REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

INVESTIGATION OF DATA MINING FOR MARKETING PURPOSES

Abdulsalam Husham Shukur Mahmood AGHA

Master's Thesis

Supervisor

Asst. Prof. Sefer KURNAZ

Istanbul, 2022

INVESTIGATION OF DATA MINING FOR MARKETING PURPOSES

Abdulsalam Husham Shukur Mahmood AGHA

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titied INVESTIGATION OF DATA MINING FOR MARKETING PURPOSES prepared by ABDULSALAM HUSHAM SHUKUR MAHMOOD and submitted on 12/12/2022 has been **accepted unanimously** for the degree Master of Science in Electrical and Computer Engineering

Asst. Prof. Dr. Sefer KURNAZ

Supervisor

Thesis Defense Committee Members:

Asst.Prof.Dr.Sefer KURNAZ	Department of Computer Engineering, Altınbaş University	_____
Asst.Prof.Dr.Oğuz KARAN	Department of Software Engineering, Altınbaş University	_____
Assoc.Prof.Dr.Adil Deniz DURU	Department of Trainer Education Programmer, Marmara University	_____

I hereby declare that this thesis meets all format and submission requirements of a Master`s Thesis.

Submission date of the thesis to the Institute of Graduate Studies: ____/____/____

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Abdulsalam Husham Mahmood AGHA

Signature

DEDICATION

The gentle soul who taught me to trust Almighty Allah and the steadiness and tackle the task with perseverance and enthusiasm.

My beloved mother

For making our life more comfortable and beautiful and for encouraging and supporting me along with all my life.

My amazing father

Who shared their words of advice encouragement to finish this study. Without my family's love and support, this thesis would not have been made possible. They were a great source of support for me.

My **Family**

PREFACE

Firstly, I thank Allah for providing me the patience and strength to achieve this work and making me believe that anything is possible with hard work and determination.

I would like to give a huge thank to my supervisor Asst. Prof. Dr. Sefer Kurnaz for serving on a graduate committee and for all of the helpful advice and directions he gave me.

I am extremely grateful to my parents for always encouraging me and pushing me to be better. I would also like to thank my brothers and sisters for their love and support for me.

ABSTRACT

INVESTIGATION OF DATA MINING FOR MARKETING PURPOSES

Mahmood Agha, Abdulsalam

M.Sc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Prof. Dr. Sefer Kurnaz

Date: February /2023

Pages: 46

To make sure that a product reaches the type of costumers that are more likely to buy it, in the last decades online Ad-based marketing started to bring more costumers for a certain product. However, in this work by using some information about a bank's costumers including their age and expected salary we were able to predict whether they will respond positively or negatively to our product's Ad by Firstly using Lineplot and Lmplot visualizations in python. And to prove that for more complexed datasets data visualization is not enough to predict the costumer's behavior, therefore, powerful machine learning models have been applied on the same dataset to make more accurate predictions, three deferent classification models have been used: Naïve-Bayes, KNN and SVM, these models showed accuracies of 88%, 90%, and 88.75%, respectively. Finally, predictions of new data that do not exist in the data set were predicted by the same models.

Keywords: Data Mining, Machine Learning, Naïve-Bayes, Support Vector Machine, K-Nearest Neighbor, Python, Data Visualization.

TABLE OF CONTENTS

Pages

ABSTRACT ..... **vii**

LIST OF FIGURES ..... **xi**

ABBREVIATIONS ..... **xii**

1. DATA MINING **1**

1.1 DATA MINING LAWS 1

1.1.1 First Law: Business Goals Law 1

1.1.2 Second Law: Business Knowledge Law 1

1.1.3 Third Law: Data Preprocessing Law 2

1.1.4 Fourth Law: Right Model Law 2

1.1.5 Fifth Law: Amplification Law 2

1.1.6 Sixth Law: Pattern Law 2

1.1.7 Seventh Law: Prediction Law 3

1.1.8 Eighth Law: Value Law 3

1.1.9 Nineth Law: Change Law 3

1.2 MACHINE LEARNING 3

1.2.1 Regression 4

1.2.2 Classification 5

1.2.3 Clustering 5

1.3 MARKETING DATA 5

1.4 DATA-DRIVEN MARKETING 6

1.5 COSTUMERS' BEHAVIOUR 6

1.5.1 Setting up a Marketing Goal 6

1.5.2 Use the Right Dataset 7

1.5.3 Build a Simple Model 7
1.5.4 Testing the Prediction Model 8
1.5.5 Set up the Model into the Market 8
2. LITERATURE REVIEW 9
2.1 DATA MINING FOR MARKETING 9
2.2 STUDY OF BUSINESSES AFFECTION BY PANDEMICS AND NATURAL DISASTERS 11
2.3 MACHINE LEARNING BASED MARKETING 11
2.3.1 Revolutionizing Marketing by Machine Learning 13
2.4 DATA GRAPHING IN MARKETING 13
3. DATASET, EXPIREMENTS AND RESULTS 16
3.1 MACHINE LEARNING IN PYTHON 16
3.2 DATASET 16
3.3 DATASET VISUALIZING IN PYTHON 19
3.4 DATA PREPROCESSING OF ADS DATASET 22
3.5 MACHINE LEARNING ON ADS DATASET 23
3.5.1 Naïve Bayes 23
3.5.2 K-Nearest Neighbor (K-NN) 25
3.5.3 Support Vector Machine (SVM) 27
3.6 ROC AND AUC SCORE 29
4. CONCLUSIONS AND RECOMMENDATIONS 31
4.1 CONCLUSION 31
4.2 RECOMMENDATIONS FOR FUTURE WORK 33
REFERENCES 34
APPENDIX A 38

LIST OF TABLES

Page

Table 2.1: Business-related data mining researches 10

Table 3.1: First 10 rows of the dataset. 17

Table 3.2: The output of info function 18

Table 3.3: The output of describe function. 18

Table 3.4: The libraries needed for ad dataset and their corresponding abbreviations 22

Table 3.5: Confusion matrix of naïve-bayes classifier. 24

Table 3.6: Confusion matrix of k-nn classifier. 26

Table 3.7: Confusion matrix of naïve-bayes classifier. 27

Table 4.1: ML models and their corresponding accuracies. 31

Table 4.2: FP, FN, TP and TN of ML models. 32

Table 4.3: AUC score of ml models. 32

LIST OF FIGURES

Pages

Figure 1.1: Machine learning tree. 4

Figure 2.1: The relationship between Data Engineering, Data Science and Data-Driven Decision Making. 9

Figure 3.1: line plot of costumers who responded negatively to the Ad. 19

Figure 3.3: Line plot of costumers and their respond to the Ad (Blue = 1, Orange = 0) 20

Figure 3.4: Lmplot of costumers who responded positively to the Ad. 21

Figure 3.5: Graph of Naïve Bayes classifier for training set. 24

Figure 3.6: Graph of Naïve Bayes classifier for test set. 25

Figure 3.9: Graph of SVM classifier for training set. 28

Figure 3.10: Graph of SVM classifier for test set. 28

Figure 3.11: ROC Curve of Naïve bayes classifier. 29

Figure 3.12: ROC Curve of K-NN classifier. 30

Figure 3.13: ROC Curve of SVM classifier. 30

ABBREVIATIONS

Ad	:	Advertisement
AI	:	Artificial Intelligence
AUC	:	Area Under Curve
FN	:	False Negative
FP	:	False Positive
K-NN	:	K-Nearest Neighbor
ML	:	Machine Learning
NB	:	Naïve Bayes
ROC	:	Receiver Operating Characteristic Curve
SVM	:	Support Vector Machine
TN	:	True Negative
TP	:	True Positive

1. DATA MINING

Data Mining (DM) is the exploration of patterns and examples in huge and complex informational collections or what called (Datasets) in terms of looking for irregularities or little neighborhood structures in information, with the tremendous mass of the information being irrelevant. Surely, one perspective for some huge scope information mining exercises is that they basically establish filtering and data reduction processes. The whole idea of DM is making sense of huge data blocks that do not make sense to the observer. Usually, the person who is responsible of working with datasets is called data scientist or data engineer.

1.1 DATA MINING LAWS

Each field has its own core values, thoughts that give design and direction in regular work. Data mining is no special case. Following are nine basic laws to manage whoever getting down to work and turn into a data miner. These laws are originally Thomas Khabaza's idea and generally they are 9 Laws in total [1]. The 9 laws of data mining recommend, and their clarifications illustrate, that these advancements won't change the idea of data mining cycle in whatever development is going to take place in the future, they just ought to be utilized to pass judgment, notwithstanding their instructive incentive for data miners.

1.1.1 First Law: Business Goals Law

You must foster a propensity for distinguishing business objectives prior to doing whatever else and zeroing in on those objectives at each progression in the information mining measure. It's huge that this law starts things out. Everybody ought to comprehend that information mining is a cycle with a reason.

1.1.2 Second Law: Business Knowledge Law

Information mining offers capacity to individuals -finance managers- who utilize their business information, experience, and knowledge, alongside information mining strategies, to discover importance in information.

1.1.3 Third Law: Data Preprocessing Law

Statisticians usually prepare or collect the datasets, yet after doing that they actually invest a ton of energy and time cleaning and getting ready information for examination. Whether I the data mining case, we quite often need to work with whatever information is accessible. We have to utilize existing business records, public information, or purchasable data. Odds are, all that information was accumulated for some reason other than data mining, and with no thorough arrangement or cautious information assortment measure. So, data miners invest a great deal of energy on data arrangement.

1.1.4 Fourth Law: Right Model Law

This law according to Thomas Khabaza [1], can be concluded by "The right model for a given application can only be discovered by experiment" which indicates that there is no right model to be applied directly to any dataset that fits it 100%, but, some models can be useful and show good results and the only way to find them is to experiment on the dataset.

1.1.5 Fifth Law: Amplification Law

The reason behind using the word "amplification" according to Khabaza is because of *"Data mining amplifies perception in the business domain."* Means that data mining is empowering the revelation of impacts that would be troublesome or difficult to identify through standard revealing of datasets.

1.1.6 Sixth Law: Pattern Law

All the datasets are somehow interconnected, hence there should be a certain pattern that connects the data, and it is undiscoverable unless we thoroughly study and experiment our dataset.

1.1.7 Seventh Law: Prediction Law

Prediction or estimation is finding general or global information about the data from local data points. By replacing the incorrect estimations by an interconnected data-driven one and even make new predictions if needed.

