Lexicon Based Extraction of Twitter Sentiments and Comparison of Machine Learning Algorithms for Improved Classification of Polarities

Author

Samina Fazilat

00000118268

Supervisor

Dr. Muhammad Abbas

DEPARTMENT OF COMPUTER & SOFTWARE ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD. JANAURAY, 2020

Lexicon Based Extraction of Twitter Sentiments and Comparison of Machine Learning Algorithms for Improved Classification of Polarities

Author

Samina Fazilat

MS-16-CSE(CSE)

00000118268

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Software Engineering

Thesis Supervisor:

Dr. Muhammad Abbas

Thesis Supervisor's Signature:

____

DEPARTMENT OF COMPUTER & SOFTWARE ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY, ISLAMABAD

31st January, 2020

Copyright Statement

Declaration

I hereby declare that I have written this thesis titled as "**Lexicon Based Extraction of Twitter Sentiments and Comparison of Machine Learning Algorithms for Improved Classification of Polarities**" completely on the basis of my personal efforts under the sincere guidance of my supervisor **Dr. Muhammad Abbas**. All citations with references to all sources used in this thesis have been mentioned clearly and contents of this thesis have not been plagiarized. I certify that this work contains no material which has been accepted for the award of any degree, or in any university or any previously published material except where due references have been made in the text.

Signature of Student

Samina Fazilat

00000118268

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical and spelling mistakes. Thesis is also according to the format given by the university.

Signature of Student

Samina Fazilat

00000118268

Signature of Supervisor

Dr. Muhammad Abbas

Acknowledgements

I am eternally grateful to Allah Almighty; The Creator of All Universes and everything that has ever existed. I would never have been able to complete this insurmountable task without His guidance. I could really feel the divine help all along my thesis. Indeed it sounded an uphill task, but Allah was always there whenever I found myself in any blind alley.

I want to extend my deepest gratitude to my parents. They provided me the moral support and always encouraged me whenever I got stuck at any point. I want to thank my sister, she instilled in me the self-belief that with hardwork there will be no limits and I do possess the requisite talent to brave any hardship and adversity.

I want to thank my supervisor **Dr. Muhammad Abbas**. Their immense guidance played an instrumental role in shaping up my dissertation work. He taught me how to carry out research actively and effectively. Veritably, he is the architect of this work and I want to give him maximum credit for that. I used to bother him over and over again, usually numerous times in a matter of hour or so, but he always obliged. You have my utmost respect, sir.

Lastly, I want to pay my earnest gratitude to my friends who always boosted my morale whenever I was depressed. They are the true heroes and equally deserve all the recognition and praise. There are certain unsung heroes as well and I can't name them all but I do like to take this opportunity to send my warmest regards to whosoever played his part in making me what I'm today. I owe it to all of them.

ABSTRACT

Twitter is amongst globally used micro-blogging applications by millions of users on daily basis for sharing their thoughts regarding diverse topics of different occasions as well as their opinions on the hottest trends. Which makes it a rich source for decision making and sentiment analysis. Over the recent years, multiple frameworks and models have been proposed to extract people's sentiments against a topic, an individual or an organization with to help decision making process and still a lot of work is on the way to get accurate models. Sentiment analysis focuses on the polarity calculation from a tweet/text and classifying them as positive, neutral and negative. The primary focus of this methodology is to address the neutral tweets along with positive and negative tweets and their automated polarity generation. Here we proposed a system which extracts the tweets from Twitter against a profile and after applying preprocessing techniques, calculating sentiments and applying Random Forest, Naïve Bayes, Support Vector Machine, Multinomial Logistic Regression Classifier and XGBoost for prediction of sentiments from the tweets. The highest accuracy achieved with this methodology is 81% accuracy.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

POS	Parts of Speech
SVM	Support Vector Machine
KNN	K Nearest Neighbors
TSA	Twitter Sentiment Analysis
NLP	Natural Language Processing
NLTK	Natural Language Tool Kit (library)
TFIDF	Term Frequency and Inverse Document Frequency
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative

Chapter 1 Introduction

The world today has become a global village due to emergence of internet and various social media applications. Sending messages, expressing feelings and giving opinions has become very easy and affordable in no time. This vast range of expression has led to use of social media in variety of ways. Public figures, universities, journalists and politicians not only use it for entertainment purpose. They rather use it to share their ideas, important news and events to generate feedback from the public. Public as we know is very vocal and expressive now a days. Their opinionated approach and comments had always been an intriguing point for researchers and analysts.

People around the world generating 2.5 quintillion bytes of data everyday has created a rage and desire among researchers around to world to make use of this wide spread of data. The researchers have been applying various techniques to get insights and meaningful information out of this spread of which could lead to important and useful conclusions for future purposes. One of these techniques is called “Sentiment Analysis”.

Sentiment Analysis is a complex phenomenon. Analyzing human emotions using natural language processing is not as easy as it may sound. Human language complexity is not easy to be solved by the machines. Humans can be intuitive in understanding emotions, tone and context of a particular opinion but the same does not apply to machine learning. Suppose a person expressing in a sentence: “My flight’s been delayed. Brilliant!”. Now a normal person would consider this a sarcastic comment because this person could judge the tone and context behind the sentence but a machine by just breaking down the sentence into words and judging each word could not make sense of sarcasm possibly. It is not impossible to make a machine learn about the context behind an opinion but it is sort of difficult and challenging.

Even though human analysis and judgment is far off better than machine analysis and processing but humanly it is not possible to analyze the opinions of people around the world and make conclusions. Therefore, as a solution to this problem sentiment analysis plays a vital role. Till now there have been a range of researchers working in this field and looking into opinion mining. One of the major sources of acquiring opinions and digging into this is a social media platform known as Twitter. It is world's biggest platform where most renowned and famous people around the world give their opinions and express their thoughts on various issues. These opinions and expressions are often used to analyze the sentiment and hence many researchers have exploring the context of these tweets. These tweets have been categorized into 3 types: Positive, Negative and Neutral tweets. Researchers have until now focused keenly on positive and negative tweets and there has been some major work regarding this. I however tend to focus on neural tweets in this study. There has been a great impact on the study of positive and negative tweets but I intend to identify the importance of neural tweets and their impact in opinion mining and sentiment analysis. This study will demonstrate the methodology which uses the all polarity values (positive, negative, neutral) and perform sentiment analysis using ensemble classifiers which have provided good results.

1.1 Motivation

Today world is connected through the internet. The message can be sent anywhere around the globe at any time. However, private material or message sharing is not the sole aim of social media. Public figures, universities, journalists, companies and politicians use social media as a platform to share their ideas and generates feedback from the public. Opinions and moods of people are pressed out through internet as the social media users are rapidly rising. User's reaction on different areas like crisis, disasters, entertainment and news, etc. have pulled in an expanding level of intrigued within the researchers.

Monthly people do a multitude of publications and no polls are created out from public about certain matters because everything cannot be done manually. This plays up the fact that there is a need for automated intellectual methods which analyze the information of the text. These methods permit us to process large data in a short passage of time and message meanings are understandable. Usage of Modern methods and technologies in artificial intelligence and big data, helps the researchers to make the process automated in the analysis of content through data collection, preparation, management and visualization of data.

Existing work is done on sentiment analysis through the fields of computational linguistics, methods of machine learning, text mining, natural language processing and rule-based learning methods. The methods of machine learning encompass training the model. Of late, through social media numerous data are collected for sentiment analysis in research. Renowned techniques of machine learning are used for Data classification and data clustering. Nevertheless, the neutral tweets have not been acted many time and researchers mostly put neutral tweets data in their future work.

The primary intent was to address the neutral tweets and show their importance in sentiment analysis. The research shows the methodology which uses the all polarity values (positive, negative, neutral) and perform sentiment analysis. Likewise their automated polarity generation. There is one inclination that the method's performance is dependent on datasets and their complexity.

1.2 Problem statement

Much research has been done on sentiment analysis. The data which is in textual form requires the sentiment analysis process to address the sentiment polarity that either sentiments are positive, negative or neutral. For classification many researchers used different techniques of supervised learning, features based on polarity and Part of speech (POS). With domain restrictions different classifiers are proposed which are automated. But mostly researches focus on positive and negative polarity and they often skip neutral polarity. There is always a room available for new improvements. The major focus of this research is to focus on all three sentiment polarities positive negative and neutral and provide an automated polarity generation and mechanism for classifying polarities with enhanced accuracy for future tweets.

1.3 Scope

Previously there has been studies on classification of positive and negative tweets but however this study is to focus on advanced step i.e. that is classification of neutral tweets. This study will be classifying the positive, negative and neutral tweets.

1.4 Thesis Organization

The rest of the thesis is organized as following:

- **Chapter 2** gives detail about the procedure and use of sentiment analysis.
- **Chapter 3** is based on literature review.
- **Chapter 4** explains the methodology used for this study.
- **Chapter 5** explains the results obtained and discussion on them.
- **Chapter 6** discusses conclusion and future work and the end is supported by references.

Chapter 2 Sentiment Analysis

This chapter gives an overview of sentiment analysis, its procedure, its need and drawbacks related to it.

2.1 Social Media

Social media can be defined as an electronic communication channel through which people communicate with each other online and share their information, thoughts, feelings, messages, videos or files etc. Around the world it is an essential part of people's lives nowadays. Now it's easy to connect with people around the world because of social media platforms. It's completely based on computer technology. People connect with each other through Personal computers, laptops, tablets or smartphones.

Figure 1 The Social Media[1]

Social media becomes new norm of society as people communicate with each other through messages or via email and remain up to date with ongoing information. Social media has changed the past communication method and gathered people in a single platform easily. Not only communication but social media has become this huge platform where opinions are formed and each and every opinion can add value to a certain object or event. These opinions are further used by different companies, researchers and bloggers to get and overall review about certain events, products and situations occurring. Social media has given a voice to every individual and hence made it easy for researchers to get reviews and public opinions.

Figure 2 Applications of Social Media[2]

Social media can be divided into two terminologies such as social media and mobile social media. Social media connects people with each other through web-based internet network. Six classification categories are defined for social media and these are “joint projects (Wikipedia)”, “Microblogging (Twitter, Blogspot)”, “content communities (Youtube)”, “social networking sites (Facebook)”, “virtual game worlds (World of Warcraft)”, and “virtual social worlds (mIRC)”.

Figure 3 Mobile Social Media[2]

The main aim social media is to create your own content or extend or recreate others content and this has become a major trend these days. Mobile social media connects people through applications. Mobile social media is also classified into 4 categories: "space-timers (Foursquare)", "space-locator (Qype)", "quick-timers (posting in Twitter)", and "slow-timers (watching video on Youtube)". Both provides the mean of communication and to shared data with each other. Mobile social media has made it easier to voice up an individual's opinion as the applications used are more easily accessible. This has led to an increased volume of opinions raising the need of sentiment analysis.

2.2 Sentiment Analysis

Social Media has made people more opinionated and connected. These days' social media could be a part of standard of living. Individuals use it a minimum of once each day. This usage has provided in huge volumes of data which is needed to be explored and looked upon. Sentiment analysis extracts the subjective information from these text or sentences being studied. Sentiment analysis determines the attitudes of a person concerning some topics. The attitudes are often directed towards positivism, negativity or are often neutral contexts. The advanced sort of sentiment analysis determines the emotional states like happy, sad, and angry.

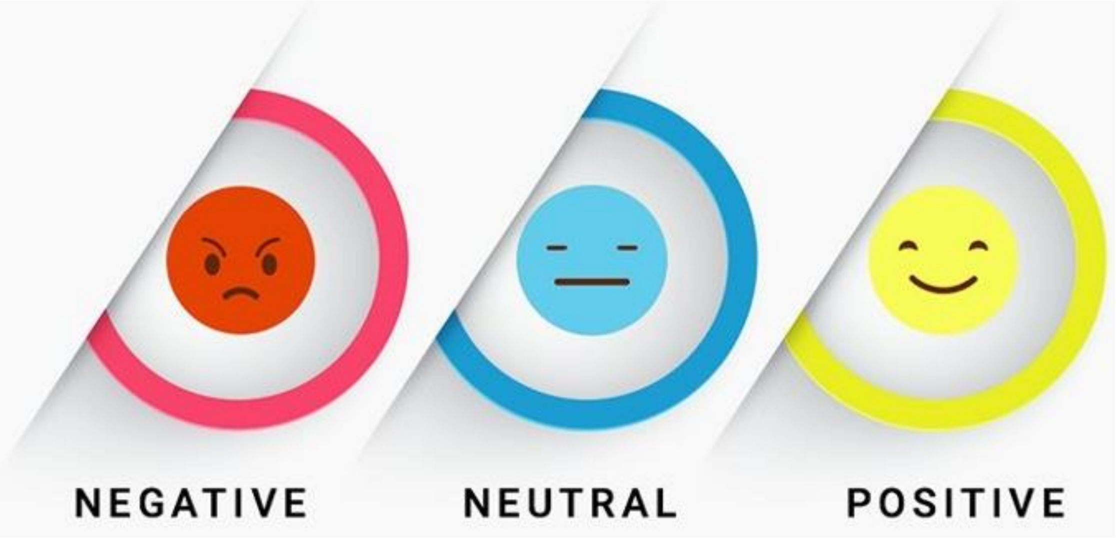

Figure 4 Sentiment Polarities[1]

The latest development within the sentiment analysis is especially targeted on towards the detection of polarity of product review and therefore the review of the films. It’s a really necessary tool within the field of computation mining for understanding the people’s plan and attitudes towards different problems, topics, events and response towards product or services being determined.

There are different enterprises which often want to enhance at intervals itself and its services for the shoppers and needs to urge some feedbacks from them in order that it will improve within the future. Therefore, social media analytics tools are employed by several huge firms like Bank of North American nation, Whirlpool etc. to act with customers and obtaining feedback from them. The Sentiment Analysis procedure helps to observe the supply, target and kind of perspective and helps to tell apart the perspective of texts.

The people or organizations are progressively using opinions from the media sources like personal blogs, review sites and social networks for his or her decision-making. However, current internet consists of many social media applications and every user with the various opinions. It’s forever difficult to accurately summarize the opinions and data from those media. We have as humans have a tendency to have totally different nature and opinion towards same product or services. Opinions dissent from individual-to-individual and from time to time on constant content.

As a personality's nature, we have a tendency to take our opinions quite seriously since it's our own and shut to our thoughts. Thus, it's forever a giant challenge to own consistent opinions on a product or a service. There ought to be some mechanism that automates and summarizes all those human opinions to urge unbiased and proper info. This can be achieved with the help of sentiment analysis.

In this chapter, the study is aim to manage basics of social media analytics and advance to the opinion mining drawback, additionally numerous analysis and challenges within the field of sentiment analysis and the key technical problems that require to be addressed.

