x3 matrix of CSPs

	Performance		
CSPs	**Wintac**	**Atum**	**Optmyzr**
Wintac	1	0.2	1
Atum	5	1	5
Optmyzr	1	0.2	1
Net Sum	**7**	**1.4**	**7**

Table 6.11. Pair-Wise Comparison 3x3 matrix of CSPs

	Price		
CSPs	**Wintac**	**Atum**	**Optmyzr**
Wintac	1	1	0.2
Atum	1	1	0.2
Optmyzr	5	5	1
Net Sum	**7**	**7**	**1.4**

Table 6.12. Pair-Wise Comparison 3x3 matrix of CSPs

	Service		
CSPs	**Wintac**	**Atum**	**Optmyzr**
Wintac	1	5	5
Atum	0.2	1	1
Optmyzr	0.2	1	1
Net Sum	**1.4**	**7**	**7**

Now, finding the Normalised Pair-wise Comparison matrix for the above three tables, Tables 6.13, 6.14 and 6.15 calculate Normalised Pair-wise Comparison matrixes with respect to each aspect i.e. Performance, Service and Price.

Table 6.13. Normalised Pair-Wise Comparison 3x3 matrix of CSPs

Performance				
CSPs	**Wintac**	**Atum**	**Optmyzr**	**Criteria Weights**
Wintac	0.142	0.142	0.142	**0.142**
Atum	0.714	0.714	0.714	**0.714**
Optmyzr	0.142	0.142	0.142	**0.142**

Table 6.14. Normalised Pair-Wise Comparison 3x3 matrix of CSPs

Price				
CSPs	**Wintac**	**Atum**	**Optmyzr**	**Criteria Weights**
Wintac	0.142	0.142	0.142	**0.142**
Atum	0.142	0.142	0.142	**0.142**
Optmyzr	0.714	0.714	0.714	**0.714**

Table 6.15. Normalised Pair-Wise Comparison 3x3 matrix of CSPs

Service				
CSPs	**Wintac**	**Atum**	**Optmyzr**	**Criteria Weights**
Wintac	0.714	0.714	0.714	**0.714**
Atum	0.142	0.142	0.142	**0.142**
Optmyzr	0.142	0.142	0.142	**0.142**

Thus, the Criteria Weight Table for the CSPs is formed in Table 6.16 as follows:

Table 6.16. Criteria Weights of CSPs w.r.t attributes

CSPs	Performance	Price	Service
Wintac	0.142	0.142	0.714
Atum	0.714	0.142	0.142
Optmyzr	0.142	0.714	0.142

The last step includes multiplying Table 6.9 and 6.16 to get the list of CSPs suitable for aspects. Rows of Table 6.16 are multiplied by the columns of Table 6.9 and the results are shown in Table 6.17. It displays the considered Cloud Service Providers i.e. Wintac, Atum and Optmyzer on the basis of their importance using all the attributes of Performance, Service and Price.

Table 6.17. Priority Weights of CSPs and percentages of importance

CSPs	Criteria Weights	Percentage
Wintac	0.218	21.8%
Atum	0.496	49.6%
Optmyzr	0.276	27.6%

It is evident from Table 6.17 that Atum CSP is the most important cloud service attribute with 49.6% priority weight followed by Optmyzer having 27.65 importance. While Wintac CSP is the least important CSP with 21.8% priority weight.

7. CONCLUSIONS

Sentiment Analysis plays a critical role in different Natural Language Processing (NLP) tasks. Aspect-Based Sentiment Analysis assesses the user's experiences, perspectives, and opinions, highlighting bold attributes in various social media platform reviews or newspapers and articles. According to the studies, ABSA is applied to various applications and a required field to explore and study. It is required by commercial organizations to analyze how their goods are regarded and by future customers for making decisions.

The thesis implements state-of-the-art Deep Learning approaches such as RNN, LSTM and GRU, identifying important attributes such as Performance, Price and Service within the 6259 Cloud Service Consumers' reviews extracted using the HaaS framework to perform Aspect-Based Sentiment Analysis. The deep learning approaches achieved an average of 89.66% accuracy rate for analyzing the aspects. The results revealed a list of CSPs reviewing their performances, services and cost for the offered services. Furthermore, an AHP Decision Model prioritized the extracted aspects according to the requirements of the CSCs stated in a case study. The study's calculations showed that the CSC emphasized achieving maximum Performance for CSPs, followed by the Price. However, the Service factor had minimum importance among the selected aspects.