1.1.8 Eighth Law: Value Law

Making estimation after studying and experimenting on the dataset may give a good results and high accuracies but that does not mean that it has to make a business sense. In other words, data-driven accuracy may not give the same business-driven one

1.1.9 Nineth Law: Change Law

Models that make sense today may not do so in the future, because the science along with whole world is changing, therefore models, data and nowadays prediction can change as well. So the idea is not to get any model for granted and to understand that everything in the data mining field is changeable.

1.2 MACHINE LEARNING

Machine learning (ML) creates a numerical model based on example data, sometimes referred to as "preparing information," to make predictions or decisions without being explicitly configured to carry out the task. In any case, Machine learning is an overall term and it contains a few sub-models as demonstrated in figure 1.1.

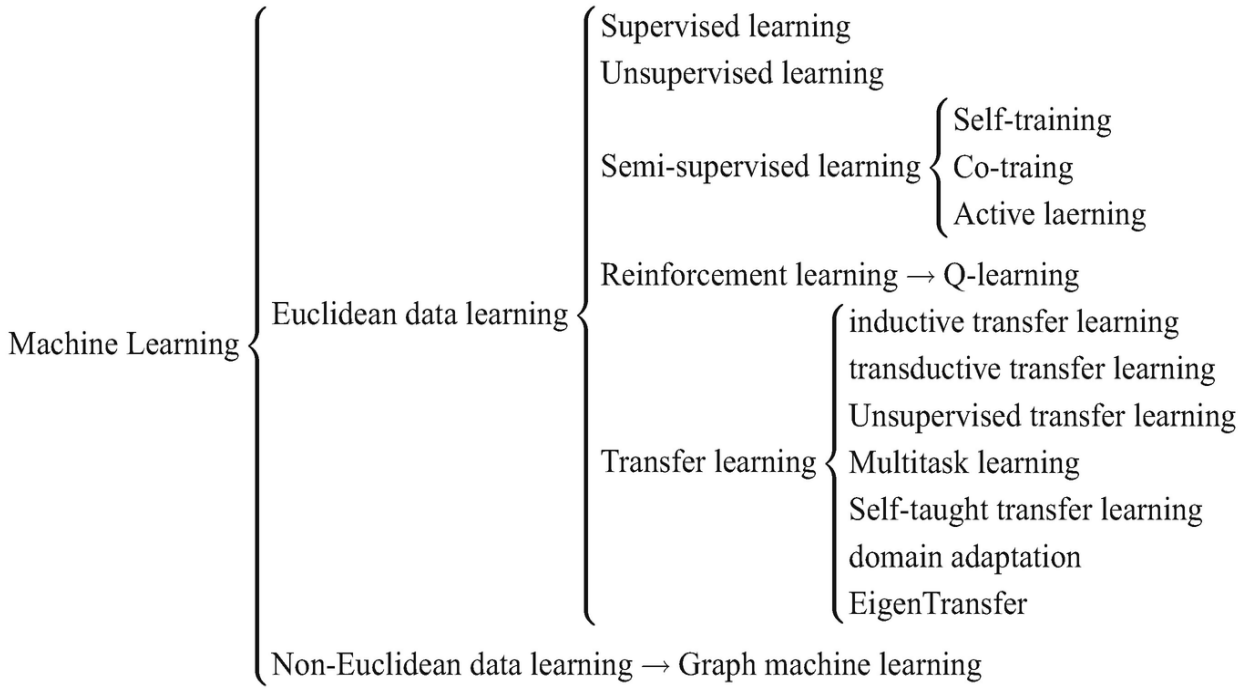

Figure 1.1: Machine learning tree [2].

To uncover patterns in an existing dataset, data mining is utilized. Contrarily, machine learning is educated using a collection of "training" data, which teaches the computer how to interpret data and subsequently make predictions about fresh sets of data. Data mining is basically a technique for doing research to identify a certain result based on the sum of the collected data. However, the tasks that a machine learning model is going to perform can be divided into three parts: Regression, Classification and Clustering.

1.2.1 Regression

In statistical modeling and machine learning, regression is to fit a measurable cycle for assessing the connections among factors. It centers around the connection between a reliant variable x and at least one autonomous factor (named as "indicators"). All the more explicitly, relapse is to investigate how the regular worth of the reliant variable changes when anybody of the autonomous factors is fluctuated, while the other free factors are fixed.

1.2.2 Classification

It is quite possibly the most oftentimes experienced dynamic errands of human movement. A Classification issue happens when an item should be allocated into a predefined gathering or classes identified with that article subsequent to bunching various gatherings or classes dependent on a bunch of noticed traits. At the end of the day, one necessity to choose which class among various classes a given testing test has a place with. This progression is known as the testing stage, contrasted and the past preparing stage. Unmistakably, this is a type of administered learning.

1.2.3 Clustering

Leave w alone the weight vector of a group which focuses on designs inside a bunch being more like each other as opposed to are designs having a place with different bunches. Clustering strategies are utilized to track down the real planning among examples and names (or targets). This arrangement has two AI techniques: directed realizing (which makes unmistakable naming of the information) and unaided learning (division of the information into classes), or a blend of more than one of these errands. The over three phases (information portrayal, include determination, and bunching) have a place with the preparation stage. Having such a bunch, one may make an arrangement in the following testing stage.

1.3 MARKETING DATA

Data plays a huge role in marketing especially in the last decade with the huge development in digital marketing and E-Commerce alongside with the traditional marketing tasks, as one can notice, most of the recently done researches in the data mining field are done on marketing data. However, the type of data highly affects the model building, for instance, in machine learning tasks the data type and its related data columns can totally change the performance of a certain model. Usually, the datasets for data mining tasks pass through some data preprocessing steps as we will explain in the next chapters. The ascent of non-understandable data is reshaping markets and the administration of showcasing exercises. However, this expanded part remains by numerous organizations, recommending the potential for additional examination

advancements. Offering a bringing-together definition and conceptualization of this type of data; connecting disjoint writing with a getting sorted out system that combines different subsets of important for advertising the executives through an integrative survey and distinguishing meaningful, computational, and hypothetical holes is the whole idea about our work.

1.4 DATA-DRIVEN MARKETING

Usually, data-driven marketing implies higher understanding of the market before applying a marketing strategy. So, it would make more sense to do so. But to collect a beneficial data for our strategy is not an easy process [3]. However, data-driven marketing can be considered as a way to understand the costumers before offering them the products. In general, this approach tells us that there are two types of consumers or “costumers”: ones that follow behavioral patterns whom their online and even offline behavior is predictable and ones who are marketing-consumers. Especially if a machine learning approach is being applied, an inside-data patterns can be found, ergo, the service providers can have pre-ideas about a specific costumer depending on specific data about him taken from his previous behavior or as a variable type data.

1.5 COSTUMERS’ BEHAVIOUR

Business owners are trying to know what their costumers’ behavior will be like in the future so they can plan for their business future accordingly, this is usually are being done by thoroughly studying costumers’ data and previous behaviors. But, making smarter data-driven decisions is not an easy task especially with no fixed tangible data about the changing costumers. However, there are some steps for making predictions about the costumers’ future behavior [4]:

1.5.1 Setting up a Marketing Goal

Our objective should have the option to deliver testable forecasts (Predictions). Having the option to say "sells will increment" is excessively ambiguous. When will they increment? By what amount? Your predictions should pass the "Perceptiveness Test". A superior objective is to "distinguish which clients will make a retail buy inside the following for example 10 days

with 0.9 accuracy". Obviously, there are numerous different practices you might need to anticipate, for example, clients' historical buying or reaction to a specific ad.

1.5.2 Use the Right Dataset

After setting up the goal, we must search what information we have to accomplish it? A decent method to do this is to work in reverse. We need to anticipate a sell goal, so it will be exceptionally convenient to think about verifiable buys. It may likewise assist with knowing the number of things clients normally purchase, how regularly do they make exchanges, do they purchase during discounts, after remunerations, before their birthday. The most sensible indicators are generally the most instructive, however there is no assurance, and picking the right information (known as highlight choice) can regularly be more workmanship than science. What amount of information shall we require? Shockingly, there is no unmistakable guideline for the measure of information we shall require, yet on account of retail drifts, you will need somewhere around 2 years of chronicled information to have the option to consolidate occasional patterns into your model.

1.5.3 Build a Simple Model

Selecting the right modelling tool is important along with the convenient libraries, for example, Scikit-Learn in Python or Caret package in R. the two of them have the ability to execute a huge assortment of complex AI calculations, and keeping in mind that it is enticing to go extravagant, the main thing is to begin with a basic model. Not on the grounds that these are the least complex to fit and generally interpretable, but since you can turn them around rapidly. This is basic when building prescient models since this is an iterative interaction and the greatest increases can be made rapidly. The danger of beginning with a confounded model is that you don't have the opportunity to develop it, or more awful the complex model is no greater than the straightforward model and it's less interpretable.

1.5.4 Testing the Prediction Model

For our situation, We needed to identify which customers are most likely to make a purchase in a specific amount of time. The consequence of our model is just the expansion of another segment or variable to our information with the mark 'Buy' or 'No Purchase' for every client in our data set. Practice great information cleanliness. Ensure you test your forecasts on a free dataset that has not been utilized in preparing the model. As a rule, this will be founded on an example on a 'wait' set of the information, regularly 20%. This will give you the most dependable gauge of how your model will act in reality'.

1.5.5 Set up the Model into the Market

To utilize your model, we need to put it into the market, ergo we will have the following arrangement choices:

(1) Keep it neighborhood: The least complex technique is to just run the model on the machine that created the model. This requires a minimal measure of work to set up.

(2) Clump it: Set up a Cron work, or other robotized administration to run the model each hour/day/week and add the forecasts to your data set.

(3) Allow it to rest: By making a RESTful API, anybody in your association can undoubtedly get to the model by means of HTTP and produce expectations progressively.

Whatever your arrangement choice, remember that a prescient model is just however great as the information it seemed to be based on. This implies that on the off chance that you just let the model sit for a really long time, the information it was prepared on will turn out to be more out of information, giving progressively more terrible forecasts. Continuously make sure to keep your models prepared with the latest information or, more than likely they will rapidly lose their worth without you understanding. Finally, Prescient demonstrating is an iterative cycle. Follow these 5 stages and you will actually want to rapidly create expectations, however to stand the trial of time, return to your model habitually.

2. LITERATURE REVIEW

2.1 DATA MINING FOR MARKETING

Data science for marketing has a long custom of accepting new difficulties, new strategies, and new trains. The field today is based upon the assorted endeavors of analysts who, for very nearly 50 years, have incorporated arrangements from an assortment of controls to give new knowledge to advertising issues. More regularly than not, the cauldron of promoting science has rewarded different controls, models, and strategies that are better and more powerful [5]. As shown in figure 2.1 below the relationship between data engineering, data science and data-driven decision making is ubiquitous.