2.3 Opinion Definition

Opinions are often given on something and therefore the given opinions are often summarized as having positive, negative or neutral attitudes towards the text or sentence into consideration. Opinion extraction and gathering information from various website in order to summarize them to get the right results is one among the massive challenges of sentiment analysis technique. There’s still a necessity to own clear and higher mechanism to mine these immense amounts of knowledge to urge the right results. Opinion-mining systems analyze the components of the text or sentence within the aspects of Who are the author, what the opinion is and how that half is especially expressed [2].

Objects and Opinions:

Objects are referred to as the target entity that is being discussed or used or talked about in a certain opinion. Associate degree object will have a group of elements and a few sets of attributes. The thing with elements and attributes is understood as feature of that object. Let’s say “I like Samsung galaxy III mini. It's a good bit screen”, the first part shows a positive opinion on Samsung phone whereas the second part shows a positive opinion about its screen that describes the feature of galaxy III mini [4].

Opinion Holder:

Opinion holder is a person who is actually categorizing the opinion concerning the entity being determined. Opinion holders are in most cases the author of certain post or organization that holds the opinion regarding a product or service [4].

Opinion and Orientation:

Opinions tend to have different dimension or orientations. An opinion or a feature can be positive, negative or neutral. These positive, negative and neutral views are known as opinion orientations. Opinions are often of two varieties specifically, regular opinion and comparative opinion [4]. These are mentioned below.

i. Direct Opinions: They refer to a type of opinion given explicitly regarding certain entity. It can be a positive or a negative one but it should be directly stating whatever the opinion in context to that entity would be.

ii. Comparative/ Indirect Opinions: These types of opinions are often expressed in as comparison of a target entity and another entity. The other entity is brought into comparison with the target entity to express opinions about in a better form.

2.4 Why Sentiment Analysis?

Sentiment analysis is becoming progressively necessary and attributable to emergence of social media. Sentiment analysis are often utilized in all styles of task. A number of them are within the field of public sentiment to understand client’s confidence, to understand what individuals deem the new product, within the field of politics to understand however individuals deem the candidate and their problems. Some firms use sentiment analysis for market research prediction and flicks industries use it to urge the reviews of the films to urge the feedback to understand whether or not the audience have positive, negative or neutral read.

Sentiment analysis determines the sound judgment, polarity and therefore the polarity strength of the content within the text or within the sentence. The polarity defines positivism or negativity whereas the polarity strength defines however the opinion is motivated like defile positive, negative or neutral. Sentiment analysis approaches are often categorized into keyword recognizing, applied math analysis, lexical affinity and conception level ways [3]. Earlier sentiment analysis was mainly targeted on product reviews and flicks reviews however currently it's targeted on excess of application starting from forum, social networks, blogs, product reviews and so on. The most basic objective of the sentiment analysis is to observe the information and verify the mental attitude of the author towards the subject to be studied.

The whole purpose of sentiment analysis is to determine the perspectives on the idea of holder, target of the perspective, attitude of perspective from set of types like love, hate or just from polarity like positive, negative or neutral. This will be finished with the assistance of some text or typically we would contemplate whole documents.

2.5 Tasks in Sentiment Analysis:

Following are the tasks in sentiment analysis.

Classification of Subjectivity

This methodology helps in identification of subjectivity or objectivity of a sentences. The sentence can be split into these two categories. This classification can be done with the help of the context in which the sentence has been used. Identifying objectivity and removing objective sentences before polarity often helps in improving the performance of the process. [7]

Sentiment Classification

Sentiment classification identifies the nature of the sentence if it shows positivity or negativity. It can be used to identify two categories i.e. positive or negative, or it can be used to identify multiple categories like extremely negative, negative, neutral, positive, extremely positive and so on. Major work of this study however focuses on the classification of positive, negative and neutral tweets so we will be dealing with three categories in this case.

Opinion Holder Classification

The sentiment analysis process is can be used to classify the opinion holders who are a vital part of this whole process. Detecting opinion holders often aids in accessing the sources of these opinions. This procedure of classifying opinion holders can be very helpful where same opinion holders have multiple categorical opinions. In these scenarios, opinion holders are mostly referred by their names and login credentials.

Object/ Feature Extraction

The main things whereas analyzing sentiment is to work out target entity. The opinion in blogs, social media and in review sites have specific intention towards topic therefore to search out the target entity is critical in such situations to extract the options. A reviewer will have totally different opinions concerning the options and elements of the target entity thus feature based mostly analysis are necessary problems in sentiment analysis [11].

2.6 Levels of Sentiment Analysis

Sentiment Analysis has 4 different levels as shown in below figure:

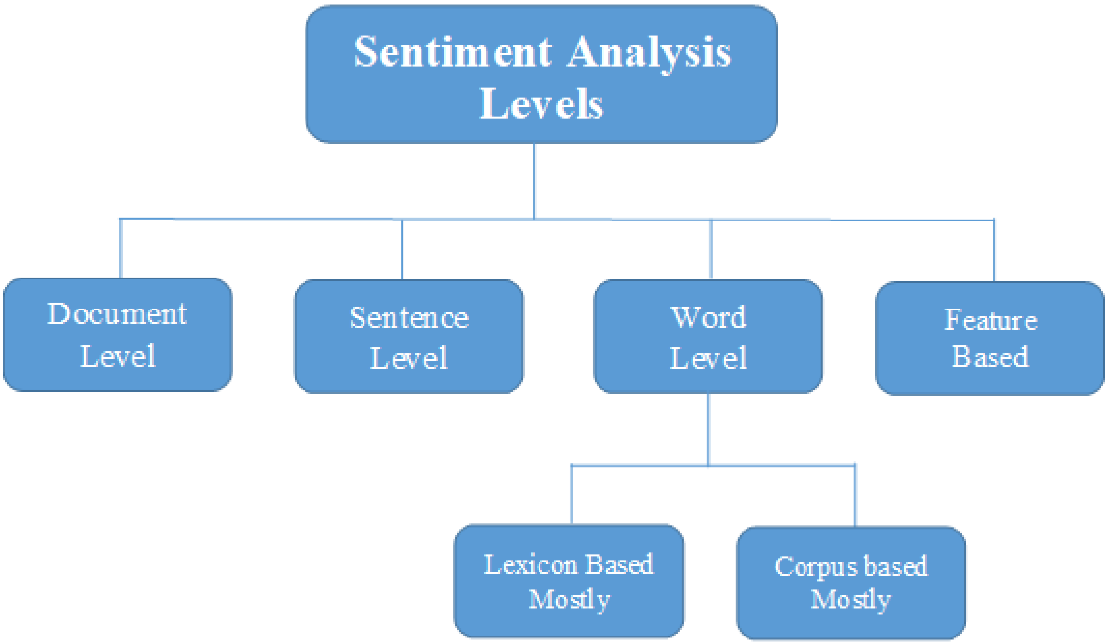

Figure 5: Levels of Sentiment Analysis

Document Level:

The document whose sentiment must be determined is taken into account as a basic unit for sentiment analysis purpose. This approach assumes that single opinion holder holds the opinion. The positive and negative reviews are often classified by mistreatment numerous obtainable machines learning approach.in [9] 3 classifiers (Naive Thomas Bayes, most entropy, and support vector machines) and options like unigrams, bigrams, term frequency, term presence and position, and parts-of-speech are experimented. They need terminated that SVM classifier works best which unigram presence info was simplest [9]. Pang and Lee developed Document level sentiment analysis as a regression drawback [9]. Supervised learning was accustomed predict rating scores.

Sentence Level:

This methodology relies on distinctive the subjective sentence from the mixture of sentences. The most drawback with document level analysis is that it will extract info from objective sentence therefore sentence level sentiment analysis is required for subjective analysis. The supervised learning methodology is employed to spot the subjective sentence.

Word Level:

It uses largely adjectives as options. Scientist conjointly uses some verbs, nouns and adverbs as options [7, 12]. The 2 ways of mechanically annotation sentiment at the word level are:

i. Lexicon Based: This approach relies on the list of words with previous polarity. During this methodology, an inventory of word is made and is extended with synonyms and antonyms mistreatment on-line lexicon. The sentiment of the word is set by however the unseen words interacted with the antecedently outlined words within the list. The positive and negative sentiments of the words and therefore the orientation of words is calculated with the assistance lexical relation. A word's linguistic orientation can by calculated by checking the relative distance from the 2 seed terms i.e. smart and dangerous. As the values lie among the ranges of [-1,1], this indicates strength of the orientation calculated till now [16]. There is a slight drawback with this methodology. This isn't specifically domain specific so this cause problems in polarity classification.

ii. Corpus Based: Corpus based mostly ways relies on legendary polarity and depends on syntactical and applied math techniques. There exists a relationship between certain unknown words and manually selected seeds. This association is classified as a positive or negative depending on the words and context in a particular situation. There is a certain degree which is achieved from this associated relationship representing it which is further used as a result [13].

Feature Based:

Document based approach and sentence level approaches are not really different from each other, however individually they consider positive and negative reviews, just like reviewers likes some options and dislike others. Due to this reason, feature based analysis is of great importance. Feature based technique extracts the feature of the product and then performs analysis on it. Yi et al. [12] wrote a paper and his approach was again the feature-based analysis. He took out noun phrases that were individual or came before a particular article. He also focused on extracting a definite base phrase that came at a start of sentence and if it was followed by predicate. For each sentiment phrase that was detected, there was a target set and the polarity was such that it finally supported the sentiment pattern information. Hu and Lui [7] used a heuristic methodology in which they had to extract the frequent noun or phrase and did association mining on it. They assigned the closest possible opinioned word to the selected feature which would then give its sentiment orientation. Etzioni [6] has improved this task with great efforts. The algorithm made by this person helped in elimination of nouns and phrases which were not related to a particular product, hence improving the technique.

2.7 Recent Trends:

The recent trend shows that the automatic methodology is the most generally used techniques for analyzing the opinions. It's supported the phenomena like linguistic communication process, Text Mining, Machine Learning and Computing, Automatic Content Analysis, and balloting Advise Applications. The rise and use of social media results in increased data within the amount of unstructured knowledge. We should be grateful that social media platforms do not require or apply any sort of restriction for application of Machine Learning algorithms. This however has resulted in mixture of increased volume of knowledge which is obtainable and a lot of complicated ideas to research. In recent years there has been a decrease in interest on semantic-based application and increase in use of statistics and visualization. Even as the other science, automatic content analysis is additionally turning into a data-intensive science. There are still number of challenges being faced on these social media platforms and they are listed as below:

- Detecting reliable content has become crucial and sentiment analysis can help in identifying same of fake reviews.
- Recognition of Name Entity has become important to identify the opinions of people.
- Limitation in filtering has allowed people to give unsuitable opinions which mostly results in wrong kind of sentiment being analyzed.
- Sentiment Analysis requires domain independence so that if words expressing a certain type of sentiment in one domain, they should be the same for the other domain.
- Opinion Orientation is very crucial in sentiment analysis because it helps in identifying the context of the opinion by comparing different words and checking their polarity if it becomes positive or negative.
- Language is of great importance in this case of sentiment analysis. Sometimes a sarcastic comment using positive words might give a false idea about the context.

2.8 Sentiment Analysis Currently:

Sentiment analysis are often used from on-line retail to blogging and conjointly in numerous applications in politics. Today clients connected with business use sentiment analysis not only for product review however conjointly for client services and whole name management. Equally sentiment analysis is beneficial in obtaining the feedback from subject and which therefore results in campaign messages and policy announcements by political parties.

Ways of determining Sentiments:

The growing use of the net media makes it troublesome for sentiment analysis. There is a scaling system methodology which uses the scoring system to work out the non-public appreciation. As we can see on most film review sites and other product reviews, the bloggers or writers have expressed about a particular film, product or a service in a form of score. So, this helps in determination of sentiments about a particular thing.

Subjectivity and Objectivity Identification:

This methodology relies on some kind of documentation. The text or the sentence will have subjective or objective opinion. Differentiating subjective and objective opinion from the text or from the sentence is troublesome task as different words can be used in different context and in different situations and scenarios. Therefore, this is often one among the troublesome ways to implement [15].

Bales Interaction Method Analysis:

This methodology identifies and records the character of every interaction. It's not accustomed to live content of the interaction. Bales Interaction Method is based on a scoring system in which scores are applied to set of classes. These units are generally created from one sentence expressing one plan. There are certain complicated sentences which are scores according the quantity of freelance clauses they have. Sentences often have separate fragments which combine to form a sentence. These fragments are scored such that one point is given to each but separately fragments get complicated to interpret. A better approach would be to take the fragments in the context in which they have been used. Just like when we are studying oral or physical studies and judging a person's sigh or grunt through their facial expressions [14].

Chapter 3 Literature Review

Twitter Sentiment Analysis or as we say TSA is a very keenly researched area. Research work in the sentiment analysis deals with the opinions and perspectives of oneself with the attitudes and emotions. These emotions and attitudes are closely linked with natural language with respect to an event. Current advancements in sentiment analysis assay the achievements incurred, which can not only be entail as positive and negative results, but also embellish the domain of emotional and behavioral analysis of varied topics and communities. An enormous amount of research for predicting social opinions is also carried out examining the several techniques in sentiment analysis.

3.1. Overview

Authors in different papers and studies has approached to the problems of sentiment analysis in different ways. Some of them have focused more on preprocessing techniques while others tend to focus more on the overall methodology and presented results. We however will be giving a brief review on preprocessing techniques and they give review about the overall methodologies used by different authors in various studies.

Pre-processing

Pre-processing is a very important part of sentiment analysis. Selecting appropriate methods for pre-processing can lead to better accuracy results.

Akrivi Krouska [19] provided a study in which he explained the effect of pre-processing techniques and their selection twitter sentiment analysis. According to Akrivi, it is extremely important to properly select methodologies for pre-processing and perform this task appropriately because it can directly affect or enhance the results. Sentiment analysis requires the use of classification mechanism to complete the process and this process often requires text pre-processing and feature extraction which leads to further classification of positive and negative tweets. The authors in the study performed an experiment on three different types of data sets. One of them was not specifically from any domain while the other two represented some particular domain. Four classifiers namely Naïve Bayes, SVM, KNN and C4.5(Decision Tree) were used in this study to work upon TSA. In terms of preprocessing techniques, unigram, and 1-to-3 grams performed better than any other techniques and using feature extraction with them improved accuracy results.