REFERENCES

[1] Zhu, L., et al. (2020). *SentiVec: Learning Sentiment-Context Vector via Kernel Optimization Function for Sentiment Analysis.* 2020.

[2] Kale, M., P. Mankame, and G.J.a.p.a. Kulkarni, (2018). *Deep learning for digital text analytics: Sentiment analysis.* 2018.

[3] Abbas, S.S.I., (2019). *Automatic Text Categorization of Turkish News with Machine Learning and Deep Learning Techniques*. 2019, Ankara Yıldırım Beyazıt Üniversitesi Fen Bilimleri Enstitüsü.

[4] Sohangir, S., et al. (2018). *Big Data: Deep Learning for financial sentiment analysis.* 2018. **5**(1): p. 3.

[5] Naderalvojoud, B., B. Qasemizadeh, and L.J.P.o.t.G. Kallmeyer, (2017). *HU-HHU at GermEval-2017 Sub-task B: Lexicon-Based Deep Learning for Contextual Sentiment Analysis.* 2017: p. 18-21.

[6] Kratzwald, B., et al. (2018). *Deep learning for affective computing: Text-based emotion recognition in decision support.* 2018. **115**: p. 24-35.

[7] Ghorbani, M., et al. (2020). *ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing.* 2020. **9**(1): p. 1-12.

[8] Pathak, A.R., et al., *Empirical evaluation of deep learning models for sentiment analysis.* 2019. **22**(4): p. 741-752.

[9] Çano, E. and M. Morisio. (2018). *A deep learning architecture for sentiment analysis*. in *Proceedings of the International Conference on Geoinformatics and Data Analysis*. 2018.

[10] Markoski, Filip. (2020). Sentiment Analysis of Twitter Data with Graph Structures and Deep Learning. 10.13140/RG.2.2.14678.88644.

[11] Hassan, A. and A. Mahmood. (2017). *Deep learning approach for sentiment analysis of short texts*. in *2017 3rd international conference on control, automation and robotics (ICCAR)*. 2017. IEEE.

[12] Li, L., et al. (2020). *How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis.* 2020. **32**(9): p. 4387-4415.

[13] AlSurayyi, W.I., N.S. Alghamdi, and A. Abraham, *Deep Learning with Word Embedding Modeling for a Sentiment Analysis of Online Reviews.*

[14] Mahendhiran, P., S.J.I.J.o.I.T. Kannimuthu, and D. Making, (2018). *Deep learning techniques for polarity classification in multimodal sentiment analysis.* 2018. **17**(03): p. 883-910.

[15] Kaman, S.J.S.P., *News Sentiment Analysis By Using Deep Learning Framework.* 2020.

[16] Jangid, H., et al. (2018). *Aspect-based financial sentiment analysis using deep learning*. in *Companion Proceedings of the The Web Conference 2018*. 2018.

[17] Rani, S., P.J.A.J.f.S. Kumar, and Engineering, *Deep learning based sentiment analysis using convolution neural network.* 2019. **44**(4): p. 3305-3314.

[18] Rehman, A.U., et al. (2019). *A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis.* 2019. **78**(18): p. 26597-26613.

[19] Othan, D., Z.H. Kilimci, and M. Uysal, *Financial Sentiment Analysis for Predicting Direction of Stocks using Bidirectional Encoder Representations from Transformers (BERT) and Deep Learning Models.*

[20] Mittal, N., D. Sharma, and M.L. Joshi. (2018). *Image Sentiment Analysis Using Deep Learning*. in *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*. 2018. IEEE.

[21] Yadav, A. and D.K.J.M.S. Vishwakarma, (2020). *A deep learning architecture of RA-DLNet for visual sentiment analysis.* 2020.

[22] Kalarani, P. and S.S.J.S.C. Brunda, (2019). *Sentiment analysis by POS and joint sentiment topic features using SVM and ANN.* 2019. **23**(16): p. 7067-7079.

[23] Vassilev, A. (2019) *BowTie-A deep learning feedforward neural network for sentiment analysis*. in *International Conference on Machine Learning, Optimization, and Data Science*. 2019. Springer.

[24] Shrestha, N. and F.J.a.p.a. Nasoz, (2019) *Deep Learning Sentiment Analysis of Amazon. Com Reviews and Ratings.* 2019.

[25] Chakraborty, K., et al. (2020) *Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media.* 2020. **97**: p. 106754.