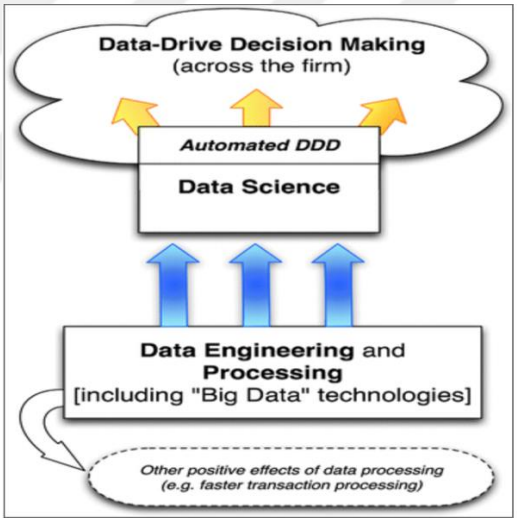

Figure 2.1: The relationship between Data Engineering, Data Science and Data-Driven Decision Making [6].

Big data on the other hand, has received considerable attention in the last decade alongside with the exponentially increasing number of internet users i.e., costumers, and due to the technological development, and the consideration of digital platforms for advertising the amount of data is becoming enormously high. Therefore, the adoption of faster data analysis techniques is urgent. The researchers around the world are putting a lot of effort and investing time and money to find considerable solution to this amount of data. Table 2.1 below shows some of the researches done in this field.

Table 2.1: Business-related data mining researches

Ref.	Method	Data Type	Study Objective
[7]	Big Data	Marketing data	Overview
[8]	Marketing Science	Big Data	Moving from products-based to customer support-based economy
[9]	Big data analysis	Marketing data	New product success
[10]	Data analysis	Social media Marketing data	Promotions of web links on Facebook
[11]	Data analysis	Quick-Service Restaurant	Workers ingaging into busineeses and its affection on the agencies profit.
[12]	Data analysis	Mobile Applications Data	Restaurants' brand development in customers satisfaction and loyalty during COVID-19 period
[13]	Data analysis	Business Data	Effects of COVID-19 on businesses
[14]	Data Analysis/ Prediction	Retailers' Data	Retailers' ups and downs during the COVID-19 outbreak
[15]	Machine Learning	Airlines Customers data	Prediction of customer return visits in airline services
[16]	Machine Learning Business Value (MLBV)	Business Data of 319 corporations	Derive business value (BV) from Machine Learning
[17]	Machine Learning	Online Marketing Data	Auto-tagging online content
[18]	Machine Learning	TripAdvisor Big Visual Data	Branding luxury hotels
[19]	Internet of things Artificial Intelligence Machine learning Blockchain Marketing strategy	Marketing Data	New tech. influence on markets
[20]	Data Analysis	Social Media Data	Impact assessment of social media usage in B2B marketing

Customized marketing is becoming more prevalent, and Marketing Science techniques that take use of client heterogeneity are becoming more important. Data Technology improvements also ensure the growing importance and use of computationally focused data handling and "Big

Data." In particular, these trends have been present for more than a century and will continue to be more clearly articulated for a very long time due to the constant idea of technology advancement. These developments lead to a shift in marketing science, both in terms of the ideas to be emphasized and the methods to be used. Data analysis is widely used in business. For instance, the study in [11] demonstrates that staff engagement impacts operational, customer, and financial performance measurements in quick-service restaurants both directly and indirectly (QSR).

2.2 STUDY OF BUSINESSES AFFECTION BY PANDEMICS AND NATURAL DISASTERS

A pandemic caused by the COVID-19 virus that originated in China and spread over the whole world at the beginning of 2020 brought about a sudden realization that pandemics, like other sometimes occurring tragedies, had happened in the past and would continue to do so in the future. Even if we can't stop dangerous viruses from spreading, we should make plans to lessen their effects on society [13]. No country appears to be immune to the severe financial effects the current flare-up has had throughout the world. This has effects on all aspects of society, not just the economy, which has led to spectacular shifts in how consumers and corporations behave. This remarkable problem is an international effort to solve some of the pandemic-related problems plaguing humanity. There are a total of 13 articles that discuss various business sectors (such as the travel industry, retail, and higher education), alterations in consumer behavior and corporate practices, moral dilemmas, and viewpoints associated with employees and authority. The marketing industry also received a piece of the COVID-19 action, which led to the insolvency of some well-known companies as consumers stayed at home and economies were shut down [21].

2.3 MACHINE LEARNING BASED MARKETING

ML is quite possibly the most famous methodologies in Artificial Intelligence. Over the previous decade, Machine Learning has gotten one of the essential pieces of our life. It is executed in an errand as straightforward as perceiving human penmanship or as unpredictable as self-driving vehicles. It is likewise anticipated that in years and years, the more mechanical

monotonous assignment will be finished. With the expanding measures of information opening up there is a valid justification to accept that Machine Learning will turn out to be significantly more pervasive as a fundamental component for mechanical advancement. There are many key enterprises where ML is having a colossal effect: Financial administrations, Delivery, Marketing and Sales, Health Care to give some examples. Be that as it may, here we will talk about the execution and utilization of Machine Learning in exchanging. Advertisers use machine learning which assist them with anticipating the further conduct of clients and rapidly upgrade publicizing offers by discover patterns in client exercises on a site. However, what is the capability of these costumers' patterns? In brain research, a pattern is a specific arrangement of social responses or a typical grouping of activities. Hence, we can discuss designs as to any space where individuals use layouts (which is most everyday issues). Consider the case of an example utilized on sites. In the event that the client isn't keen on the proposal in the spring up window displayed underneath, they can close this window by: tapping on the X sign, clicking No much obliged, clicking anyplace on the site that is outside the spring up window. Notwithstanding these three moves the client can make, the spring up window will close all alone after a specific timeframe. So, we get four potential client activities: Click on X sign: Can be valid/bogu, Click No much appreciated: Can be valid/bogus, click past the spring up: Can be valid/bogus. And Spring up survey time is 5 seconds. At the point when many such boundaries are gathered, the gathered information acquires esteem since it contains examples of conduct and conditions. It shrouds the colossal capability of conduct information, permitting us to enhance client information with the missing boundaries dependent on the information we as of now have for different clients. For instance, the easiest method to characterize an intended interest group is by gender and age. However, imagine a scenario in which clients round out this information just in 10% of cases. How might you see what number of your site clients fall into your intended interest group? Examples of conduct can help. You can utilize gender and age information from 10% of clients to decide designs explicit to a specific gender and age. Then, at that point you can utilize these examples to foresee the gender and age of the leftover 90% of clients. Having total information about gender and age, you would now be able to make customized offers to all site guests.

2.3.1 Revolutionizing Marketing by Machine Learning

According to a research done by Columbus in [21], Machine Learning is “revolutionizing” the marketing field. 58% of businesses prioritize customized customer service and new product development over solving the most difficult marketing challenges using AI and machine intelligence. Improvements in customer support and experiences, according to 57% of business executives, will result in the biggest growth opportunities for AI and machine learning. By 2022, real-time personalized advertising across digital platforms and improved message targeting accuracy, context, and precision will accelerate. Additionally, assess customer churn and dramatically lower it using machine learning to improve risk prediction and intervention models. Beyond sectors with constrained inventories, like travel and lodging, price optimization and price elasticity are spreading into manufacturing and services. Retail marketing improvements in demand forecasting, assortment effectiveness, and pricing have the potential to result in an increase in Earnings Before Interest & Taxes (EBIT) of 2%, a 20% stock reduction, and a decrease of 2 million product returns annually. the development and improvement of propensity models to direct cross-sell and up-sell strategies which prospects through which channels is another way machine learning is revolutionizing marketing.

2.4 DATA GRAPHING IN MARKETING

Graphs at times don't generally get the credit they merit in the business world. Frequently, they are kidded about as being senseless visual guides. Actually, they offer incredible benefit. Illustrations are ordinarily used to all the more likely address a bunch of results or examples and Contribute to the development of an inquiry. They can work on psychological thinking and increase the extent of how an assessment has ended up by filling in as illustrative images. The concept of information representation is a fantastic tool for surveying company execution. In the corporate world, the executive's graphical evaluation is critical in bringing critical facts and making an appropriate therapeutic action. Making diagrams and graphs to image information and appreciate insights is the ideal way to manage revealing and following the market focuses of firms. You may ask how a simple outline can accomplish this objective, yet you will be shocked by how amazing a visual example is in understanding monetary reports that simple

numbers and figures. Charts intelligently address data along a few metrics based on how the available insights are shown. The primary function of charts is to highlight relationships between things, which can include everything from profit and loss statistics to agreements and advertising figures in the business sector. The normal sorts of diagrams are line and structured presentations, pie outlines, dissipate plots and bar charts. Overall outlines address one kind of data, for instance, you may show the level of benefits from different states in the country. Diagrams then again show one bunch of factors addressed in a persistent stream against another variable element, for example, the yearly marketing projections of the previous 10 years or something almost identical. The expanding ease with which charts would now be able to be made just as the extent of appealing visuals has made an effect in the business field. It is fascinating to take note of that charts can disguise or uncover data as is wanted and will rely upon the sort of diagram picked and the degree of detail organized. For example, the pie outline may give an image of relative amounts of every division, except if an exact mathematical figure or rate share is required it very well may be smarter to go in for a plain configuration than a chart. Consequently, understanding the reason for introducing the data is basic to choosing the right kind of graphical showcase. Consider a straightforward line graph to address the example of products sold throughout some undefined time frame. A diagram, for example, this viably uncovers the example of deals, and can likewise be utilized to think about the qualities for a few producers. So how does this assist with improving the business you wonder? It's a genuinely clear methodology truly. If one somehow happened to see the individual deals upsides of an organization throughout the long term, accepting there has been a consistent move in deals, then, at that point one is probably going to reason that the organization is promoting its items right. Presently that is quite essential. Be that as it may, an examination of comparing information from organizations inside a similar industry may show a stamped distinction, which implies your business isn't working out quite as well as you expected! In spite of the fact that you might have the option to deduce this little snippet of data by contemplating pages and pages of organization reports, the straightforwardness with which a solitary diagram can recount the entire story is unquestionable. So presently you realize where your organization remains as well as have the option to quantify and set future focuses for the following year. The interaction of compelling graphical development starts with a basic investigation of the data accessible.