Another study conducted by Dimitrios Effrosynidis also discussed about pre-processing techniques and their results [20]. Dimitrios took 2 different data sets and applied 15 different pre-processing techniques on it to get certain results. After processing different machine learning algorithms like Naïve Bayes, Logistic Regression and Linear SVC were applied to finalize the best of pre-processing techniques. So, after applying sentiment on twitter-based data, there were some techniques like stemming, replacement of repeating punctuation and removal of numbers resulted in better performance. While some other techniques like handling capitalized words, replacement of slangs and some other techniques resulted in low performances and accuracies. Hence the fore mentioned techniques resulting in good accuracies and better performances were recommended by the authors. The authors have also mentioned to work with the combination of some other classification algorithms to improve results. Going into the details of the results provided, SS-Twitter data set achieved 61.4% with Linear SVC through the pre-processing technique of replacement of long words. However, Naïve Bayes came up with the accuracy of 60.6% after stemming technique and 61 % was achieved through Logistic Regression when lowercases were used. For the other dataset naming SemEval, Linear SVC resulted 59% of accuracy, Naïve Bayes resulted in 60.6% and Logistic Regression resulted in 60.7%. These scores were the highest of all achieved after replacing URLs and user mentions which is one of the pre-processing techniques.

Giving opinions and has evolved over time and now increasing trends of emoticons have completely changed it. Emoticons are really common and Katarzyna Wegrzyn presented a study which was based on pre-processing techniques used with sentiment analysis [21]. Katarzyna compares 3 techniques involved in pre-processing of emoticons and emoticon weight lexicon method using tokenizer of Twitter and NB model. The used alpha factor for integration of lexicon method with a classifier. This combined strategy showed results which proved that NB Model was actually better for twitter sentiment analysis. The authors managed to achieve results with accuracy of up to 80%.

Methodology:

The section of literature review covers the summary of the studies and researches based on the over-all methodologies of Twitter Sentiment Analysis or shortly termed as TSA. These studies are the base to build up our own methodology techniques.

The first study which came across was based on NLP and Machine Learning approaches which were used to classify various tweets as being positive, negative or neutral. Hamed AL-Rubaiee performed sentiment analysis on product reviews of Mubasher Products [22]. The author used SVM and Bayesian methods for classification and different experiments were carried out after splitting of data into two different subsets showing independence in terms of their time periods. So, the opinion words given in Arabic were translated into English and further used for Sentiment Analysis. The authors performed different sets of experiments on the data sets and got a good accuracy for Naïve Bayes with N-Gram when Neutral tweets were also being classified. The accuracy was up to 75%.

Aliza Sarlan, in her study has shared the results of a protype built to extract results of sentiment analysis by combining NLP and Machine Learning techniques [23]. The authors have developed a prototype which can classify tweets into positive and negative classes showing perspective of different customers. The results are shown in the form of a pie chart on an html page. The study also claimed to have a plan of developing a web application for this purpose but there were certain limitations to Django frame work due to which it did not proceed but however a major part of future works was. The main lack was in terms of classification of tweets. According to the results shared, out a full data set 9.4% of tweets were classified as positive tweets and 6.5% were classified as negative tweets but the remaining 84.1% was classified as null. So, a major chunk of data was taken as null. This is where I feel lies the importance of classification of neural tweets which we will be covering in this study.

Priyanka Tiyagi in her research has done a review of twitter sentiment analysis techniques and presented us the results [24]. The authors have mentioned use of Tweepy for extraction of data from twitter. There were different machine learning algorithms which were used for classification of positive, negative or neutral sentiments. The algorithms were used were SVM, KNN and Naïve Bayes and implemented using python. The process of sentiment analysis comprised of 4 steps namely pre-processing, tokenization, feature extraction and classification. Tokenization was done by finding out the polarity in the existing data set. For classification the fore mentioned algorithms were used. As per the study of this paper SVM gave the accuracy of 80% but this can be improved if a combined model of KNN and Naïve Bayes is applied. This task however is the future work of this study.

Ali Hasan and Sana Moin have presented an interesting approach in their study for TSA [25]. Their study is focused on analysis of exiting sentiment analysis tools with the help of machine learning algorithms. Application of machine learning algorithms help in getting the accuracy rate of sentiments related to elections. The authors tend to focus on lexicon -based sentiment analysis in which words, phrases or sentences in a document are used to find out semantic orientation. Next is the polarity which can be found out by using the dictionary having the semantic scores of specific words. The authors came up with the idea to compare sentiment lexicons i.e. W-WSD, SentiWordNet and TextBlob , by using machine learning. The validation of their sentiment analysis was done by a hybrid approach in which SVM and Naïve Bayes were used. Results showed that W-WSD worked the best in all cases. The accuracy achieved with SVM was 79% and with Naïve Bayes, it was 62% which concluded that W-WSD worked the best in cases.

Another study was conducted to get the sentiment analysis of reviews related to an airline service [26]. The authors Yun Wao and Dr. Dr. Qigang Gao have claimed this research to be one of the empirical contributions in the field of sentiment analysis. The authors have used an ensemble approach in comparison with the other classification methods and proved to have gotten better accuracy results. Their improved results were such that the high accuracy could have been used for the implementation of investigation customer satisfaction. Not only specific to airline services but this also helps in analysis of other twitter data sets. The accuracy of two class data set and three class data set was measured in terms of F1-Measure. The final accuracy for two class data set in case of ensemble model was 91.7% and in case three class data set was 84.2 %. These results were better and improved than other classification models used previously.

In [27] authors discovered that if he could use a pipeline of different classifiers which could be used as a hybrid model for classification of tweets into positive, negative or neutral classes. This scheme was helpful according to Khan in improving the accuracy of the model. The study dealt with the pre-processing, emoticons and SWN based sentiment scoring. Once the data was pre-processed properly then it was used for classification using the models EC and SWN which were classifiers selected for this study. The results were better than the previous models but there is a great scope for improvement if they could also involve certain domain specific words, some emoticons and slangs then perhaps the accuracy could improve.

In [28] authors proposed an approach based around lexicons which was used to combine different lexicons and dictionaries which could further be used for sentiment analysis. Their main target was to classify the tweets using slangs. Their study proposed a method having three different modules. The first module was related to tweet capturing and filtering. Second module was based on subjectivity detection and third and final module was sentiment scoring. These modules used lexicons i.e. opinion lexicon, word net, SWN and emoticon repositories. For binary classification the accuracy that was achieved was up to 92% and for multi-class classification the accuracy was up to 85%. But in terms of recall for a negative class, system could have improvement.

There was another study on classification of sentiments [29]. They used lexicon-based approach in which he extracted information using pre-processing techniques for classification of user sentiments from on line communities. Some famous lexicons like SWN and user defined dictionaries were used for this purpose. The basic classifiers used for this purpose were, i) SWN-based classifier, ii) a modifier and negation classifier and iii) a domain specific classifier. Since there was an issue of domain dependency, it was solved through these classifiers. The limitations in this study was that the approach was not handling compound sentences, clangs and sarcastic sentences.

A study is proposed in which he collected tweets on the basis of different words[30]. These words were further processed according to the regular expressions and machine learning algorithms so that feature selection and classification can be applied further. This classification was helpful in monitoring the public concern and disease information in Portugal and Spain. The achieved F-measure values were quite better than the baseline methods. The values could be found the range of 0.8 to 0.9. Again, the problem with this system was that the writers did not involve the support for using emoticons, slang and domain specific words.

In [31] they proposed a technique which was based on unsupervised learning. This was a comparatively a new idea considering the previous techniques as previously there was reliance on classification procedures and models. First of all, sentiment scor5es were computed on word level using SWN and a ransom walk technique, which was used to analyze the weights of the tweets. Again, the proposed methodology was far better than the baseline methods since it had no dependency on the labelled data set. But, the set of limitations revolved around negation handling and manual annotation of tweets. Also, there was certain inconsistency found in the final sentiment score.

Internet slang was the major problem which was not being solved by the authors previously[32]. They worked on this problem and found out a solution for it in 2014. The authors presented a framework which based on lexicon techniques used to find out slangs in the tweets based on certain polarity score. Th major limitations in this proposed study were not enough study on the emoticons and the modules working on spelling correction and context aware can be made more sophisticated. Again, domain specific words were not focused.

Author Aleksovski along with his fellow authors Calarelli, Grčar, & Mozetič [33] worked upon investigation of relationship between stock returns and Twitter platform. As per their study, Twitter had a huge impact on stock returns. So, to analyze the effect of Twitter platform, twitter text was analyzed based on an "event study". This was as economic technique which was used to automatically identify the events based on Twitter text. This further helped in analyzing the sentiments and classifying them into positive or negative class. Once the sentiments were classified then study was used in the end to find the relationship between Tweets and stock returns.

In [34] proposed a study which was based on an approach consisting of four modules named as i) Data Collection, ii) noise Reduction, iii) lexicon generation iv) sentiment classification. The modules were based on four different algorithms which provided better results then the existing possible solutions. The limitation for this study was based on the size of dataset used for the approach proposed. As for this study, dataset of iphone6 was used but the scale of the dataset can enhance to improve the effect. Tweet scrapping can be effective technique for acquiring data.

In [35] Sentiment detection of twitter messages with the helpfulness of linguistic features have been inspected by them. They assessed the value of existing lexical assets and also includes that catch data about the casual and inventive dialects utilized as a part of microblogging. They adopt a directed strategy to the issue, yet use existing hashtags in the twitter information for building training information.

In [36] utilize worldly Latent Dirichlet Allocation (LDA) to break down and approve the relationship between points removed from tweets and related occasions. They built up the term co-occurrence recovery procedure to follow sequentially co-occurring terms and accordingly make up for LDAs restrictions. At last, creators distinguish topical knowledge.

In [37] They stated "For each extracted entity (e.g. iPhone) from tweets, we add its semantic concept (e.g. "Apple Product) as an additional feature, and measure the correlation of the representative concept with negative/positive statements."In [38] utilized Sentiment Diffusion Forecasting by utilizing dataset of 23 million tweets from more than 16 million Twitter users. Researchers get the Sentiment Scores of tweets by adjustment of the AVA (Adjective-VerbAdverb) algorithm of sentiment analysis.

In [39] examine whether it is conceivable to foresee varieties in vote expectation in light of sentiment time arrangement extricated from news remarks, utilizing three Brazilian decisions as contextual investigation. They Mainly accentuate on a way to deal with anticipate surveys vote goal varieties that is satisfactory for situations of meager data. Researcher created trials to evaluate the impact on the gauging precision of the proposed highlights, and their separate arrangement.

In literature, numerous methods of Twitter Sentiment Analysis are available which use base classifiers [47]. A survey on Sentiment Analysis algorithms and the application has been presented in [40]. Through emoticons and hashtag sentiments of tweets was determined [40] [41]. Many linguistic resources are used with lexicon and learning and lexicon-based techniques like POS to gets sentiments [42] [43] [44] [45]. Through dataset, feature set and bootstrap an ensemble learning technique is introduced which accurately classify the sentiments as compared to many classifiers. [11]. another ensemble classifier approach is introduced which is trained on features for sentiment calculations [46].

From twitter different student's data were gathered to classify their different problems they face on daily basis [48]. A hybrid approach is used to where machine learning algorithms are merged with sentiment lexicons to identify the polarity of tweets in the business domain to check the consumer reviews on products [49]. To check the performance of different base classifiers like Random Forest. SVM, etc. the authors performed detailed analysis [50]. Not just classify the tweets, but also removes the anomalies from tweets an enhanced sentiment classifier is developed [51].

Knowledge based rules are not defined in black box classification [56]. Examples of such classifications are Naïve NB, Artificial Neural Networks, K nearest neighbor, Support Vector Machine etc. [52][53][54][55].Automated sentiment classifier is proposed which classify sentiments into positive and negative for TV shows in Brazil and they achieved higher accuracy which reached around 90% [57]. With knowledge domain the loss reduction is higher, so the author proposed a technique which is based on domain oriented technique for reduction of information loss [58].

The study was based on NLP and Machine Learning approaches which were used to classify various tweets as being positive, negative or neutral. Hamed AL-Rubaiee performed sentiment analysis on product reviews of Mubasher Products [59]. The author used SVM and Bayesian methods for classification and different experiments were carried out after splitting of data into two different subsets showing independence in terms of their time periods. So, the opinion words given in Arabic were translated into English and further used for Sentiment Analysis. The authors performed different sets of experiments on the data sets and got a good accuracy for Naïve Bayes with N-Gram when Neutral tweets were also being classified. The accuracy was up to 75%.

Twitter also has a solid impact on stock exchange and stock revenues. To go more in depth and investigate with twitter relationship with stock market the author in [60] goes with in depth study. The “event study” financial technique was implemented for twitter text analysis to identify the events automatically based on tweets volume load. Analysis of positive and negative sentiments were conveyed with the help of it during loads. At long last, to distinguish the connection amongst tweets and stock values "occasion contemplate" was connected. The principle restriction of this was the less modules of emotion and slang, because of this issue the results were not as accurate as expected.

To perform twitter sentiment analysis, a unified method was proposed in [61]. The method is divided into 4 modules “data collection”, “noise reduction”, “lexicon generation” and “sentiment classification”. Four new algorithms were created for the implementation of these four modules. “iPhone 6” dataset is used to get experimental results. As compare with other similar methodologies this methodology is considered more powerful. Be that as it may, to accomplish more adequacy, additionally tries are required on bigger datasets with tweet scratching in gushing mode.

In [62], different sentiment polarities were compared on a comprehensive level for text of twitter tweets based on different classification methods. In addition, for the purpose of methods evaluation multiple classifier are combined to compare them with individual results .Also evaluation was done on the collected tweets and the also use manually marked tweets. The authors considerate as a major contribution in sentiment analysis with respect to emoticons detection. As emoticon detection not always expressed the overall sentiments, particularly in non-sentiment ("neutral") case in the text.

Tan et al. [63] said that the users that shared similar opinions are likely to be connected. The authors proposed the model that were generated from either by following the network that has been made by tagging different user with the help of "@" or by analyzing the network of twitter follower/followee. The authors explained that by employing information of link of twitter there will be improvement in user-level sentiment analysis.

A neuron-based feed-forward BPN network proposed by Chen et al. [64], a methodology based on neural network, which uses orientation of sentiment to calculate the results at each neuron. The proposed methodology is combination of machine learning classifiers and semantic orientation indexes. In order to obtain efficiency in methodology, semantic orientation indexes used as inputs for neural network. The proposed methodology outperforms other neural networks and traditional approaches by increasing efficiency in both training as well as classification time.

Supervised machine learning techniques in combination with the artificial neural networks was employed by Malhar and Ram [65] on the twitter data for the tweet classification. The case study encores the results of SVM with all other classifiers and used and analyzed data of presidential and assembly elections. The authors proposed a methodology to predict the outcome of election results by utilizing the user influence factor. To carry out reduction in dimension the authors combined the Principle Component Analysis with SVM.