[26] Chaturvedi, I., et al. (2019). *Fuzzy commonsense reasoning for multimodal sentiment analysis.* 2019. **125**: p. 264-270.

[27] Do, H.H., et al. (2019). *Deep learning for aspect-based sentiment analysis: a comparative review.* 2019. **118**: p. 272-299.
[28] Nassif, A.B., et al. (2020). *Deep learning for Arabic subjective sentiment analysis: Challenges and research opportunities.* 2020: p. 106836.
[29] Ouerhani, N., et al. (2020). *Smart Ubiquitous Chatbot for COVID-19 Assistance with Deep learning Sentiment Analysis Model during and after quarantine.* 2020.
[30] Hossain, E., et al., *SentiLSTM: A Deep Learning Approach for Sentiment Analysis of Restaurant Reviews.* 2020.
[31] Wang, Y., et al. (2016). *Attention-based LSTM for aspect-level sentiment classification*. in *Proceedings of the 2016 conference on empirical methods in natural language processing*. 2016.
[32] Rysbek, D. (2019). *Sentiment analysis with recurrent neural networks on turkish reviews domain.* 2019.
[33] Rao, G., et al. (2018). *LSTM with sentence representations for document-level sentiment classification.* 2018. **308**: p. 49-57.
[34] Cunanan, K. (2018). *Developing a Recurrent Neural Network with High Accuracy for Binary Sentiment Analysis.* 2018.
[35] Dhande, L.L., G.J.I.J.o.S.E. Patnaik, and T. Research, (2014). *Review of sentiment analysis using naive bayes and neural network classifier.* 2014. **3**(7): p. 1110-1113.
[36] Murthy, D., et al., *Text based sentiment analysis using LSTM.* 2020. **9**(05).
[37] Souma, W., I. Vodenska, and H.J.J.o.C.S.S. Aoyama, (2019). *Enhanced news sentiment analysis using deep learning methods.* 2019. **2**(1): p. 33-46.
[38] Deepa, M.D.J.T.J.o.C. and M. Education, *Bidirectional Encoder Representations from Transformers (BERT) Language Model for Sentiment Analysis task.* 2021. **12**(7): p. 1708-1721.
[39] Alaparthi, S. and M.J.a.p.a. Mishra, (2020). *Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey.* 2020.
[40] Yadav, A. and D.K.J.A.I.R. Vishwakarma, (2020). *Sentiment analysis using deep learning architectures: a review.* 2020. **53**(6): p. 4335-4385.
[41] Qaisi, L.M. and I. Aljarah. (2016). *A twitter sentiment analysis for cloud providers: a case study of Azure vs. AWS*. in *2016 7th International Conference on Computer Science and Information Technology (CSIT)*. 2016. IEEE.
[42] Agüero-Torales, M., et al. (2019). *A cloud-based tool for sentiment analysis in reviews about restaurants on TripAdvisor.* 2019. **162**: p. 392-399.
[43] Singh, H. and R.J.I.J.A.E.R. Randhawa, (2015). *CPSEL: Cloud provider selection framework for ranking and selection of cloud provider.* 2015. **10**(7): p. 18787-18810.
[44] Keijsers, M., C. Bartneck, and H.S. Kazmi. (2019). *Cloud-Based Sentiment Analysis for Interactive Agents*. in *Proceedings of the 7th International Conference on Human-Agent Interaction*. 2019.
[45] Karamitsos, I., et al., *Tweet Sentiment Analysis (TSA) for Cloud Providers Using Classification Algorithms and Latent Semantic Analysis.* 2019. **7**(4): p. 276-294.
[46] Sankar, H., et al. (2020). *Intelligent sentiment analysis approach using edge computing-based deep learning technique.* 2020. **50**(5): p. 645-657.
[47] Hussain, W., et al. (2018). *Comparing time series with machine learning-based prediction approaches for violation management in cloud SLAs.* Future Generation Computer Systems, 2018. **89**: p. 464-477.
[48] Hussain, W., F. Hussain, and O. Hussain. (2016). *QoS prediction methods to avoid SLA violation in post-interaction time phase*. in *2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA)*. 2016. IEEE.
[49] Alkalbani, A.M., W. Hussain, and J.Y. Kim, (2019). *A centralised cloud services repository (CCSR) framework for optimal cloud service advertisement discovery from heterogenous web portals.* IEEE Access, 2019. **7**: p. 128213-128223.
[50] Hussain, W., F.K. Hussain, and O.K. Hussain. (2015). *Comparative analysis of consumer profile-based methods to predict SLA violation*. in *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. 2015. IEEE.
[51] Hussain, W., F.K. Hussain, and O.K. Hussain. (2014). *Maintaining trust in cloud computing through SLA monitoring*. in *International Conference on Neural Information Processing*. 2014. Springer.
[52] Hussain, W., F.K. Hussain, and O.K. Hussain. (2016). *Risk management framework to avoid SLA violation in cloud from a provider's perspective*. in *International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*. 2016. Springer.
[53] Alghamdi, A., et al. (2017). *The need of an optimal QoS repository and assessment framework in forming a trusted relationship in cloud: A systematic review.* in *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)*. 2017. IEEE.