Example location comes in extremely convenient to choose the right sort of visual that will best address your information. Chart development is an iterative cycle implying that there is plentiful extension for experimentation to evaluate what works best. Given the ubiquity and adaptability of illustrations and the significance of the examples uncovered by utilizing pictures, diagrams are key dynamic devices for any endeavor.

3. DATASET, EXPIREMENTS AND RESULTS

3.1 MACHINE LEARNING IN PYTHON

Artificial Intelligence and Machine Learning-based tasks are clearly what's in store these days since we need better personalization, more intelligent proposals, and further developed pursuit usefulness. Now our systems, devices and applications can see, hear, and react, that is the thing that machine learning has brought while improving the client experience and making esteem across numerous businesses. Presently we are probably facing two inquiries: How would we be able to rejuvenate these encounters? what's more, what programming language is utilized for machine learning? Python programming is considered as a powerful tool to compensate for machine learning needs thanks to its powerful libraries that are designed to deal with big and complicated datasets. NumPy library provides dealing with huge, arrays and matrices, as well as a wide variety of high-level mathematical functions for working with these arrays. [22]. Sci-kit Learn library on the other hand contains easy and efficient tools for predictive data analysis such as machine learning and data preprocessing [23]. Another useful library is SciPy which is used for scientific computing and technical computing. Which contains modules for optimization, linear algebra, integration, interpolation, and special functions [24]. However, to build a machine learning model few steps are needed to be done. These steps are usually referred to as “data preprocessing”. In data preprocessing stage NumPy is used to apply some modifications like feature scaling, deal with missing data and splitting the dataset into training and testing sets. However, these steps might differ according to the type and the size of the dataset as well as the number of features (Column) in it.

3.2 DATASET

Our dataset describes the behavior of a certain group of customers from different ages with different salaries on a certain Online Advertising material. However, the dataset contains 400 rows and three columns (Age, Estimated Salary and Purchased), the content of the column “Age” are the ages of the costumers, the column “Estimated Salary” contains a range of salaries dedicated to each one of the costumers and finally, the column “Purchased” contains 0’s and 1’s

where 0 refers to no response from the costumer toward the Ad and 1 for the costumers who interacted with the Ad, Table 3.1 below shows the first 10 rows of the dataset.

Table 3.1: First 10 rows of the dataset.

#	Age	EstimatedSalary	Purchased
0	20	17000	0
1	35	25000	0
2	25	44500	0
3	30	56300	0
4	19	77700	0
5	27	58000	0
6	27	84000	0
7	32	150000	1
8	25	33000	0
9	35	65000	0

We have checked if the dataset contains any null data or (missing data) by using the function "info" and check the "is null" column the output of this function is shown in Table 3.2., we can notice that the all the columns are integers.

Table 3.2: The output of info function

Column Name	No. of Rows	Is Null	Data type
Age:	400	NO	Integer
Estimated Salary:	400	NO	Integer
Purchased:	400	NO	Integer
D types:	Integer		
Data Columns:	Total 3 columns		

Then, we checked the content of the columns by using the function "Describe" and the output is shown in Table 3.3. We can notice that minimum age of the studied costumers is 18 and the maximum is 60, whether the minimum estimated salary is 15,000 and the maximum is 150,000.

Table 3.3: The output of describe function.

	Age	Estimated Salary	Purchased
count	400	400	400
mean	Thirty-seven (37)	70000	/
std	Ten (10)	34000	/
min	Eighteen (18)	15000	/
25%	Twenty- nine (29)	43000	/
50%	Thirty Seven (37)	70000	/
75%	Forty-six (46)	88000	1
max	Sixty (60)	150000	1

3.3 DATASET VISUALIZING IN PYTHON

In Python programming there are two main libraries for data visualization, Matplotlib and Seaborn. Matplotlib is the base plotting library in python [25], whereas seaborn is more advanced plotting library as well [26]. In figure 3.1 and figure 3.2 we show plots of costumers who responded negatively and positively to the Ad, respectively.

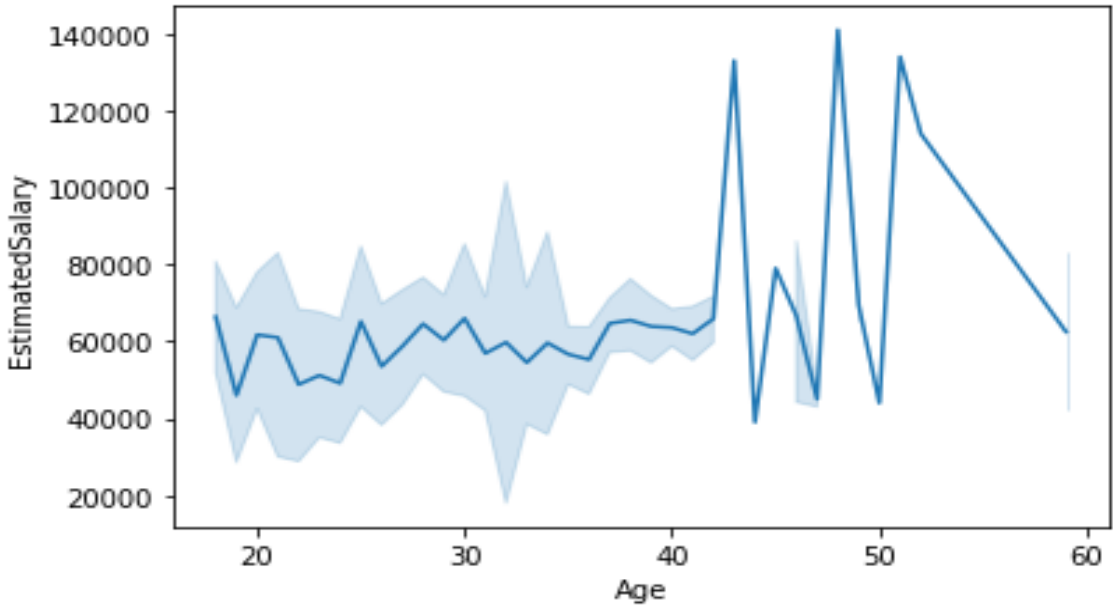

Figure 3.1: line plot of costumers who responded negatively to the Ad.

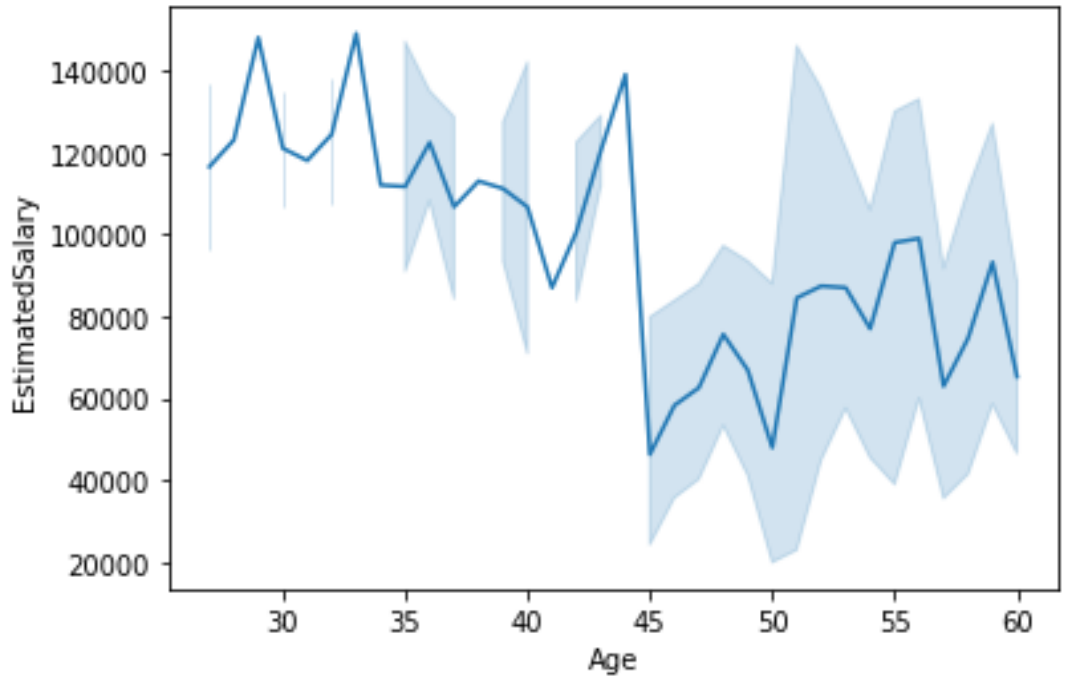

Figure 3.2: line plot of costumers who responded positively to the Ad.

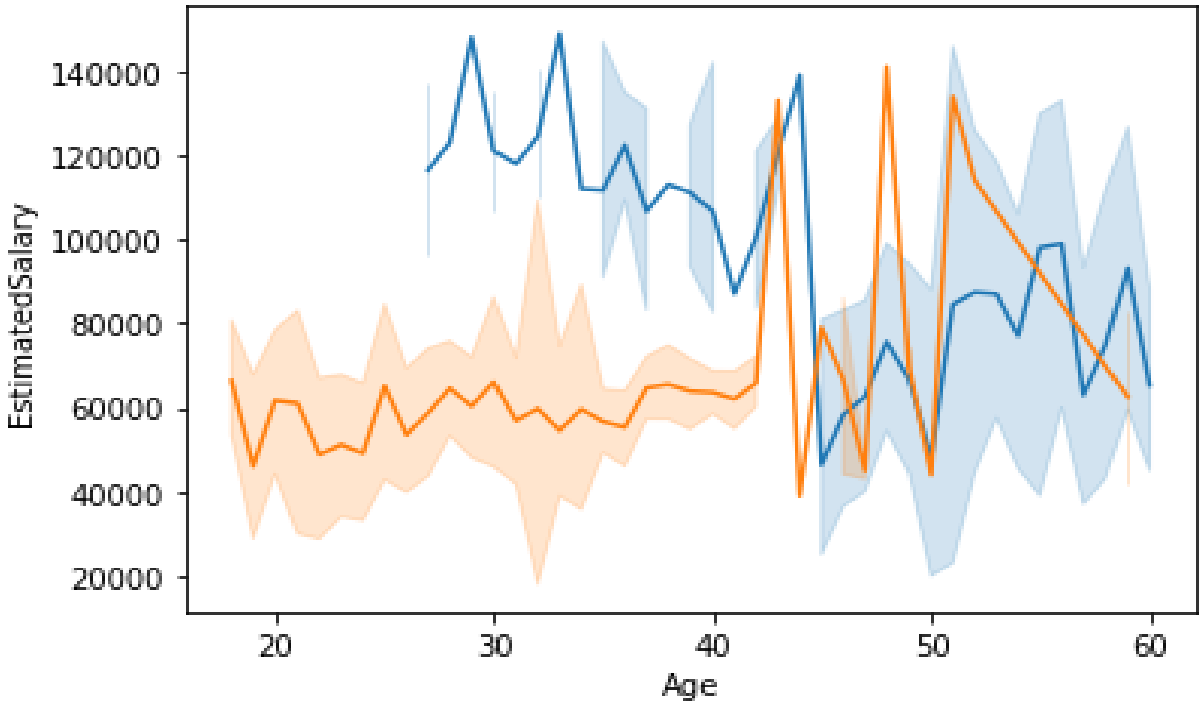

Figure 3.3: Line plot of costumers and their respond to the Ad (Blue = 1, Orange = 0)

All the figures (3.1, 3.2 and 3.3) are done using Seaborn Library and Line Plot function. However, we can notice that when the costumers are younger (25-40) and only when their salary are high, they are more likely to respond positively to the Ad, and when they are (<27) no matter how their salaries are, they will probably respond negatively. While when the ages are (>40) the costumers will respond positively to the ad. Such conclusions were so hard to make without these Seaborn powerful graphs, however, we can notice that although some costumers have high salaries and their ages are over 40, they responded negatively. So just by data visualization it is almost impossible to predict the costumer's behavior, ergo, powerful machine learning models will be used to predict the costumer's behavior more accurately.