A hybrid feature, which is the combination of unigram, bigram and jointly, was reviewed by Anton and Andrey [66] while exploring the existing methodologies. The authors developed a paradigm for twitter message to analyses sentiments automatically. This work outlays the techniques and methodologies for analysis of social media message through accent categorization. This study incorporated the work done in the existing literature for automatic sentiment analysis. The work also includes a study of character feature predicated in statements made on social media to immerse itself with the context of current methods.

Document analysis-based technique for building a sentiment classifier was studied by Pak and Paroubek [67]. The method uses the linguistic analysis and determinates the sentiments as positive, neutral and negative, respectively, in a document. The process presented for collection of corpuses is automatic which trains the classifiers for the sentiments. The use TreeTagger in the study accommodates to assay the disparity between the positive, neutral and negative results.

It is very pertinent to use neutral messages for the purpose of experiencing knowledge about the polarity as explained by Kopel and Schler [68]. The study emphasizes that only examining the positive and negative message will not have a significant impact on prediction of neutral messages. Consequently, it is cardinal to speculate about the neutral messages by clarifying the difference between positive and negative messages. The study investigates and put emphasizes on having one of the corpus containing neutral documents resulting with no sentiment, which, can be used as indicator or counter to test positivity or negativity of any document.

Go et al. [69] proposed a technique to analyze twitter message through an automatic sentiment classifier, which classified the messages as positive or negative based on the respective query term. The authors used the machine learning algorithms for the sentiment analysis of twitter posts and results are displayed using the process of distant supervisions. The employed technique employed the SVM, Naïve Bayes and Maximum Entropy methods for data training. The used data sets also contained emoticons and the technique resulted in the accuracy of about 80%. The authors credit the steps involved in preprocessing of the data for the improved efficacy.

Overall, the study incorporated the idea of data containing emoticons and involved the methods of distant supervised learning.

Another study by Christianini and Taylor [70] hind-sighted the knowledge of machine learning algorithms, i.e. SVM. The study presented a deep overview about the SVM for the readers to garner the understanding about algorithm at implementation level for solving pragmatic scenarios.

Burger et al. [71] explored the computation handling capacity of computing devices in large scale application environments. Application of pattern recognition and statistical estimation lie in the category of large scale; however, modern day computers are powerful enough to handle such applications. Maximum entropy principle, which, described in detail, is employed for statistical modelling in the study. The authors used the methods of maximum likelihood using example problems in natural language processing to automatically construct maximum entropy models. This process chooses the best model having highest entropy from a set of consistent models. The maximum likelihood parameter serves as defining criteria to obtain optimal indicators of specified constraints on a given training data.

Romero et al. [72] studied the social systems such as twitter and identified the most common features of the data sets i.e. hashtags. The data signifies the use of new terms regularly affecting the meaning of original ones. The study highlighted the structural disparities among the data sets containing hashtags of irregular types. The study presented simulation-based models, although quite generative by self, for the study of hashtags expansions and adoption dynamics. The hashtags expand and interact with the latest adopters and adoption dynamics.

Text mining endeavors sentiment analysis as a prosperous domain in concatenation with emotional polarity computation as presented by Li and Wu [73]. The authors study the hotspot detection and forecast method using sentiment analysis and text mining. Emotional polarity of a message is descripted through an algorithm which assigns a value for each word of message. The work is embedded with K-Means clustering and SVM to categorize the method as unsupervised text mining. The results of study for four top hotspots of the year using K-Means and SVM are depicted with experimentation.

Until this date a minimal of researches carried for the Chinese documents on sentiment analysis as described by Tan and Zhang [74]. The conducted study proposes a method to categorize the sentiments using the data sets of 1021 Chinese documents by embodying the various selection methods. The authors used the varying selection methods, such as CHI, document frequency, mutual information and information gain along with machine learning methods such as SVM, Winnow classifier, Naïve Bayes, Centroid classifier and KNN classifier. The results depict that the information gain outperforms others selection methods for sentiment terms while SVM performs exceptionally well on the classification of sentiments.

Another method called Delta TFIDF is proposed by Martineau and Finin [75] to gauge word scores effectively prior to classification. The method is easy on understanding and computation for sentiment classification. The study used the data sets of movie reviews for classification through SVM for more accurate results. Hence, authors concluded the supremacy of Delta TFIDF over TDFIF feature. The study employed the method to measure sentiment polarity and subjectivity detection by counting the term raw for all sizes of documents. Delta TFIDF is described as first measuring technique before employing classification. This technique boosts and highlights the effectiveness of choosing selective words using calculated supervised distribution.

A study by Nielson [76] for sentiment analysis adopted a method to analyze messages by assigning scores to the effective words occurring from a labeled word list. The author revealed a technique, ANEW (Affective Norm for English Words), for effective term listing exclusively for micro blogs. This was prior to the incumbent use of sentiment analysis and micro blogging. The ANEW method is best utilized for detecting sentiment strength from a micro blog in comparison with another list of words. The method was tested in comparative manner using the data sets of messages posted on Twitter.

Two SVM classifiers were descripted by Mohammad et al. [77] in sentiment analysis study. The presented method bifurcated the technique into the term level and message level. Word sentiment in a message is determined by term level module while the message level module explores the sentiments of SMS or tweet messages. The authors took part in a competition where 44 teams came in their submissions stood fist in work on tweets, getting 88.93 F-core in term-level task and 69.02 F-score in message-level task. The model executed the features at sentiment and semantic levels while addressing the two term-sentiment associations; i.e. emoticons and hashtags.

Kouloumpis et al. [78] analyzed the Twitter messages while taking benefit of employing semantic features. The study examined the informal and intuitive language along with the benefits of existing lexical resources while investigating the knowledge collecting features in microblogging. The hashtags data is used and fed to supervised learning methods to identify the solutions to the problem at hand. The results suggest that that the feature of part-of-speech in experimentation did not performed as per expectation in the domain of microblogging. The harbingers from results also conclude that data sets containing hashtags are essential to classify emoticons, present in messages, as positive or negative.

Denecke [79] elucidate the multilingual sentiment analysis framework in decision for work polarity. The idea of sentiment analysis in this approach was originated on the compelling use of lexical English resources. The approach; classified as a feasible approach by authors for sentiment analysis in multi-lingual frameworks; first translates the input data sets from original language the English language through a benchmarked software for translation. The next step involves the process of classification of documents into positive or negative categories based on present adjectives in the documents for the sentiment analysis.

Gokulkrishnan et al. [80] proposed a methodology for the preprocessing of publically generated tweets from twitter online microblogging site and on the basis of their opinion content of irrelevant, negative or positive sentiment classified can be done; and investigating the performance different classifying methods based on precision and recall. The authors handled the skewness of the datasets by exclusively new approach called SMOTE oversampling method which helped by increases the accuracy of the classifier. The authors also explained limitations and applications of the research. Random Forest, SVM and SMO generates better performance compare to Naïve Bayesian classifier.

Neri et al. [81] performed sentiment analysis on newscast over more than 1000 Facebook posts and then compared the sentiment for dynamic company La7 and Rai – the Italian social broadcasting company which is emerging company. The authors experiment done by Knowledge Mining System which is used by security related agencies and institution of government in Italy to control information contained Web Mining and OSINT. The authors observations were mapped with the study conducted by the Italian research institute highly specialized in study of media at empirical and theoretical level, occupied in the study of communication of politics in the mass media known as Osservatorio di Pavia.

Wilson et al. [82] said that the methodologies for automatic sentiment analysis start with a big set of terms noted with their respective polarity. The main purpose of this study is to easily differentiate between contextual and prior polarity, with prior knowledge of understanding which are the necessary features for this task. The experiment covers the feature performance for multiple algorithms of machine learning. Except one algorithm, features when combined together gives the best results. The evaluations shows that when natural instances are present the performance of features degraded on great pace. The authors suggested that indicating features that described more complex interdependencies between polarity clues can be considered as future research work.

Godbole et al. [83] proposed a system which contains phase of identification sentiment in which for a particular topic which displays some opinions and scoring phase and a sentiment aggregation that will scores relative entities in the same class. At last the authors investigates the importance of methods for scoring involving huge blogs and news dataset. The authors interested in the fact that sentiments can vary according to the geographic location, news source or demographic group. As future work the authors are studying in evaluating the extent to which we predict the changes of future in behavior of market or popularity.

Benamara et al. [84] explored the strength of subjective statements within a document or expressions which uses the special part of speech such as nouns, verbs and adjectives. The authors said that until their contribution there was not a single related to adverbs in sentiment analysis nor use of adverb-adjective combinations (AACs) in sentiment analysis. The authors proposed a sentiment analysis method which is based on AACs which uses a linguistic evaluation of degrees of adverbs. The authors explained the experimental results on dataset of 200 news articles and compares the proposed technique with existing techniques of sentiment analysis. Based on Pearson correlation with human objects their experimental results gives higher accuracy.

Boyd and Ellison [85] stated that social networking sites (SNSs) are regularly seeking the attention of industry and academic researchers fascinated by their reach and affordance. The authors described in the introductory article the functions of SNSs and introduce a complete definition. The authors presented an aspect on the history of such sites, explaining development and key changes. The authors finally concluded that the condition is changing drastically and people should aware of which sites is using and why and for what purposes, especially other countries than U.S.

Another study by Agarwal et al. [86] also examined the twitter data for sentiment analysis. The authors proposed functions polarity prior POS- specific and studied the usage of a tree kernel to prevent the necessity for hectic feature engineering. The tree kernel and the new functions performed approx at the same level both surpassing the state of the art baseline. The authors concluded that for twitter data sentiment analysis is not that different as sentiment analysis for different genres.

Nasukawa and Yi [87] proposed an approach for sentiment analysis to find sentiments connected with negative or positive polarities from a document for specific subject, rather than classifying the whole document into negative or positive. The major problems in sentiment analysis are whether the statements points negative or positive behavior towards the subject and to be found how sentiment are described in texts. The authors stated that it is essential to clearly find out the semantic relationships between the subject and the sentiment expressions to increase the accuracy for the analysis of sentiment. A precision of 75-95% was achieved using the proposed system while applying the proposed mechanism on web pages and news articles for the identification of sentiments.

Wang et al. [88] proposed a system in U.S. elections 2012 for presidential candidates using real-time evaluation of sentiment on online microblogging site twitter. In order to collect the poll data the traditional analysis of election takes much time, but with the help of this system it takes data from more people with help of twitter, a microblogging service. It helps the social people like scholars, media and politician to broadcast their future perspective of the public opinion and electoral process. The authors finally concluded that the system and approach are generic, and should be adopted easily and spread across various other domains.

Wilson et al. [89] proposed a methodology for sentiment analysis using contextual polarity. The method identifies the statement polarity by detailed examination through phase-level evaluations of sentiments. The certainty of polarity is described as neutral or polar after evaluations. The evaluations of results show that the technique is capable to articulate statements in terms of contextual polarity of sentiments automatically. The evaluations of study are carried out on huge datasets resulting outputs greater than the baselines.

Kanayama and Nasukawa [90] proposed an unsupervised lexicon building approach which detected the clauses of polar that grant negative or positive effect in a particular domain. The entries that are lexical in nature to be received are called polar atoms, the lesser human-recognizable semantic models that justify clause polarity. By the usage of precision and overall density of consistency in the dataset, the statistical approximation selects necessary polar atoms through candidates, without change in the threshold values. The obtained result shows that the applied method is robust enough for datasets with different domains and also for weight of initial lexicon and the precision of polarity report from the automatically received lexicon was on average of 94%.

Choi and Cardie [91] studied that the essential cooperation in event of compositional semantics and presents a learning based technique that connects structural aasumption by compositional semantics for learning method. The authors conducted experiments that shows compositional semantics based on natural heuristics that can outperform the learning based techniques which does not integrate compositional semantics, whereas a technique which consolidate semantics compositional onto learning which is greater that other all alternatives. The authors also studied that for describing expression-level polarity, content word negator plays an important role. Finally the authors concluded that accuracy of classification of expression level linearly decreases as context that is gradually determined.

Melville et al. [92] presented a uniformed framework with respect to world-class associations using background lexical information and improve the information by using one of the available training examples to a particular domain. Experimental results shows that the authors methodology better performs than using training data or background knowledge within separation and text classification with lexical knowledge using to optional methodology. The authors concluded that they made two contributions. Firstly, they described a uniformed framework for combining knowledge of lexical for categorization of text in supervised learning and secondly, successfully applied the described methodology to analysis of classification of sentiment.

Paltoglou and Thelwall [93] stated that a large number of sentiment analysis methodologies have used support vector machines as their baselines with the weights of binary unigram. The authors in this paper explored if there is any reliable feature weighted schemes which can improve accuracy of classification with the help of retrieving the information. The authors shows that alternatives of the tf.idf scheme gives notable increase in accuracy for sentiment analysis, with the use of sublinear function for smoothing of document frequency and term frequency weights. The methodology was tested on large data set and obtained highest accuracy.

Fernandez et al. [94] developed a system introduced for the Subtask B Sentiment analysis of twitter i.e. SemEval 2014 task9. The authors system comprises of supervised methodology using techniques of machine learning, which using the text in dataset as features. This work is totally independent of any external resources and knowledge. The originality of author methodology depends on the use of skipgrams, n-grams and words as features. In the experimental study, it is clearly proves that skipgram shows better results than the ngram or word for the given datasets.

Mullen and Malouf [95] developed initial tests of statistics on a fresh datasets postings of group of political discussion that indicates the post that made response direct to post of others that having a greater likelihood that presents the perspective of opposing politics that of original post. The authors concluded that's the approaches of traditional text classifications is insufficient for this task in this dataset of sentiment analysis and the improvement can be made by utilizing information about how posters cooperate with one another.

Harb et al. [96] stated that the previous approaches until this paper were written suffered from drawback i.e. for a particular topic either the adjective is not available or from another topic it meaning is different. The authors proposed a new methodology which consists of two steps. Firstly, for a particular topic the authors extract a learning dataset from the internet. Secondly the authors extracting from the dataset, they made two classes that are negative and positive adjectives with respect to the topic. The experimental study on the real dataset shows the importance of authors methodology. The experiments are performed on dataset that are cinema reviews and blogs shows that with the author methodology, it is easy to extract the desired adjectives for a particular topic.

Kim and Hovy [97] stated that the identification of a sentiment was challenging problem. The authors developed a system for a topic it automatically searches the views posting users over an identified topic along with the sentiments expressed views. The system consists a component for describing sentiment derived from the expression and to merge the sentiments into statements. The authors experimented varying classification techniques and merged the sentiment over word and sentence level. For the improvement of recognition of Holder, the authors are using parser to attach areas that are more reliable with Holders. The learning techniques that are used is this system are support vector machines and decision list.