[54] Hussain, W., et al. (2015). *Profile-based viable service level agreement (SLA) violation prediction model in the cloud*. in *2015 10th international conference on P2P, parallel, grid, cloud and internet computing (3PGCIC)*. 2015. IEEE.

[55] Hussain, W., et al., (2020). *Cloud marginal resource allocation: A decision support model.* Mobile Networks Applications, 2020. **25**(4): p. 1418-1433.

[56] Hussain, W., F.K. Hussain, and O.K. Hussain. (2016). *SLA management framework to avoid violation in cloud*. in *International Conference on Neural Information Processing*. 2016. Springer.

[57] Alkalbani, A.M. and W. Hussain, (2021). *Cloud service discovery method: A framework for automatic derivation of cloud marketplace and cloud intelligence to assist consumers in finding cloud services.* International Journal of Communication Systems, 2021. **34**(8): p. e4780.

[58] Hussain, W., et al. (2017). *Formulating and managing viable SLAs in cloud computing from a small to medium service provider's viewpoint: A state-of-the-art review.* Information Systems, 2017. **71**: p. 240-259.

[59] Hussain, W., et al. (2016). *Provider-based optimized personalized viable SLA (OPV-SLA) framework to prevent SLA violation*. 2016, British Computer Society.

[60] Alrashed, B.A. and W. Hussain. (2020). *Managing SLA Violation in the cloud using Fuzzy re-SchdNeg Decision Model*. in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. 2020. IEEE.

[61] Hussein, D.M.E.-D.M.J.I.S.D.F.o.C. and E. Information Cairo University, *Analyzing scientific papers based on sentiment analysis.* 2016.

[62] Bravo-Marquez, F. (2017). *Acquiring and exploiting lexical knowledge for twitter sentiment analysis*. 2017, University of Waikato.

[63] Ahmed, S., et al., *The World Of Mobile Cloud Computing (A Survey To The Concept, Challenges and Solutions).* International Conference on Industrial Engineering and Management Applications.

[64] Raza, M.R., et al. (2021). *Sentiment Analysis using Deep Learning in Cloud*. in *2021 9th International Symposium on Digital Forensics and Security (ISDFS)*. 2021. IEEE.

[65] Alkalbani, A.M.K.R. (2018). *A methodology for automatic derivation of cloud marketplace and cloud intelligence*. 2018.

[66] Hussain, W., et al. (2021) *Integrated AHP-IOWA, POWA Framework for Ideal Cloud Provider Selection and Optimum Resource Management.* IEEE Transactions on Services Computing, 2021(01): p. 1-1.

[67] Saaty, T.L. (1977). *A scaling method for priorities in hierarchical structures.* Journal of mathematical psychology, 1977. **15**(3): p. 234-281.

[68] Gompf, K., M. Traverso, and J. Hetterich, (2021). *Using Analytical Hierarchy Process (AHP) to Introduce Weights to Social Life Cycle Assessment of Mobility Services.* Sustainability, 2021. **13**(3): p. 1258.

[69] Tsagdis, A., *The use of the Analytical Hierarchy Process as a source selection methodology and its potential application within the Hellenic Air Force*. 2008, NAVAL POSTGRADUATE SCHOOL MONTEREYCA.

CURRICULUM VITAE

Muhammad Raheel RAZA

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

ACADEMIC ACTIVITIES

Proceedings:

1. Muhammad Raheel Raza, Walayat Hussain, Erkan Tanyıldızı, and Asaf Varol. "Sentiment Analysis using Deep Learning in Cloud." In *2021 9th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1-5. IEEE, 2021.