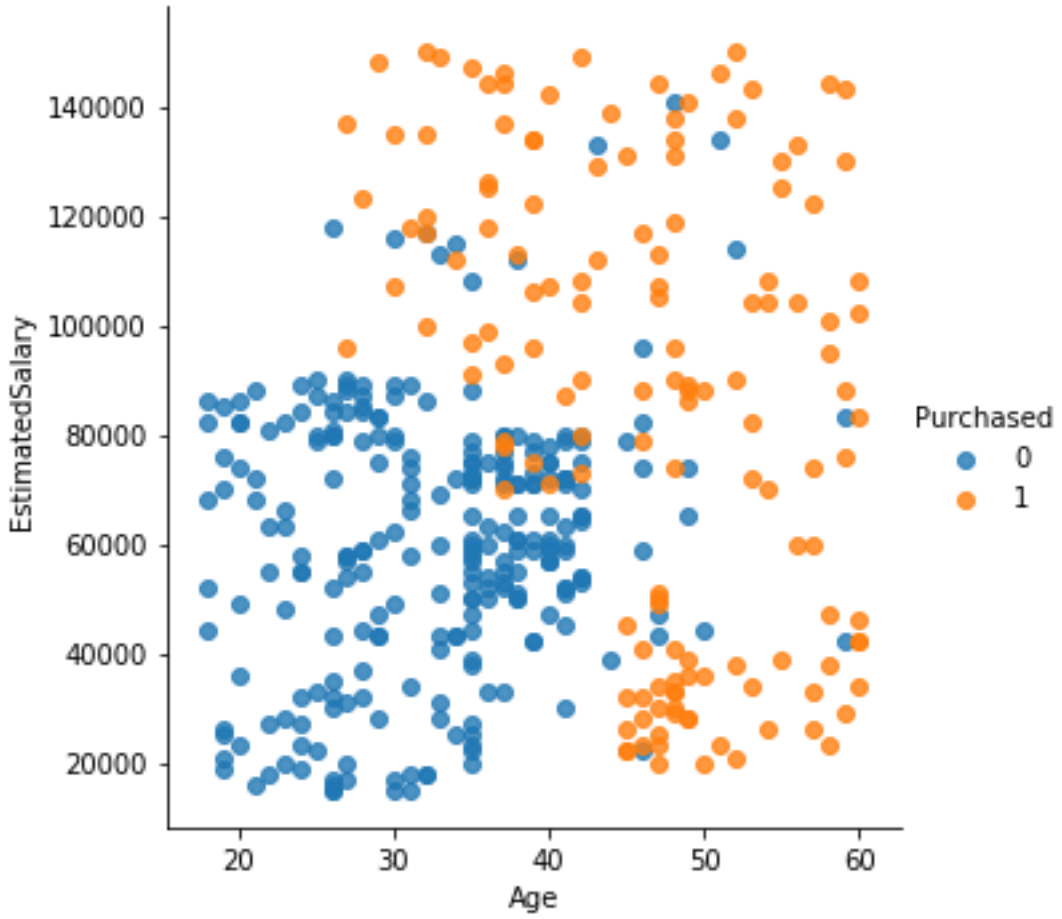

Figure 3.4: Lmplot of costumers who responded positively to the Ad.

Another powerful graph is the LmPlot shown in Figure 3.4 which show each independent costumer and whether he responded positively (1) or negatively (0), here one can see the power of using Python and its Seaborn since this plot has been drown by only one line code.

3.4 DATA PREPROCESSING OF ADS DATASET

Before applying any machine learning model to the dataset that we have, there are some data-preprocessing steps that need to be done. First, we need to import the libraries shown in table 3.4, it should be mentioned that these libraries will be used via their abbreviation.

Table 3.4: The libraries needed for ad dataset and their corresponding abbreviations.

Library	Abbreviation
Numpy	np
Matplotlib. pyplot	plt
Seaborn	sns
Pandas	pd
Scikit learn	sklearn

After importing the libraries, we need to specify which columns (Feature) of our dataset are the independent variables and which one is the dependent variable by using “iloc”. However, in our case as in Table 3.1 the dependent variable is the one indexed -1, since in Python if we want to select the last column instead of start counting from the right to reach the last column, we can easily use the negative indexing which will start from the left by -1, in this case the -1 indexed column is our dependent variable. Then we need to split the data set into train and test sets by importing “train_test_split” function from “sklearn.model_selection”. Which will split the dataset by a percentage we give to the test set by specifying the value of the parameter “test_size” in our case we used 0.2 which will split our data into 80% for training set and 20%

for testing set. Later on, from “sklear.preprocessing” we imported “StandardScaler” which will be fit transformed into our X_train and X_test this will normalize all the dataset such that its distribution will have a mean value 0 and standard deviation of 1, which will help the classifier in giving better predictions since it is hard to discover patterns between data points there are big gaps between their values.

3.5 MACHINE LEARNING ON ADS DATASET

Machine Learning has different types, supervised learning, unsupervised learning, and reinforcement learning. Choosing which one to use highly depends on the dataset as well as the type of the independent variable of our dataset. for our dataset and since we are grouping the costumers into categories to expect their behavior which is either 1 or 0 i.e., positive or negative response, the appropriate type to use is classification. For our dataset, to predict the costumer’s behavior three different classifiers have been used: Naïve Bayes, K-Nearest Neighbors (KNN) and Support Vector Machine (SVM).

3.5.1 Naïve Bayes

The core of Naïve-Bayes classifier depends on Bayes hypothesis to build a probabilistic machine learning model that’s used for classification task. The name naive is utilized on the grounds that it expects the columns that go into the model is free of one another. That is changing the worth of one element, doesn't straightforwardly impact or change the worth of any of different highlights utilized in the calculation. By its hints, Naive Bayes is by all accounts a straightforward yet amazing calculation. However, it so mainstream. That is on the grounds that there is a critical benefit with NB. Since it is a probabilistic model, the calculation can be coded up effectively and the forecasts made genuine speedy. Constant speedy. Along these lines, it is effectively versatile and is traditionally the calculation of decision for true (applications) that are needed to react to client's solicitations immediately. For our data set the classifier showed an accuracy of 0.8875 which means that 88% of the predictions were correct. Table 3.5 shows the confusion matrix of Naïve bayes classifier.

Table 3.5: Confusion matrix of naïve-bayes classifier.

	0	**1**
0	51	4
1	5	20

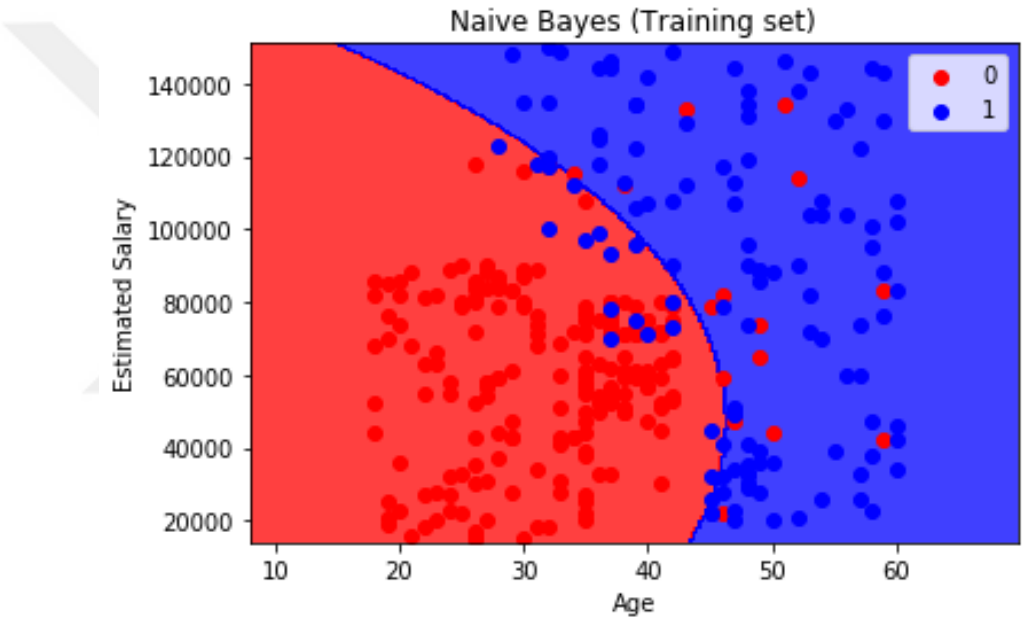

Figure 3.5: Graph of Naïve Bayes classifier for training set.

From Table 3.5 we can notice that the classifier predicted 51 true negative, 20 true positive, 5 false negative and 4 false positives. Naïve Bayes classifier's performance for training and testing sets are shown in figures 3.5 and 3.6, respectively.