Martalo et al. [98] investigated how factors that are affective impact on the dialogue patterns and whether this impact may be explained and identified by Hidden Markov Models (HMMs). The goal of the authors is to study the chance of applying this model to classify behavior of users for the purposes of adaptation. The authors obtained the initial results of their research and present a debate of problems that are open. With the help of the results, the author claims that the complicated interaction between the pragmatic level and the acoustic level comprises an important facet of emotions contained in voice expressions.

Daoud [99] proposed a classifier and the introduced classifier contains four components which are AdaBoost which is a piece of an algorithm, Bayesian neural network, support vector machine and a technique for feature selection that is Signal-to-Noise. To confirm the efficiency of introduced classifier, the authors applied seven traditional classifiers to four datasets. The experimental study shows that applying the introduced classifier increases the rates of classification for all datasets. The author stated that SVMs key features are the control over capacity attained by margin optimization, sparseness of solution, the lack of local minima and the usage of kernels.

Yessenov and Misailovic [100] presents study of effectiveness of techniques of machine learning in text message classification by semantic meaning. The authors use comments of movie reviews from Digg that is social network which is popular as authors dataset and text classification can be done by negative or positive and objectivity or subjectivity attitude. The authors suggested different methodologies in text feature extraction such as using knowledge of WordNet synonyms, bounding word frequencies by threshold, handling negations, restricting to adjectives and adverbs, using large movie review corpus and a bag-of-words model. The authors analyze their performance on accuracy using four methodologies of machine learning that are K-Means clustering, Maximum Entropy, Decision Trees and Naïve Bayes. Finally, the authors concluded that bag-of- words model perform better than relative models.

Kang et al. [101] stated that the senti-lexicon existed does not properly adopt the word sentiment used in the restaurant review. The author introduced a senti-lexicon of restaurant reviews for the sentiment analysis. Using supervised learning technique a review document is classified as negative sentiment and positive sentiment, hence there is chance for the accuracy of positive classification to greater than 10% than the accuracy of negative classification. The author also introduced an improved version of Naïve Bayes to deal with these types of problems. The authors improved Naïve Bayes had managed to low the gap between positive and negative accuracy by 3.8% when applied with unigram + bigram and 28.5% when compared with SVM.

An experimental technique proposed by pang and lee [102] is comprised of a system that examines the sentiments through analysis of opinions by figuring out the ratio of positive words to total number of words and bifurcating the opinions as positive and negative. Further studies in 2008 proposed the methods which can classify the tweets based on the tweet term. Comparatively, machine learning techniques result pretty well than human generated baselines. SVM outperform the results as compare to Naïve Bayes. Regardless of using different types of features the authors did not attain desired accuracies over topic based categorization.[01 in original]

Jiang et al. [103] focus on target-dependent sentiment classification. Here target-dependent is the classification of sentiment as neutral, positive or negative depending on nature of the question that is asked. The authors proposed to make better target-dependent sentiment classification by joining features of target-dependent and considering related tweets. The authors also proposed that there is need of consideration current tweets to the related tweets by employing graph based optimization. As claimed by authors experimental results, the graph-based optimization increases the performance. .

3.2 Research Gaps:

Sentiment analysis could be a topic with personal and technical challenges. Opinions are expressed by multiple numbers of individuals and whenever there are tons of individuals that's forever an opportunity to own multiple opinions within the same subject. Thus, deciphering the moods are troublesome not just for humans however conjointly for computers. Equally we all know that opinions dissent from person to person and to research the actual text or sentence comes up with some technical challenges. Contemplate a situation that states that low machine is on the aspect of the reception. This statement could also be positive negative or neutral betting on matters and therefore the individuals. To research those variations from person to person and time to time is often a good challenge. The various experiments show that automatic sentiment analysis could be a smart tool for sentiment analysis however it cannot forever be trusty and that we cannot say that it forever offers correct analysis on the information.

The most of the researches within the field of sentiment analysis are in the main targeted on product reviews and flicks reviews however we have a tendency to be still behind in developing an honest model that perceive human language and interpret it well. Additional work is often wiped out increasing the techniques and algorithmic program to handle a lot of general writing, analysis on short sentences like abbreviations to perform cross domain analysis. The sentiment analysis algorithmic program but are often used in spam detection, detection of the context, analysis of the expression and to observe human language. Researches in rising the word identification, bipolar sentiment and developing completely automatic tools are often done additionally. The sentiment analysis algorithms use easy terms and expression however thanks to sizable amount of opinion, opinion orientation and therefore the different context as an entire are a giant task for computers to urge it done and extract the right sentiment.

Opinions are necessary to everybody as a result of whenever we'd like to form any call we would like to listen to different people’s opinions. This is often not solely true for a personal however conjointly true for any organization. Within the past, once anyone required to form a choice, he/she generally accustomed kindle opinions from friends and families. Once a corporation wished to search out opinions of the final public concerning its product and services, it was accustomed to conduct surveys and focus teams. However today with the explosive growth of the social media content on the online, within the past few years, the globe has taken a distinct form. Individuals will currently post reviews of totally different and various products on different bourgeois sites and give their views on nearly something in discussion forums and blogs, and conjointly in social network sites.

Currently if somebody desires to shop for a product, he/she is not any longer restricted to his/her friends and families' reviews as a result of there are many users' reviews on the online. For a corporation, it should not go to conduct any surveys or focus teams so as to assemble consumers' opinions concerning its product and people of its competitors' as a result of there are many resources obtainable where it can collect that information. However, finding opinion sites and observing them online will still be a formidable task. As a result of which there are an outsized range of numerous sites, and every website could have an enormous volume of narrow-minded text. It's troublesome for a reader's personality to search out relevant sites, extract connected sentences with opinions, read them, summarize them, and organize them into usable forms.

3.3. Literature Review Conclusion:

Sentiment analysis could be a field with giant space of application and provides scientist and educational organization millions of analysis challenges. With the ascension of net and net enabled applications sentiment analysis became thus common among totally different communities thus a lot of innovative, automatic and effective account techniques are needed that ought to overcome these challenges faced by Sentiment Analysis.

So the major focus of this research is to focus on all three sentiment polarities positive negative and neutral and provide an automated polarity generation and mechanism for classifying polarities with enhanced accuracy for future tweets.

Chapter 4 Methodology

Sentiment Analysis has become exceedingly useful in various domains. Any marketer, organization or researcher who has an interest in users live feedback on any object or event cans benefit from sentiment extraction and analysis. Traditional method of obtaining user opinions has been through reviews. The task of sentiment analysis began with most popular application in field of reviews for products and other entities for the purpose of marketing or benchmarking services. However, it soon became clear that reviews only are not very useful because of possibility of fake data that can be seen in many cases in form of closely related dates on all reviews, all five stars reviews, and review including exact names of related people in company etc.

Subjectivity analysis or opinion mining, on the other hand is the process of computationally determining linguistic expression of somebody opinion, sentiment, emotions etc. hidden inside the user generated texts Pang and Lee [34]. The application areas include marketing intelligence, topic surveys, tracking political activities or topics, advertisement placements etc.

Micro blog services such as Twitter, allow millions of its subscribers to post their views and comments in the form of posts called Tweets. Tweets limited into maximum of 140 characters in length. Tweets are mostly used by users to communicate their opinion about certain topic, personality, object, event or product on a social platform. Therefore, tweets are a rich source of mixture of subjective and objective information with no separation between positive or negative opinion.

Natural language processing is a process which requires a lot of cleaning in the data. Analyzing sentiments has always been complex and analyzing them through natural language processing is kind of a complex task. This chapter explains how different steps were carried out to complete the process. The methodology is explained through the over-view of the system diagram. Let’s go into the details of each step to find out about the methodology used in this study.

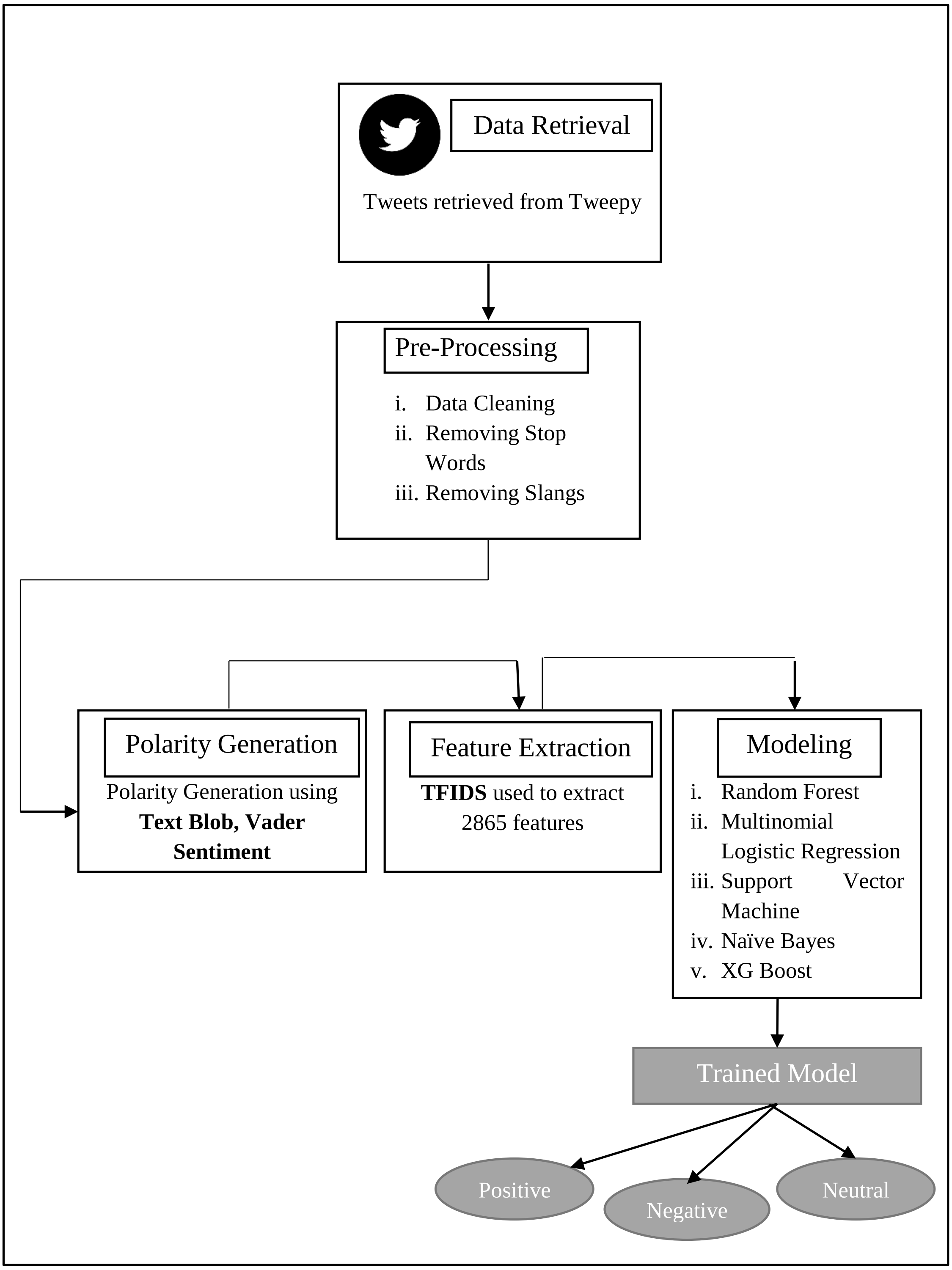

Figure 6 Methodology Overview

4.1. Data Retrieval

Data for this study has been accessed from a library named "Tweepy". Tweepy is an open source library available on GitHub which is used to access twitter data via OAuth authentication. The accession of data through tweepy has now become limited and it allows to access only 100 tweets maximum. But earlier there was no such constraints and we managed to extract 3000 tweets. We specifically targeted Donald Trump's id for accessing scrapping the text [18].

4.2. Data Pre-Processing

Data pre-processing is extremely important in natural language processing because it basically transforms the data so that it can be further used for processing. For this study the pre-processing steps include data cleaning using regular expressions, stop words removal and slang removal. These steps are explained in detail as following.

i. Data Cleaning:

Data cleaning is extremely important part of pre-processing. This step helps in removal of unnecessary words like URLS and numbers etc. Tweets contain such type of data often therefore we used regular expressions for identifying the exact pattern through data and then removing them. We managed to remove URLs, numbers, user names and special characters through regular expressions for removing the data.

ii. Removing stop words:

Stop words are normally the words which do not provide any useful information in a particular sentence or tweet. Removal of such words is important so that the actual content can be derived from the sentence. The **nltk** library consists of a corpus module which provides the objects and functions helpful in removing the stop words. We used this library to do so.

iii. Removing slangs:

A set of slang words with their abbreviations was defined in a separate file. The words from this file were then compared with the words of tweets. As a result of the comparison of each fore-mentioned, the final data we achieved was free of all the possible slang words as defined. These words were done replaced with the abbreviation of it mentioned in the file that was to be compared.

4.3. Polarity Generation

Polarity generation is a very important step because it identifies the type of tweets and their counts in a way. For polarity generation we used two kind of libraries, **Text Blob** and **vaderSentiment**. Text blob was used to create a score for which was further used to classify the tweet as positive, negative or neutral. If the score was greater than 0 than it was supposed to be a positive tweet, if the score was less than 0 than it was supposed to be a negative tweet and further if the score was equal to 0 than it was supposed to be a neutral tweet. Vader Sentiment technique was also used to find a polarity score which further helped in determining the class of the tweet. In our case if the vader score was greater then 0.05 then the tweet was positive and if the score was less than -0.05, than the score was negative. If the score is between the range (-0.05 to 0.05) then it was supposed to be neutral.

4.4. Feature Extraction:

For feature extraction, we used TFIDF technique which basically helps to create bag of words. It is used to create features such that each feature represents the frequency of words appeared. A total of 2865 features were made out of this.

4.5. Splitting:

Splitting was done on data to train the model and also to later test it. Different proportions were checked and finally the ratio of 70-30 was decided as it gave best results in that scenario.

4.6. Modelling:

Different classification models were used for classification of tweets into positive, negative and neutral tweets. We used the following models for classification.

- Random Forest
- Multi-Nomial Logistic Regression
- Support Vector Machine
- Naïve Bayes and

- XG-Boost

Random Forest:

Random forest is an ensemble model which works as a classifier and also as a regressor. It is a classic example of CART models. But for this study we will be focusing on Random Forest Classifier. The random forest classifier makes multiple decision trees on different variables and then combines the result. The final result is obtained from the majority voting of each decision tree.