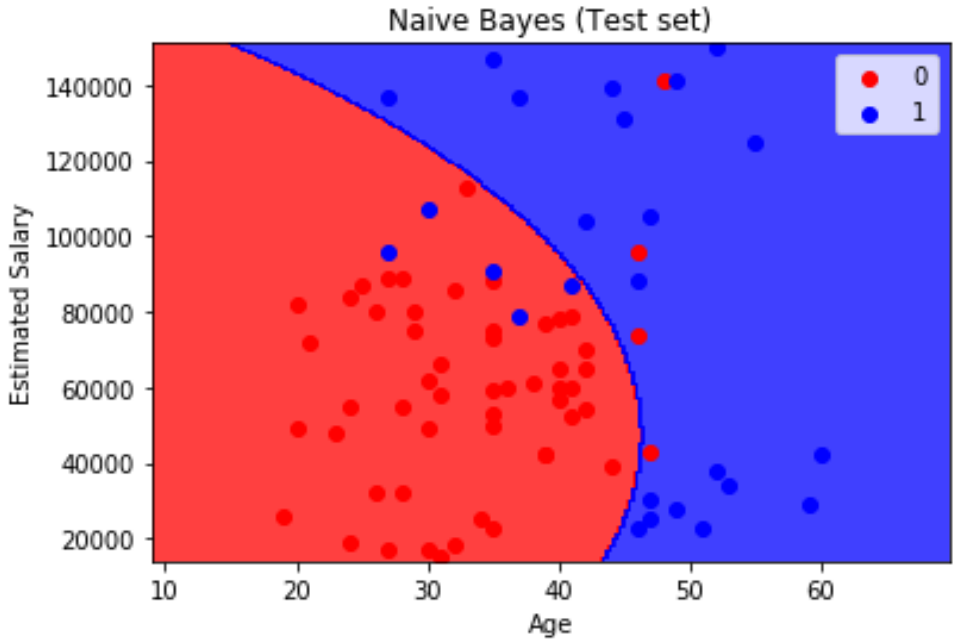

Figure 3.6: Graph of Naïve Bayes classifier for test set.

3.5.2 K-Nearest Neighbor (K-NN)

For relapse and order issues, K-Nearest Neighbors (KNN) is likely the least difficult calculation used in machine learning. KNN computations rely on similitude measures to characterize new information focuses and make use of existing information. A single data point's classification is mostly determined by the votes of its neighbors [29] The class with the closest neighbors receives the information. Exactness may increase as you increase the number of nearest neighbors and the value of k. The KNN algorithm believes that related things are located nearby. In other words, related things are located close to one another. 90% of the predictions made by the classifier for our data set were accurate, or 0.9 accuracy.shows the confusion matrix of K-NN classifier.

Table 3.6: Confusion matrix of k-nn classifier.

	0	**1**
0	50	5
1	3	22

From Table 3.6 we can notice that the classifier predicted 50 true negative, 22 true positive, 3 false negative and 5 false positives. K-NN classifier's performance for training and testing sets are shown in figures 3.7 and 3.8, respectively.

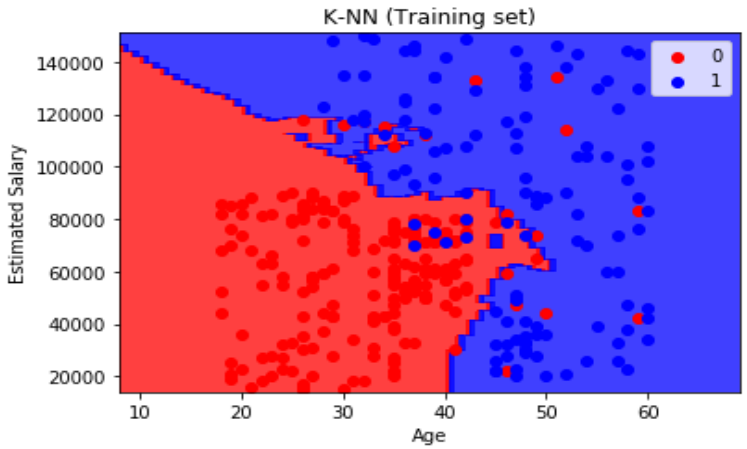

Figure 3.7: Graph of KNN classifier for training set.

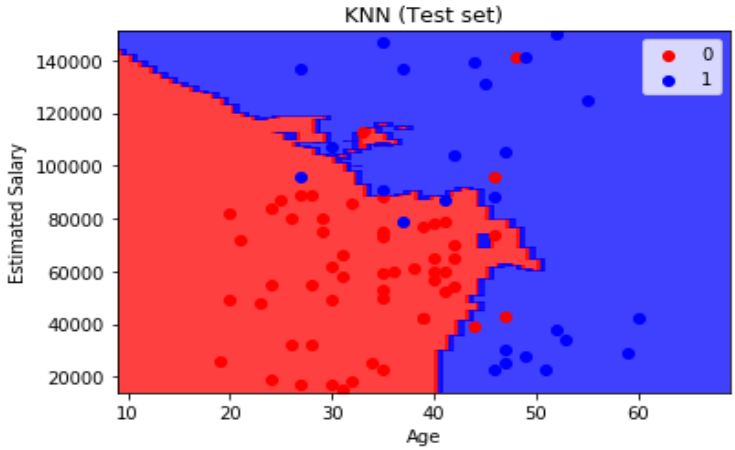

Figure 3.8: Graph of KNN classifier for test set.

3.5.3 Support Vector Machine (SVM)

The goal of the support vector machine approach is to locate a hyperplane in an N-dimensional space that clearly classifies the data points for an N-features dataset. There are a variety of potential hyperplanes that could be chosen to separate two classes of information points. The plane with the greatest edge—that is, the extreme separation between the information points of the two classes—is the one we are trying to find. Enhancing the edge distance provides some support, enabling more accurate characterization of upcoming information focuses. The classifier's accuracy for our data set was 0.8875, which indicates that 86.75% of the predictions were accurate. The confusion matrix for the Naive Bayes classifier is displayed in Table 3.7.

Table 3.7: Confusion matrix of naïve-bayes classifier.

	0	**1**
0	49	6
1	4	21

From Table 3.7 we can notice that the classifier predicted 49 true negative, 21 true positive, 4 false negative and 6 false positives. K-NN classifier’s performance for training and testing sets are shown in figures 3.9 and 3.10, respectively.

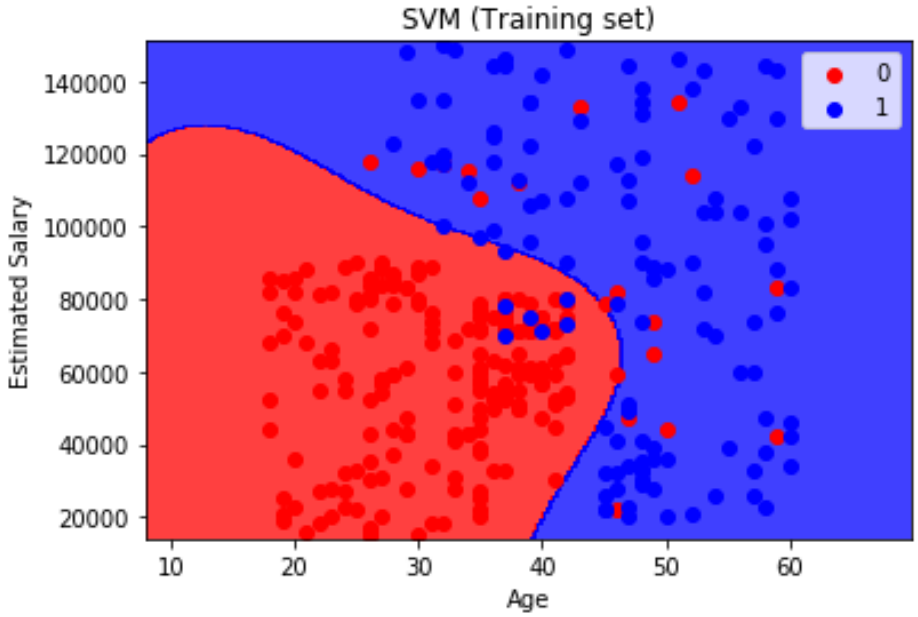

Figure 3.9: Graph of SVM classifier for training set.

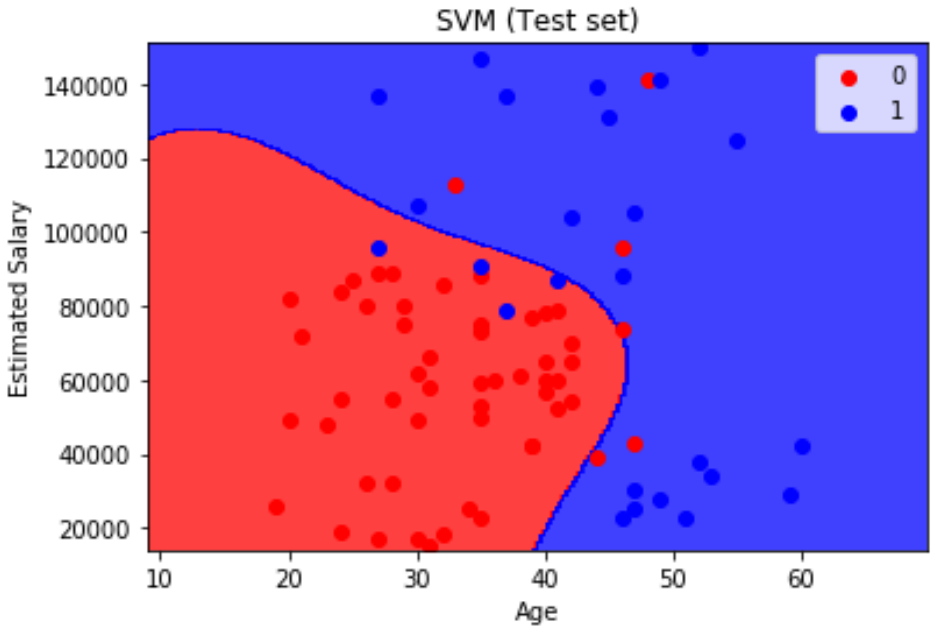

Figure 3.10: Graph of SVM classifier for test set.

3.6 ROC AND AUC SCORE

AUC (Area Under Curve) score addresses the likelihood that a True positive (Blue) is situated to one side of a true negative (ruddy). AUC goes in regard from 0 to 1. A demonstrate whose expectations are 100% off-base has an AUC of 0.0 though the one who got 100% right has an AUC of 1.0. Whereas a ROC bend could be a graph appearing the introduction of a characterization show at all gathering limits. This curve plots two boundaries: True Positive Rate and False Positive Rate [27]. However, for our models the AUC scores of the three models were 0.86, 0.89 and 0.86 for Naïve-Bayes, K-NN and SVM models, respectively. While the ROC curve for the three models were as shown in figures 3.11, 3.12 and 3.13.

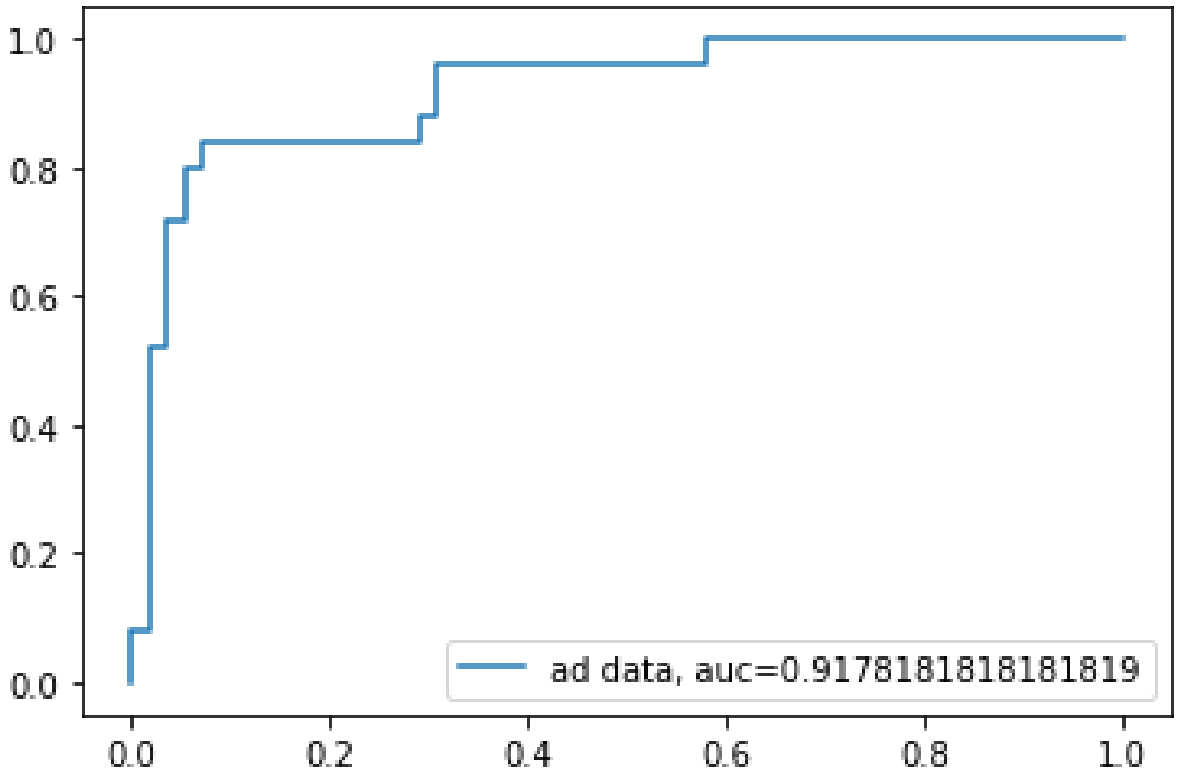

Figure 3.11: ROC Curve of Naïve bayes classifier.

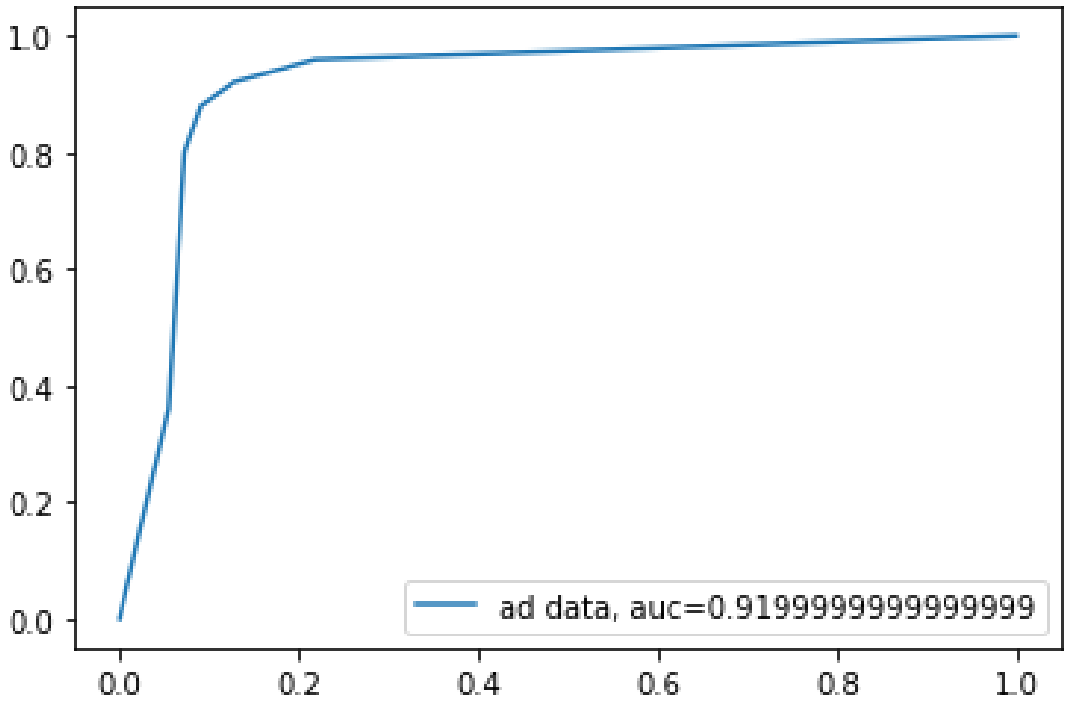

Figure 3.12: ROC Curve of K-NN classifier**.**

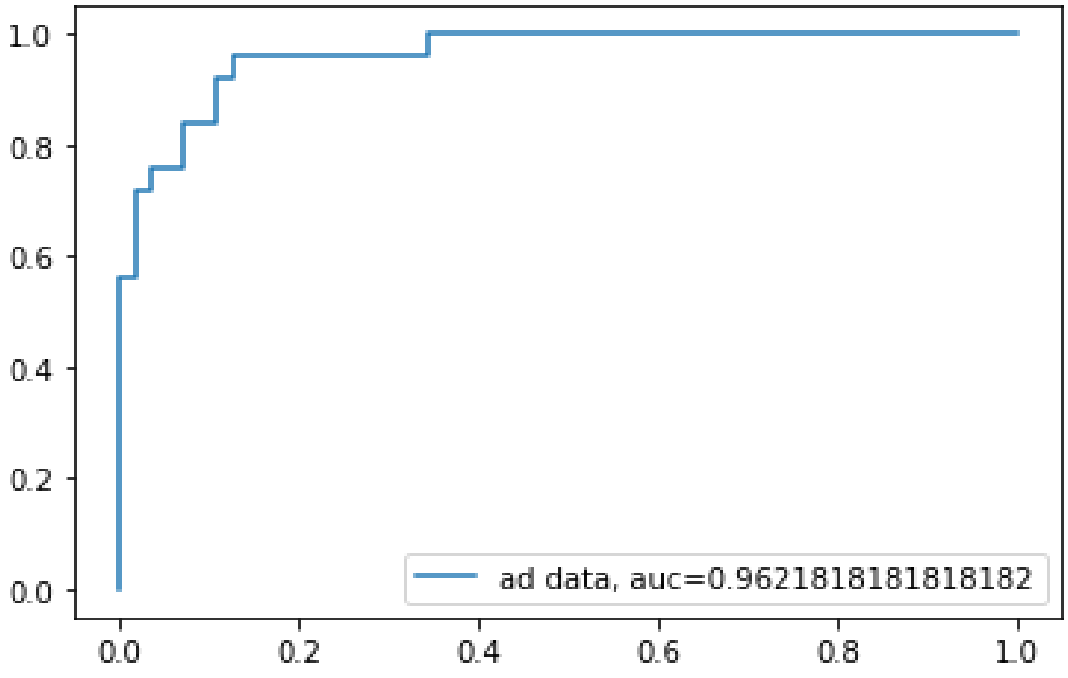

Figure 3.13: ROC Curve of SVM classifier.

4. CONCLUSIONS AND RECOMMENDATIONS

4.1 CONCLUSION

Marketing is the key to sell a product if it is done smartly, if we assumed that it has been done this way, it means that our product will reach the type of costumers that are more likely to buy it, in the last decades online Ad-based marketing is the main tool to bring more costumers for a certain product. However, in our research by using some information about the costumers like their age and expected salary we were able to predict whether they will respond positively or negatively to our product's Ad. First, we used Lineplot and Lmplot visualizations from Python's Seaborn to show the relation between the costumers' data and their respond to the ads, which showed that when the costumers are younger (25-40) and only when their salary are high, they are more likely to respond positively to the Ad, and when they are (<27) no matter how their salaries are, they will probably respond negatively. While when the ages are (>40) the costumers will respond positively to the ad. we can notice that although some costumers have high salaries and their ages are over 40, they responded negatively. So just by data visualization it is almost impossible to predict the costumer's behavior, ergo, powerful machine learning models will be used to predict the costumer's behavior more accurately. Three deferent classification models have been used: Naïve-Bayes, KNN and SVM, these models showed accuracies as shown in Table 4.1.

Table 4.1: ML models and their corresponding accuracies.

Model	Accuracy
Naïve-Bayes	88%
K-NN	90%
SVM	88.75%

The number of false positives, false negatives, true positives and true negatives in each model were as shown in Table 4.2.

Table 4.2: FP, FN, TP and TN of ML models.

Model	False positives (FP)	False negatives (FN)	True positives (TP)	True negatives (TN)
Naïve-Bayes	4	5	51	20
K-NN	5	3	50	22
SVM	6	4	49	21

Finally, we checked the performance of each model by calculating the AUC score (Area Under Curve score) and the results were as shown in Table 4.3.

Table 4.3: AUC score of ml models.

Model	AUC Score
Naïve-Bayes	0.86
K-NN	0.89
SVM	0.86

We can say that all of our three classifiers showed good results since they all have accuracies of over 88%, but K-NN classifier is the best with accuracy of 90% and AUC score of 0.89 out of 1.

Finally, we double checked whether our models are really predicting in real life by using the function “Predict” to predict how the costumer will respond to our Ad., however the test three

models showed that a costumer who is 20 years old and has a salary of 11,000 will positively respond to our Ad.