Multinomial Logistic Regression:

Similar to Binary Logistic Regression, Multinomial Logistic Regression is used for the purpose of classification but as in case of binary logistic regression, there is classification on the basis of yes or no. For MLR there is classification for multiple classes. Similar to simple logistic regression, it calculates the maximum likely hood to find the probability of any relationship to a specific class.

SVM:

A support vector machine is also a classification (Supervised) algorithm which helps in classification problem. It works by using a hyperplane to separate the classes. If we are provided with the labelled data then considering a two-dimensional plane, the hyperplane will be a line which will be separating the plane in to two parts such that each class is separated on the side of hyperplane.

Naïve Bayes:

Naïve Bayes is another classifier which is based on the Naïve Bayes theorem. It is not just a single but it forms a collection of algorithms where all of them follow the Naïve Bayes theorem to get probabilities. These probabilities help in classification process.

XG-Boost:

XG Boost is a boosting algorithm which works really good for classification. It works on the basic of boosting where it attempts to convert the weak learners to strong learners by assigning weights to them. All the weak learners combine to form a strong learner. This is also a classic example of ensemble methods.

4.7. Model validation using new data:

For verifying our data in terms of either the labels made were correct or not, we got another labelled data and validated our model according to that. This was done by repeating the same methodology from which we acquired the labels. After applying the methodology on the labeled data, we got almost 70 – 80 % of accuracy which validated our methodology.

4.8. Results and Scoring Metrics:

The metrics used for results and scoring were Accuracy, F1-Score, Precision and Recall for measuring the performance of the model.

Chapter 5 Results

This chapter explains detailed discussion about metrics and model performance evaluation. The results did vary according to splitting of data into training and testing set. The details of which will be discussed further in this chapter followed by the performance metrics used in this study.

5.1. Performance metrics:

As mentioned before, the following performance metrics are used for measuring the performance of classification models in this study. They are

- Accuracy
- Precision
- Recall
- F1-Score

Accuracy:

Accuracy provides the score of how accurate the predictions are made by the model. Its check is majorly on the True Positives and True Negatives achieved out of the total data set. Accuracy of a model can be calculated according to the following formula.

$$Accuracy=\frac{tp+tn}{tp+tn+fp+fn} \qquad \text{(i)}$$

Precision:

Precision refers to the correctness of predictions made by the model. It basically tells how often the predicted values as 'yes' are correct. Precision for a classification model is calculated according to the following formula.

$$Precision=\frac{tp}{tp+fp} \qquad \text{(ii)}$$

Recall:

Recall refers to the check of the predicted values with the actual values. It checks and compares the predicted 'yes' values with the 'actual' yes values like if the values are actually 'yes' then how often does the model predict 'yes'. This can be also called as the True Positive Rate or Sensitivity. Recall is calculated by the formula given below.

$$Recall=\frac{tp}{tp+fn} \qquad \text{(iii)}$$

F1-Score:

Sometimes the precision and recall are not enough in certain cases. So, we use F1-Score to get the comparative score. F1-Score is basically the harmonic mean between precision and recall. It can be calculated through the following formula.

$$F1\text{-}Score=\frac{2*precision*recall}{precision+recall} \qquad \text{(iv)}$$

5.2. Data Splitting and Results

Data splitting is very crucial for model performance. Training the model with the right amount of data can lead to a high performance of the model. For this study however we tested the model on all kinds of probable splits, starting from 50-50 split to 90-10 splitting and results were recorded. The final split was decided on the 70-30 ratio. The results for each splitting are given below.

50-50 Splitting

Following are the results for 50-50 splitting on all model types.

Table 1 Accuracy Results on 50-50 Data Splitting

Models	Accuracy
Random Forest	77%
Multinomial Logistic Regression	77%
Support Vector Machine	78%
Naïve Bayes	59%
XG Boost	72%

Accuracy provides the score of how accurate the predictions are made by the model. For this 50-50 data splitting we get the upmost accuracy of 78% in Support Vector Machine and the least is 59% in Naïve Bayes. So we can conclude on this data splitting of 50-50 that our best model is Support Vector Machine in this scenario and weak model is Naïve Bayes

Table 2 Positive Tweets Scores on 50-50 Data Splitting

Models	Positive Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	76%	87%	86%
Multinomial Logistic Regression	77%	91%	84%
Support Vector Machine	81%	87%	84%
Naïve Bayes	71%	63%	67%
XG Boost	72%	90%	80%

Precision refers to the correctness of predictions made by the model. Here on 50-50 data splitting we get the upmost precision of 81% in Support Vector Machine whereas the least is 71% in Naïve Bayes for positive tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 91% in Multinomial Logistic Regression whereas the least is 63% in Naïve Bayes for positive tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 86% in Random Forest whereas the least is 67% in Naïve Bayes for positive tweets.

Table 3 Neutral Tweets Scores on 50-50 Data Splitting

Models	Neutral Tweets Scores		
	Precision	Recall	F1-Score

Random Forest	58%	75%	65%
Multinomial Logistic Regression	75%	7%	12%
Support Vector Machine	71%	34%	46%
Naïve Bayes	50%	26%	34%
XG Boost	43%	7%	12%

Here on 50-50 data splitting we get the upmost precision of 75% in Multinomial Logistic Regression whereas the least is 43% in XG Boost for neutral tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 75% in Random Forest whereas the least is 7% in Naïve Bayes and Multinomial Logistic Regression for neutral tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 65% in Random Forest whereas the least 12% in Multinomial Logistic Regression and XG Boost for neutral tweets.

Table 4 Negative Tweets Scores on 50-50 Data Splitting

Models	**Negative Tweets Scores**		
	Precision	Recall	F1-Score
Random Forest	75%	57%	65%
Multinomial Logistic Regression	75%	65%	70%
Support Vector Machine	71%	71%	71%
Naïve Bayes	44%	59%	50%
XG Boost	74%	53%	62%

For 50-50 data splitting we get the upmost precision of 75% in Random Forest whereas the least is 44% in Naïve Bayes for negative tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 71% in Support Vector Machine whereas the least is 53% in XG Boost for negative tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 71% in Support Vector Machine whereas the least 50% in Naïve Bayes for negative tweets.

60-40 Splitting

Following are the results for 60-40 splitting on all model types.

Table 5 Accuracy Results on 60-40 Data Splitting

Models	Accuracy
Random Forest	79%
Multinomial Logistic Regression	76%
Support Vector Machine	69%
Naïve Bayes	61%
XG Boost	73%

Accuracy provides the score of how accurate the predictions are made by the model. For this 60-40 data splitting we get the upmost accuracy of 79% Random Forest and the least is 61% in Naïve Bayes. So we can conclude on this data splitting of 60-40 that our best model is Random Forest in this scenario and weak model is Naïve Bayes.

Table 6 Positive Tweets Scores on 60-40 Data Splitting

Models	Positive Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	84%	87%	86%
Multinomial Logistic Regression	76%	92%	83%

Support Vector Machine	73%	87%	79%
Naïve Bayes	59%	64%	61%
XG Boost	79%	84%	81%

Precision refers to the correctness of predictions made by the model. Here on 60-40 data splitting we get the upmost precision of 84% in Random Forest whereas the least is 59% in Naïve Bayes for positive tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 92% in Multinomial Logistic Regression whereas the least is 64% in Naïve Bayes for positive tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 86% in Random Forest whereas the least is 61% in Naïve Bayes for positive tweets.

Table 7 Neutral Tweets Scores on 60-40 Data Splitting

Models	Neutral Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	60%	74%	66%
Multinomial Logistic Regression	100%	4%	7%
Support Vector Machine	77%	6%	11%
Naïve Bayes	9%	8%	9%
XG Boost	60%	51%	55%

Here on 60-40 data splitting we get the upmost precision of 100% in Multinomial Logistic Regression whereas the least is 9% in Naïve Bayes for neutral tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 74% in Random Forest whereas the least is 8% in Naïve Bayes and Multinomial Logistic Regression for neutral tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 66% in Random Forest whereas the least 7% in Multinomial Logistic Regression for neutral tweets.

Table 8 Negative Tweets Scores on 60-40 Data Splitting

Models	Negative Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	81%	60%	69%
Multinomial Logistic Regression	76%	68%	72%
Support Vector Machine	61%	67%	64%
Naïve Bayes	30%	26%	28%
XG Boost	67%	61%	64%

For 60-40 data splitting we get the upmost precision of 81% in Random Forest whereas the least is 30% in Naïve Bayes for negative tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 68% in Multinomial Logistic Regression whereas the least is 26% in Naïve Bayes for negative tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 72% in Multinomial Logistic Regression whereas the least 28% in Naïve Bayes for negative tweets.

70-30 Splitting

Following are the results for 70-30 splitting on all model types.

Table 9 Accuracy Results on 70-30 Data Splitting

Models	Accuracy
Random Forest	81%

Multinomial Logistic Regression	78%
Support Vector Machine	79%
Naïve Bayes	58%
XG Boost	78%

Accuracy provides the score of how accurate the predictions are made by the model. For this 70-30 data splitting we get the upmost accuracy of 81% Random Forest and the least is 58% in Naïve Bayes. So we can conclude on this data splitting of 70-30 that our best model is Random Forest in this scenario and weak model is Naïve Bayes.

Table 10 Positive Tweets Scores on 70-30 Data Splitting

Models	**Positive Tweets Scores**		
	Precision	Recall	F1-Score
Random Forest	85%	89%	87%
Multinomial Logistic Regression	78%	93%	85%
Support Vector Machine	84%	88%	86%
Naïve Bayes	72%	63%	67%
XG Boost	83%	84%	84%

Precision refers to the correctness of predictions made by the model. Here on 70-30 data splitting we get the upmost precision of 85% in Random Forest whereas the least is 72% in Naïve Bayes for positive tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 93% in Multinomial Logistic Regression whereas the least is 63% in Naïve Bayes for positive tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 87% in Random Forest whereas the least is 67% in Naïve Bayes for positive tweets.

Table 11 Neutral Tweets Scores on 70-30 Data Splitting

Models	Neutral Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	65%	79%	71%
Multinomial Logistic Regression	83%	7%	13%
Support Vector Machine	57%	42%	49%
Naïve Bayes	46%	29%	35%
XG Boost	61%	66%	63%

Here on 70-30 data splitting we get the upmost precision of 83% in Multinomial Logistic Regression whereas the least is 46% in Naïve Bayes for neutral tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 79% in Random Forest whereas the least is 7% in Multinomial Logistic Regression for neutral tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 71% in Random Forest whereas the least 13% in Multinomial Logistic Regression for neutral tweets.

Table 12 Negative Tweets Scores on 70-30 Data Splitting

Models	Negative Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	84%	63%	72%
Multinomial Logistic Regression	80%	69%	74%
Support Vector Machine	75%	72%	73%
Naïve Bayes	43%	58%	50%

XG Boost	71%	67%	69%

For 70-30 data splitting we get the upmost precision of 84% in Random Forest whereas the least is 43% in Naïve Bayes for negative tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 72% in Support Vector Machine whereas the least is 58% in Naïve Bayes for negative tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 72% in Random Forest whereas the least 50% in Naïve Bayes for negative tweets.

80-20 Splitting

The results for 80-20 splitting are shown below.

Table 13 Accuracy Results on 80-20 Data Splitting

Models	Accuracy
Random Forest	79%
Multinomial Logistic Regression	77%
Support Vector Machine	78%
Naïve Bayes	57%
XG Boost	78%

Accuracy provides the score of how accurate the predictions are made by the model. For this 80-20 data splitting we get the upmost accuracy of 79% Random Forest and the least is 57% in Naïve Bayes. So we can conclude on this data splitting of 80-20 that our best model is Random Forest in this scenario and weak model is Naïve Bayes.

Table 14 Positive Tweets Scores on 80-20 Data Splitting

Models	Positive Tweets Scores		
	Precision	Recall	F1-Score

Random Forest	86%	87%	86%
Multinomial Logistic Regression	77%	91%	84%
Support Vector Machine	82%	89%	85%
Naïve Bayes	71%	62%	67%
XG Boost	83%	84%	83%

Precision refers to the correctness of predictions made by the model. Here on 80-20 data splitting we get the upmost precision of 86% in Random Forest whereas the least is 71% in Naïve Bayes for positive tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 91% in Multinomial Logistic Regression whereas the least is 62% in Naïve Bayes for positive tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 86% in Random Forest whereas the least is 67% in Naïve Bayes for positive tweets.

Table 15 Neutral Tweets Scores on 80-20 Data Splitting

Models	**Neutral Tweets Scores**		
	Precision	Recall	F1-Score
Random Forest	58%	77%	66%
Multinomial Logistic Regression	75%	7%	12%
Support Vector Machine	47%	34%	39%
Naïve Bayes	41%	25%	31%
XG Boost	62%	68%	65%

Here on 80-20 data splitting we get the upmost precision of 75% in Multinomial Logistic Regression whereas the least is 41% in Naïve Bayes for neutral tweets.

Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 77% in Random Forest whereas the least is 7% in Multinomial Logistic Regression for neutral tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 66% in Random Forest whereas the least 12% in Multinomial Logistic Regression for neutral tweets.

Table 16 Negative Tweets Scores on 80-20 Data Splitting

Models	Negative Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	78%	62%	69%
Multinomial Logistic Regression	75%	65%	70%
Support Vector Machine	75%	69%	72%
Naïve Bayes	40%	54%	46%
XG Boost	72%	67%	70%

For 80-20 data splitting we get the upmost precision of 78% in Random Forest whereas the least is 40% in Naïve Bayes for negative tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 69% in Support Vector Machine whereas the least is 54% in Naïve Bayes for negative tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 72% in Support Vector Machine whereas the least 46% in Naïve Bayes for negative tweets.

90-10 Splitting

Following are the results for 90-10 splitting done on data.

Table 17 Accuracy Results on 90-10 Data Splitting

Models	Accuracy

Random Forest	75%
Multinomial Logistic Regression	76%
Support Vector Machine	70%
Naïve Bayes	59%
XG Boost	78%

Accuracy provides the score of how accurate the predictions are made by the model. For this 90-10 data splitting we get the upmost accuracy of 78% XG Boost and the least is 59% in Naïve Bayes. So we can conclude on this data splitting of 80-20 that our best model is XG Boost in this scenario and weak model is Naïve Bayes.