4.2 RECOMMENDATIONS FOR FUTURE WORK

As our machine learning models showed good performance in predicting the costumers' behavior for a certain advertisement material, by giving information about only the age and the expected salary of the costumers. This can help the marketing department in any company to develop tools that target the right costumers who are more likely to buy the product. However, some banks already have information about their costumers, so by using such machine learning models they can easily target their credit cards ads to the costumers that are probably going to get it. Ergo, for future work we highly recommend that more sophisticated models can be built in order to tackle the gap between the right costumers and the marketing departments. Neural Networks can be a good candidate for this task since it deeply studies the datasets and takes all the probabilities, but the dataset size should be taking into consideration before choosing the machine learning models.

REFERENCES

[1] T. Khabaza, "9 Laws of Data Mining," no. September, pp. 1–10, 2010.

[2] X.-D. Zhang, "Machine Learning," in *A Matrix Algebra Approach to Artificial Intelligence*, Singapore: Springer Singapore, 2020, pp. 223–440. doi: 10.1007/978-981-15-2770-8_6.

[3] M. Mulvenna, M. Norwood, and A. Büchner, "Data-Driven Marketing," *Electronic Markets*, vol. 8, no. 3, pp. 32–35, 1998, doi: 10.1080/10196789800000038.

[4] T. Paris, "Predicting Customer Behaviour in 5 Steps ," in *Medium*, 2017.

[5] P. Chintagunta, D. M. Hanssens, and J. R. Hauser, "Editorial-Marketing Science and Big Data," *Marketing Science*, vol. 35, no. 3, 2016, doi: 10.1287/mksc.2016.0996.

[6] F. Provost and T. Fawcett, "Data Science and its Relationship to Big Data and Data-Driven Decision Making," *Big Data*, vol. 1, no. 1, pp. 51–59, Mar. 2013, doi: 10.1089/big.2013.1508.

[7] P. Chintagunta, D. M. Hanssens, and J. R. Hauser, "Marketing science and big data," *Marketing Science*, vol. 35, no. 3. pp. 341–342, 2016. doi: 10.1287/mksc.2016.0996.

[8] R. T. Rust and M. H. Huang, "The service revolution and the transformation of marketing science," *Marketing Science*, vol. 33, no. 2, pp. 206–221, 2014, doi: 10.1287/mksc.2013.0836.

[9] Z. Xu, G. L. Frankwick, and E. Ramirez, "Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective," *J Bus Res*, vol. 69, no. 5, pp. 1562–1566, 2016, doi: 10.1016/j.jbusres.2015.10.017.

[10] Y. Chawla and G. Chodak, "Social media marketing for businesses: Organic promotions of web-links on Facebook," *J Bus Res*, vol. 135, pp. 49–65, 2021, doi: 10.1016/j.jbusres.2021.06.020.

[11] A. Lambert, R. P. Jones, and S. Clinton, “Employee engagement and the service profit chain in a quick-service restaurant organization,” *J Bus Res*, vol. 135, no. June, pp. 214–225, 2021, doi: 10.1016/j.jbusres.2021.06.009.

[12] T. Dirsehan and E. Cankat, “Role of mobile food-ordering applications in developing restaurants’ brand satisfaction and loyalty in the pandemic period,” *Journal of Retailing and Consumer Services*, vol. 62, pp. 1–28, 2021, doi: 10.1016/j.jretconser.2021.102608.

[13] N. Donthu and A. Gustafsson, “Effects of COVID-19 on business and research,” *J Bus Res*, vol. 117, no. June, pp. 284–289, 2020, doi: 10.1016/j.jbusres.2020.06.008.

[14] E. Pantano, G. Pizzi, D. Scarpi, and C. Dennis, “Competing during a pandemic? Retailers’ ups and downs during the COVID-19 outbreak,” *J Bus Res*, vol. 116, no. May, pp. 209–213, 2020, doi: 10.1016/j.jbusres.2020.05.036.

[15] S. Hwang, J. Kim, E. Park, and S. J. Kwon, “Who will be your next customer: A machine learning approach to customer return visits in airline services,” *J Bus Res*, vol. 121, no. August, pp. 121–126, 2020, doi: 10.1016/j.jbusres.2020.08.025.

[16] C. Reis, P. Ruivo, T. Oliveira, and P. Faroleiro, “Assessing the drivers of machine learning business value,” *J Bus Res*, vol. 117, no. July 2019, pp. 232–243, 2020, doi: 10.1016/j.jbusres.2020.05.053.

[17] J. Salminen, V. Yoganathan, J. Corporan, B. J. Jansen, and S. G. Jung, “Machine learning approach to auto-tagging online content for content marketing efficiency: A comparative analysis between methods and content type,” *J Bus Res*, vol. 101, no. April, pp. 203–217, 2019, doi: 10.1016/j.jbusres.2019.04.018.

[18] S. Giglio, E. Pantano, E. Bilotta, and T. C. Melewar, “Branding luxury hotels: Evidence from the analysis of consumers’ ‘big’ visual data on TripAdvisor,” *J Bus Res*, vol. 119, no. October 2019, pp. 495–501, 2020, doi: 10.1016/j.jbusres.2019.10.053.

[19] V. Kumar, D. Ramachandran, and B. Kumar, “Influence of new-age technologies on marketing: A research agenda,” *J Bus Res*, vol. 125, no. October 2019, pp. 864–877, 2021, doi: 10.1016/j.jbusres.2020.01.007.

[20] N. K. Tiwary, R. K. Kumar, S. Sarraf, P. Kumar, and N. P. Rana, “Impact assessment of social media usage in B2B marketing: A review of the literature and a way forward,” *J Bus Res*, vol. 131, no. March, pp. 121–139, 2021, doi: 10.1016/j.jbusres.2021.03.028.

[21] L. Columbus, “10 Ways Machine Learning Is Revolutionizing Supply Chain Management,” *Forbes*, 2018.

[22] “What is NumPy? — NumPy v1.21 Manual.” https://numpy.org/doc/stable/user/whatisnumpy.html (accessed Jul. 15, 2021).

[23] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[24] “SciPy API — SciPy v1.7.0 Manual.” https://docs.scipy.org/doc/scipy/reference/ (accessed Jul. 15, 2021).

[25] “Matplotlib: Python plotting — Matplotlib 3.4.2 documentation.” https://matplotlib.org/ (accessed Jul. 13, 2021).

[26] “seaborn: statistical data visualization — seaborn 0.11.1 documentation.” https://seaborn.pydata.org/ (accessed Jul. 13, 2021).

[27] “Classification: ROC Curve and AUC | Machine Learning Crash Course.” https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc (accessed Jul. 28, 2021).

[22] L. Asmelash, “Nearly 80% of hotel rooms in the US are empty, according to new data - CNN,” 2020, Accessed: Jul. 12, 2021. [Online]. Available: https://edition.cnn.com/2020/04/08/us/hotel-rooms-industry-coronavirus-trnd/index.html .

[23] L. Columbus, “10 Ways Machine Learning Is Revolutionizing Supply Chain Management,” *Forbes*, 2018, Accessed: Jul. 26, 2021. [Online]. Available: https://www.forbes.com/sites/louiscolumbus/2018/02/25/10-ways-machine-learning-is-revolutionizing-marketing/?sh=6c3fd8e95bb6 .

[24] “What is NumPy? — NumPy v1.21 Manual.” https://numpy.org/doc/stable/user/whatisnumpy.html (accessed Jul. 15, 2021).

[25] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011, Accessed: Jun. 03, 2021. [Online]. Available: http://scikit-learn.sourceforge.net .

[26] “SciPy API — SciPy v1.7.0 Manual.” https://docs.scipy.org/doc/scipy/reference (accessed Jul. 15, 2021).

[27] “Matplotlib: Python plotting — Matplotlib 3.4.2 documentation.” https://matplotlib.org (accessed Jul. 13, 2021).

[28] “seaborn: statistical data visualization — seaborn 0.11.1 documentation.” https://seaborn.pydata.org (accessed Jul. 13, 2021).

[29] V. singh, “Machine Learning K-Nearest Neighbors (KNN) Algorithm In Python - An Introduction.”

[30] “Classification: ROC Curve and AUC | Machine Learning Crash Course.” https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc (accessed Jul. 28, 2021).

APPENDIX A

ML PYTHON CODE

```
#import the libraries
#import the dataset
dataset = pd.read_csv('Ads.csv')
dataset.head(11)
dataset.info()
dataset.describe()
dataset.isnull()
val = dataset[dataset.Purchased == 1]
vis1= sns.lineplot(data=val, x= 'Age', y= 'EstimatedSalary')
val2 = dataset[dataset.Purchased == 0]
```

```
vis2= sns.lineplot(data=val2, x= 'Age', y= 'EstimatedSalary')
sns.lmplot( data= dataset, x= 'Age', y= 'EstimatedSalary', hue ='Purchased' , fit_reg=False)

# dividing the data into two parts from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(Independent_Variables, Dependent_Variable

#Scaling the Ad data
from sklearn.preprocessing import StandardScaler
std = StandardScaler()
X_train = std.fit_transform(X_train)
X_test = std.transform(X_test)

#Training the model
from sklearn.naive_bayes import GaussianNB
NB_classifier = GaussianNB()
NB_classifier.fit(X_train, y_train)

#Predicting a new result
print(NB_classifier.predict(std.transform([[45,11000]])))

#Predicting the test set results
y_pred = NB_classifier.predict(X_test)
print(np.concatenate((y_pred.reshape(len(y_pred),1), y_test.reshape(len(y_test),1)),1))

#Training the KNN on the training set
from sklearn.neighbors import KNeighborsClassifier
KNN_classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p = 2)
KNN_classifier.fit(X_train, y_train)

# Making Predictions
```

```
print(KNN_classifier.predict(std.transform([[45,11000]])))
y_pred = KNN_classifier.predict(X_test)
print(np.concatenate((y_pred.reshape(len(y_pred),1), y_test.reshape(len(y_test),1)),1))

#Training the SV Classifier
from sklearn.svm import SVC
SVM_classifier = SVC(kernel = 'rbf', random_state = 0)
SVM_classifier.fit(X_train, y_train)

#Predicting a new result
print(SVM_classifier.predict(std.transform([[45,11000]])))
y_pred = SVM_classifier.predict(X_test)
print(np.concatenate((y_pred.reshape(len(y_pred),1), y_test.reshape(len(y_test),1)),1))

from sklearn.metrics import confusion_matrix, accuracy_score
Confusion_Matrix = confusion_matrix(y_test, y_pred)
print(Confusion_Matrix)
accuracy_score(y_test, y_pred)

#Predicting a new results using SVM
print(SVM_classifier.predict(std.transform([[20,11000]])))
#Predicting a new result using NB
print(NB_classifier.predict(std.transform([[20,11000]])))
#Predicting a new result ysing KNN
print(KNN_classifier.predict(std.transform([[20,11000]])))
```