Table 18 Positive Tweets Scores on 90-10 Data Splitting

Models	**Positive Tweets Scores**		
	Precision	Recall	F1-Score
Random Forest	79%	86%	83%
Multinomial Logistic Regression	79%	89%	84%
Support Vector Machine	68%	96%	79%
Naïve Bayes	66%	55%	60%
XG Boost	83%	84%	83%

Precision refers to the correctness of predictions made by the model. Here on 90-10 data splitting we get the upmost precision of 83% in XG Boost whereas the least is 66% in Naïve Bayes for positive tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 96% in Support Vector Machine whereas the least is 55% in Naïve Bayes for positive tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of

84% in Multinomial Logistic Regression whereas the least is 60% in Naïve Bayes for positive tweets.

Table 19 Neutral Tweets Scores on 90-10 Data Splitting

Models	Neutral Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	63%	71%	67%
Multinomial Logistic Regression	100%	8%	15%
Support Vector Machine	100%	7%	14%
Naïve Bayes	15%	16%	16%
XG Boost	62%	68%	65%

Here on 80-20 data splitting we get the upmost precision of 100% in Multinomial Logistic Regression and Support Vector Machine whereas the least is 15% in Naïve Bayes for neutral tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 71% in Random Forest whereas the least is 7% in Support Vector Machine for neutral tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 67% in Random Forest whereas the least 14% in Support Vector Machine for neutral tweets.

Table 20 Negative Tweets Scores on 90-10 Data Splitting

Models	Negative Tweets Scores		
	Precision	Recall	F1-Score
Random Forest	73%	56%	63%
Multinomial Logistic Regression	68%	68%	68%

Support Vector Machine	83%	49%	61%
Naïve Bayes	24%	23%	28%
XG Boost	72%	67%	70%

For 80-20 data splitting we get the upmost precision of 83% in Support Vector Machine whereas the least is 24% in Naïve Bayes for negative tweets. Recall refers to the check of the predicted values with the actual values. On this data we get the upmost recall of 67% in XG Boost whereas the least is 23% in Naïve Bayes for negative tweets. F1-Score use to get the comparative score. We get the upmost F1-Score of 70% in XG Boost whereas the least 28% in Naïve Bayes for negative tweets.

The best results were achieved from the split of **70-30** ratio.

5.3. Personality Judgment:

Textblob and Vaderseniment are used to calculate sentiment of tweets. For the proposed methodology Donald Trump tweets are used as a dataset. Around 3000 tweets were fetched from tweepy. After applying preprocessing techniques the results of tweets polarities are shown in graphs.

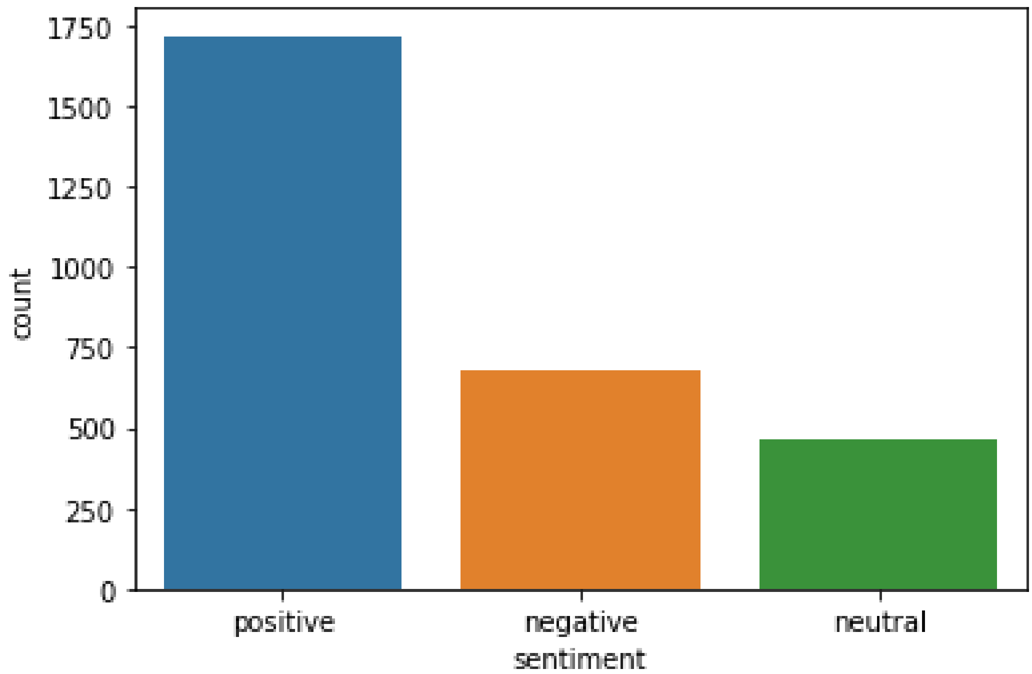

Figure 7 Tweets Polarities Division with TextBlob

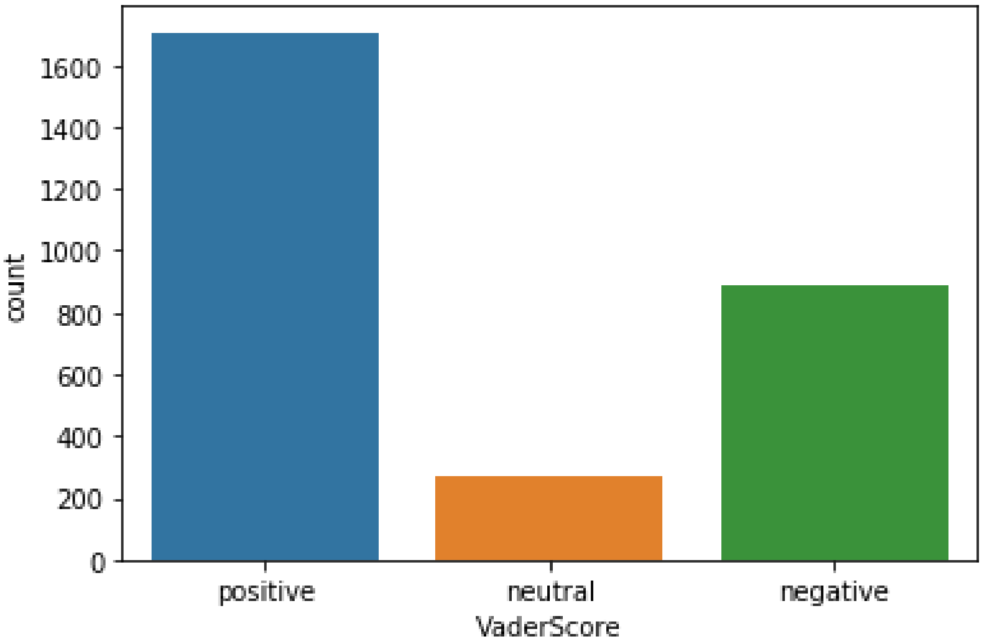

Figure 8 Tweets Polarities Division with vaderSentiment

After Applying both sentiments techniques it has been observed that Donald Trump mostly post positive tweets. Negative and neutral tweets ratio is comparatively low from Positive tweets.

5.4. Final Results

Results are shown according to the accuracies and confusion matrix.

Accuracy Results

The final metric selected was accuracy and below are the accuracies for all the models used.

Models	Accuracy
Random Forest	81%
Multinomial Logistic Regression	78%
Support Vector Machine	79%
Naïve Bayes	58%
XG Boost	78%

Table 21: Accuracy Comparison

The above results show that Random Forest gives the best accuracy on the 70-30 split.

Confusion Matrix

The confusion matrix from random forest is shown below.

	Predicted Positive	Predicted Neutral	Predicted Negative
Actual Positive	133	22	56
Actual Neutral	1	99	26
Actual Negative	25	32	466

Table 22: Confusion Matrix

This confusion matrix shows the TPs, TNs, FPs and FNs for all three classes i.e. Positive, Neutral and Negative classes. And apparently the accuracy seems to be of good count as shown in the table before.

Chapter 6 Conclusion

The following chapter consists about the summary of contributions claimed in this study and also gives a hint of the future work for this particular study.

6.1. Contributions

The study is based on a classic example of natural language processing use case. Now a days a lot of people use NLP for sentiment analysis and this has helped a lot of enterprises and large firms to get the opinions of their customers. Getting opinions of customers not only help the enterprise owners in improving the quality of business but also leads to customer satisfaction.

Sentiment Analysis has become exceedingly useful in various domains. Any marketer, organization or researcher who has an interest in users live feedback on any object or event cans benefit from sentiment extraction and analysis. Traditional method of obtaining user opinions has been through reviews.

The task of sentiment analysis began with most popular application in field of reviews for products and other entities for the purpose of marketing or benchmarking services. However, it soon became clear that reviews only are not very useful because of possibility of fake data that can be seen in many cases in form of closely related dates on all reviews, all five stars reviews, and review including exact names of related people in company etc.

Not only in business and enterprise systems, but sentiment analysis is widely used in social media to get opinion mining done on certain tweets and statuses. A huge processing is done on tweets available on Twitter where a variety of users and famous celebrities express their views and opinions on a certain topic. These tweets are used to analyze the sentiment behind them using NLP.

This study has been proposed to apply Twitter Sentiment Analysis on tweets accessed from Tweepy. Tweepy is an open source library available for python users which helps in retrieving the tweets. Almost 3000 tweets of Donald Trump were retrieved via Tweepy which was further used for processing.

The additional work in this study was based on classification of neutral tweets as well since previously there has been a variety of studies on classification of positive and neutral tweets. So, the major focused was on multi-class classification problem with Vadersentiment and TextBlob in which different machine learning algorithms are used to classify positive, negative and neutral tweets. Also personality judgment is done based on their tweets polarity. Different machine learning algorithms and Random Forest gave the best results with the accuracy of 81%.

6.2. Future Work

For this study we have just used single algorithms to get the best results possible in classifying the neutral tweets.

In the future we can use hybrid approaches to increase the accuracy of identification of labels and we can move to deep learning algorithms to achieve a better accuracy as compared to current system.

References

[1] Social media analytics [online]: http://en.wikipedia.org/wiki/Social_analytics (last accessed on 15 November 2014).

[2] [Online]: http://www.w3.org/2012/06/pmod/opinionmining.pdf (last accessed on 15 November 2014).

[3] Cambria Erik. Schuller Bj??rn. Xia Yunqing. Havasi Catherine. New Avenues in Opinion Mining and Sentiment Analysis"

[4] [Online]: http://www.cs.uic.edu/~liub/FBS/IEEE-Intell-Sentiment-Analysis.pdf (last accessed on 15 November 2014).

[5] Liu B. Sentiment Analysis and Subjectivity. Handbook of Natural Language Processing, Second edition, 2010

[6] Popescu A-M. Etzioni O. Extracting product features and opinions from reviews, Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP), 2005

[7] Aue A. Gamon M. Customizing sentiment classifiers to new domains: A case study. Proceedings of Recent Advances in Natural Language Processing (RANLP), 2005.

[8] Pang B. Lee L. Vaithyanathan S. Sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2002, (EMNLP):79'86.

[9] Pang B. Lee L. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales, Proceedings of the Association for Computational Linguistics (ACL),2005:115'124

[10] Bethard S. Yu H. Thornton A. Hatzivassiloglou V. and Jurafsky D. Automatic extraction of opinion propositions and their holders. Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text, 2004.

[11] Hu M. Liu B. Mining opinion features in customer reviews. In Proceedings of AAAI, 2004:755'760.

[12] Esuli A. Sebastiani F. Determining the semantic orientation of terms through gloss classification. In Proceedings of CIKM-05, the ACM SIGIR conference on

information and knowledge management, Bremen, DE, 2005.
[13] Turney P. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the Association for Computational Linguistics (ACL), 2005: 417’424.
[14] Bales,’Robert.”Interaction’Process’Analysis’Article.”Interaction Process Analysis.’Web.’25’ Mar.’2012.
[15][Online]:http://web.njit.edu/~da225/NetHelp/default.htm?turl=Documents%2Fsubjectivityobjectivityidentification.htm (last accessed on December 15 2014)
[16] Kamps, J., Marx, M., Mokken, R.J., de Rijke, M., Using WordNet to measure semantic orientation of adjectives In Language Resources and Evaluation (LREC),2004
[17] Sentiment analysis and opinion mining [Online]: http://stp.lingfil.uu.se/~santinim/sais/2014/bingliu.pdf (last accessed on 18 December 2014)

[18] account, Donald J. TrumpVerified. “Donald J. Trump (@RealDonaldTrump).” *Twitter*, Twitter, 27 Oct. 2019, https://twitter.com/realdonaldtrump.

[19] Krouska, Akrivi, Christos Troussas, and Maria Virvou. "The effect of preprocessing techniques on Twitter sentiment analysis." 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA). IEEE, 2016.

[20] Effrosynidis, Dimitrios, Symeon Symeonidis, and Avi Arampatzis. "A comparison of pre-processing techniques for twitter sentiment analysis." International Conference on Theory and Practice of Digital Libraries. Springer, Cham, 2017.

[21] Wegrzyn-Wolska, Katarzyna, et al. "Explore the effects of emoticons on Twitter sentiment analysis." Comput. Sci. Inf. Technol. 2 (2016): 65.

[22] Hamed, A. R., Renxi Qiu, and Dayou Li. "The importance of neutral class in sentiment analysis of Arabic tweets." Int. J. Comput. Sci. Inform. Technol 8 (2016): 17-31.

[23] Sarlan, Aliza, Chayanit Nadam, and Shuib Basri. "Twitter sentiment analysis." Proceedings of the 6th International Conference on Information Technology and Multimedia. IEEE, 2014.

[24] Tyagi, Priyanka, and R. C. Tripathi. "A Review towards the Sentiment Analysis Techniques for the Analysis of Twitter Data." Available at SSRN 3368718 (2019).

[25] Hasan, Ali, et al. "Machine learning-based sentiment analysis for twitter accounts." Mathematical and Computational Applications 23.1 (2018): 11.

[26] Wan, Yun, and Qigang Gao. "An ensemble sentiment classification system of twitter data for airline services analysis." 2015 IEEE international conference on data mining workshop (ICDMW). IEEE, 2015.

[27]Asghar, M. Z., Khan, A., Ahmad, S., Khan, I. A., & Kundi, F. M. (2015). "A unified framework for creating domain dependent polarity lexicons from user generated review"s. PloS One, 10(10). e0140204

[28] Masud, F., Khan, A., Ahmad, S., & Asghar, M. Z. (2014). "Lexicon-based sentiment analysis in the social web". Journal of Basic and Applied Scientific Research, 4(6), 238–248.

[29]Asghar, M. Z., Khan, A., Ahmad, S., Qasim, M., & Khan, I. A. (2017). "Lexicon-enhanced sentiment analysis framework using rule-based classification scheme". PloS One, 12(2). e0171649

[30] Prieto, V. M., Matos, S., Alvarez, M., Cacheda, F., & Oliveira, J. L. (2014). "Twitter: A good place to detect health conditions". PloS One, 9(1). e86191

[31]Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T., & Urena-Lopez, L. A. (2012). "Random walk weighting over sentiwordnet for sentiment polarity detection on twitter". In Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis (pp. 3–10). Association for Computational Linguistics.

[32]Kundi, F. M., Ahmad, S., Khan, A., & Asghar, M. Z. (2014). “Detection and scoring of internet slangs for sentiment analysis using SentiWordNet”. Life Science Journal, 11(9), 66–72.

[33]Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). “The effects of twitter sentiment on stock price returns”. PloS One, 10(9). e0138441

[34] Ribeiro, P. L., Weigang, L., & Li, T. (2015). “A unified approach for domain-specific tweet sentiment analysis. In Information Fusion (Fusion)”, 2015 18th International Conference on (pp. 846–853). IEEE.

[35] Na´dia F.F. da Silva, Eduardo R. Hruschka, and Estevam R. Hruschka Jr. “Tweet sentiment analysis with classifier ensembles. Decision Support Systems”, 66:170 – 179 , 2014. ISSN 0167-9236. doi: http://dx.doi.org/10.1016/j. dss.2014.07.003. URL http://www.sciencedirect.com/ science/article/pii/S0167923614001997.

[36] G. Gautam and D. Yadav. “Sentiment analysis of twitter data using machine learning approaches and semantic analysis”. In Contemporary Computing (IC3), 2014 Seventh International Conference on, pages 437–442, Aug 2014. doi: 10.1109/IC3.2014.6897213.

[37] Alec Go, Richa Bhayani, and Lei Huang. “Twitter sentiment classification using distant supervision”. Processing, pages 1–6, 2009. URL http://www.stanford.edu/~alecmgo/papers/TwitterDistantSupervision09.pdf.

[38] G.Vinodhini and RM.Chandrasekaran. “Sentiment analysis and opinion mining: A survey”. International Journal of Advanced Research in Computer Science and Software Engineering, 2(6):282–292, June 2012.

[39] Ajinkya Ingle, Anjali Kante, Shriya Samak, and Anita Kumari. “Sentiment analysis of twitter data using hadoop”. International Journal of Engineering Research and General Science, 3:144–147, November-December 2015.

[40] Davidov, D., Tsur, O., Rappoport, A., 2010. “Enhanced sentiment learning using twitter hashtags and smileys”, in: Proceedings of the 23rd international

conference on computational linguistics: posters, Association for Computational Linguistics. pp. 241–249.

[41] Go, A., Bhayani, R., Huang, L., 2009. "Twitter sentiment classification using distant supervision". CS224N Project Report, Stanford 1, 12.

[42] Mohammad, S., Kiritchenko, S., Zhu, X.D., "Building the state-of-the-art in sentiment analysis of tweets", in: SemEval@NAACL-HLT 2013 Nrc-canada.

[43] Read,J.,2005."Using emoticons to reduce dependency in machine learning techniques for sentiment classification", in:Proceedings of the ACL student research workshop, Association for Computational Linguistics. pp. 43–48.

[44] Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R., 2011. "Sentiment analysis of twitter data", in: Proceedings of the workshop on languages in social media, Association for Computational Linguistics. pp. 30–38.

[45] Zhang,L.,Ghosh,R.,Dekhil,M.,Hsu,M.,Liu,B.,2011. "Combining lexicon-based and learning-based methods for twitter sentiment analysis".

[46] Hassan, A., Abbasi, A., Zeng, D., 2013. "Twitter sentiment analysis: A bootstrap ensemble framework", in: Social Computing (SocialCom), 2013 International Conference on, IEEE. pp. 357–364.

[47] Hassan, A., Abbasi, A., Zeng, D., 2013. "Twitter sentiment analysis: A bootstrap ensemble framework", in: Social Computing (SocialCom), 2013 International Conference on, IEEE. pp. 357–364.

[48] Ankita and Nabizath Saleena, " An Ensemble Classification System for Twitter Sentiment Analysis", International Conference on Computational Intelligence and Data Science (ICCIDS 2018,pp. 937-946

[49] X. Chen, M. Vorvoreanu and K. Madhavan, "Mining Social Media Data to Understand Students' Learning Experiences", IEEE Transaction, 2014, vol. 7, no. 3, pp. 246-259.

[50] S. Bahrainian and A. Dangel, "Sentiment Analysis using Sentiment Features", in Int. joint Conf. of Web Intelligence and Intelligent Agent Technologies, 2013, pp. 26-29.

[51] S. Bahrainian and A. Dangel, “Sentiment Analysis using Sentiment Features”, in Int. joint Conf. of Web Intelligence and Intelligent Agent Technologies, 2013, pp. 26-29.

[52] Z. WANG, V. J. Chuan TONG, X. XIN and H. C. CHIN, “Anomaly Detection through Enhanced Sentiment Analysis on Social Media Data”, in 6th International Conference on Cloud Computing Technology and Science, 2014, pp. 918-922.

[53] Wang, X.; Gotoh, O.: “Accurate molecular classification of cancer using simple rules”. BMC Med. Genom. 2(1), 1 (2009)

[54] Califf, M.E.; Mooney, R.J.:” Bottom-up relational learning of pattern matching rules for information extraction”. JMLR 4, 177–210 (2003)

[55] Choi, Y.; et al.: “Identifying sources of opinions with conditional random fields and extraction patterns”. In: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 355–362. Association for Computational Linguistics (2005)

[56] Feldman, R.; Rosenfeld, B.; Fresko, M.: TEG—“A hybrid approach to information extraction”. Knowl. Inf. Syst. 9(1), 1–18 (2006)

[57] Asghar, M.Z., Khan, A., Khan, F. et al. Arab J Sci Eng (2018) 43: 857. *https://doi.org/10.1007/s13369-017-2770-1*

[58] A.C.E.S Lima. and L.N. de Castro, “Automatic sentiment analysis of Twitter messages”, in 4th Int. Conf. on Computational Aspects of Social Networks (CASoN), 2012, pp. 52-57.

[59] R. Batool, A. M. Khattak, J. Maqbool and S. Lee, “Precise tweet classification and sentiment analysis”, in 12th Int. Conf. on Computer and Information Science (ICIS), 2013, pp. 461-466

[60] Hamed, A. R., Renxi Qiu, and Dayou Li. "The importance of neutral class in sentiment analysis of Arabic tweets." Int. J. Comput. Sci. Inform. Technol 8 (2016): 17-31.

[61] Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of twitter sentiment on stock price returns. PloS One, 10(9). e0138441)

[62] Ribeiro, P. L., Weigang, L., & Li, T. (2015). A unified approach for domain-specific tweet sentiment analysis. In Information Fusion (Fusion), 2015 18th International Conference on (pp. 846–853). IEEE.)

[63] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou and P. Li, "User-level sentiment analysis incorporating social networks", Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1397--1405, 2011.

[64] L. Chen, C. Liu and H. Chiu, "A neural network based approach for sentiment classification in the blogosphere", Journal of Informetrics, vol. 5, no. 2, pp. 313-322, 2011.

[65] M. Anjaria and R. Guddeti, "Influence factor based opinion mining of Twitter data using supervised learning", Communication Systems and Networks (COMSNETS), 2014 Sixth International Conference on, pp. 1--8, 2014.

[66] A. Barhan and A. Shakhomirov, "Methods for Sentiment Analysis of Twitter Messages", 12th Conference of FRUCT Association, 2012.

[67] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining.", LREc, vol. 10, pp. 1320--1326, 2010.

[68] M. Koppel and J. Schler, "THE IMPORTANCE OF NEUTRAL EXAMPLES FOR LEARNING SENTIMENT", Computational Intell, vol. 22, no. 2, pp. 100-109, 2006.

[69] A. Go, R. Bhayani and L. Huang, "Twitter sentiment classification using distant supervision", CS224N Project Report, Stanford, vol. 1, p. 12, 2009.

[70] A. Andrew, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods20016Nello Christianini and John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000. xiii + 189 pp., ISBN: ISBN 0-521-78019-5 Hardback: £27.50", Kybernetes, vol. 30, no. 1, pp. 103-115, 2001.

[71] A. Berger, V. Pietra and S. Pietra, "A maximum entropy approach to natural language processing", Computational linguistics, vol. 22, no. 1, pp. 39--71, 1996.

[72] K. Nigam, J. Lafferty and A. McCallum, "Using maximum entropy for text classification", IJCAI-99 workshop on machine learning for information filtering, vol. 1, pp. 61--67, 1999.

[73] D. Romero, B. Meeder and J. Kleinberg, "Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter", Proceedings of the 20th international conference on World wide web, pp. 695--704, 2011.

[74] S. Tan and J. Zhang, "An empirical study of sentiment analysis for Chinese documents", Expert Systems with Applications, vol. 34, no. 4, pp. 2622-2629, 2008.

[75] J. Martineau and T. Finin, "Delta TFIDF: An Improved Feature Space for Sentiment Analysis", Proceedings of the Third International ICWSM Conference, vol. 9, 2009.

[76] FA. Nielson. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", arXiv preprint arXiv:1103.2903, 2009.

[77] SM. Mohammad, S. Kiritchenko and X. Zhu, "NRC-Canada: Building the State-of- the-Art in Sentiment Analysis of Tweets", arXiv preprint arXiv:1308.6242, 2013.

[78] E. Kouloumpis, T. Wilson and J. Moore, "Twitter Sentiment Analysis: The Good the Bad and the OMG!", ICWSM, vol. 11, pp. 538—541, 2011.

[79] K. Denecke, "Using SentiWordNet for Multilingual Sentiment Analysis", Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on IEEE, pp. 507—512, 2008.

[80] B. Gokulkrishnan, P. Priyanthan, T. Ragavan, N. Prasath and A. Perara," Opinion Mining and Sentiment Analysis on a Twitter Data Stream", Advances in ICT for Emerging Regions (ICTer), 2012 International Conference IEEE, pp. 182—188, 2012.

[81] F. Neri, C. Aliprandi, F. Capeci, M. Cuadros and T. By," Sentiment Analysis on Social Media ", Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining, pp.919—926, 2012.

[82] T. Wilson, J. Wiebe and P. Hoffman, “Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis”, Computational Lingusitics, MIT Press, vol. 35, no. 3,pp. 399—433, 2009.

[83] N. Godbole, M. Srinivasaiah and S. Skiena,” Large-Scale Sentiment Analysis for News and Blogs”, ICWSM, vol. 7, no. 21, pp. 219—222, 2007.

[84] F. Benamara, C. Caserano, A. Picariello, DR. Recupero and VS. Subrahmanian,” Sentiment Analysis: Adjectives and Adverbs are better than Adjectives Alone”, ICWSM, 2007.

[85] D. Boyd and N. Ellison,” Social Network Sites: Definition, History, and Scholarship”, IEEE Engineering Management Review, vol. 3, no. 38, pp. 16—31, 2010.

[86] A. Agarwal, B. Xie, I. Vovsha, O. Rambow and R. Passonneau, “Sentiment Analysis of Twitter Data”, Proceedings of the workshop on languages in social media, Association for Computational Linguistics, pp. 30—38, 2011.

[87] T. Nasukawa and J. Yi, “Sentiment Analysis: Capturing Favorability Using Natural Language Processing”, Proceedings of the 2nd international conference on Knowledge capture, ACM, pp. 70—77, 2003.

[88] H. Wang, D. Can, A. Kazemzadeh, F. Bar and S. Narayanan,” A System for Real- time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle”,

Proceedings of the ACL 2012 System Demonstrations, Association for Computational Linguistics, pp. 115—120, 2012.

[89] T. Wilson, J. Wiebe and P. Hoffman,”Recognizing Contextual Polarity in Phrase- Level Sentiment Analysis”, Proceedings of the conference on human language technology and empirical methods in natural language processing, Association for Computational Linguistics, pp. 347—354, 2005.

[90] H. Kanayam and T. Nasukawa, ” Fully Automatic Lexicon Expansion for Domain- oriented Sentiment Analysis”, Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, pp. 355—363, 2006.

[91] Y. Choi and C. Cardie, “Learning with Compositional Semantics as Structural Inference for Subsentential Sentiment Analysis”, Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, pp. 793—801, 2008.

[92] P. Melville, W. Gryc and RD. Lawrence, ” Sentiment Analysis of Blogs by Combining Lexical Knowledge with Text Classification”, Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, pp. 1275—1284, 2009.

[93] G. Paltoglou and M. Thelwall, “A study of Information Retrieval weighting schemes for sentiment analysis”, Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, pp. 1386—1395, 2010.

[94] J. Fernandez, Y. Gutierrez, J. Gomez and P. Martinez-Barco, ” GPLSI: Supervised Sentiment Analysis in Twitter using Skipgrams”, Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), number SemEval, pp. 294—299, 2014.

[95] T. Mullen and R. Malouf, “A Preliminary Investigation into Sentiment Analysis of Informal Political Discourse”, AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, pp. 159—162, 2006.

[96] A. Harb, M. Plantie, G. Dray, M. Roche, F. Trousset and P. Poncelet, “Web opinion mining: How to extract opinions from blogs?”, Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology, ACM, pp. 211—217, 2008.

[97] SM. Kim and E. Hovy, “Determining the Sentiment of Opinions”, Proceedings of the 20th international conference on Computational Linguistics, Association for Computational Linguistics, 2004.

[98] A. Martalo, N. Novielli and F. De Rosis,” Attitude Display in Dialogue Patterns”,Proc. AISB 2008 Symposium on Affective Language in Human and Machine, 2008.

[99] E. AL-Daoud, “Integration of Support Vector Machine and Bayesian Neural Network for Data Mining and Classification”, World Academy of Science, Engineering and Technology, vol. 64, 2010.

[100] K. Yessenov and S. Misailovic, “Sentiment Analysis of Movie Review Comments”,Methodology, pp. 1—17, 2009.

[101] H. Kang, SJ. Yoo and D. Han, ” Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews”, Expert Systems with Applications, Elsevier, vol. 39, no. 5, pp. 6000—6010, 2012.

[102] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques", Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, pp. 79--86, 2002.

[103] L. Jiang, M. Yu, M. Zhou, X. Liu and T. Zhao, "Target-dependent twitter sentiment classification", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 151--160, 2011.

[104] E. Psomakelis, K. Tserpes, D. Anagnostopoulos, and T. Varvarigou, “Comparing methods for twitter sentiment analysis,” KDIR 2014 - Proceedings of the Int. Conf. on Knowledge Discovery and InformationRetrieval, pp. 225-232, 2014.

[105] account, Donald J. TrumpVerified. “Donald J. Trump (@RealDonaldTrump).” Twitter, Twitter, 27 Oct. 2019, https://twitter.com/realdonaldtrump.