AN ENSEMBLE CLASSIFIER FOR SARCASM DETECTION USING SOCIAL NETWORKING DATA

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Computer Science and Engineering

By

Jyoti Godara

41800123

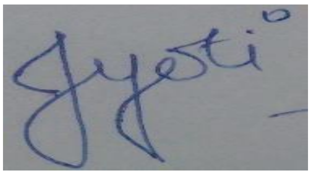
Supervised By
Dr. Isha Batra

Co-Supervised by
Dr. Rajni Aron

LOVELY PROFESSIONAL UNIVERSITY
PUNJAB
2022

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled, **AN ENSEMBLE CLASSIFIER FOR SARCASM DETECTION USING SOCIAL NETWORKING DATA** submitted for the Degree of Doctor of Philosophy in Computer Science and Engineering is the result of my original and independent research work carried out under the guidance of Supervisor Dr. Isha Batra, Associate Professor, School of Computer Science and Engineering, Lovely Professional University, Punjab and Co-Supervisor Dr. Rajni Aron, Associate Professor, Department of Data Science, Narsee Monjee Institute of Management, Mumbai. This work has not been submitted for the award of any degree, diploma, associateship, fellowship of any University or Institution.



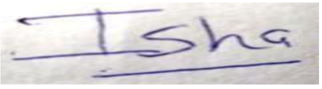
Date: 14th Sep,22

Investigator: Jyoti Godara

Registration No: 41800123

CERTIFICATE

This is to certify that the thesis entitled *"AN ENSEMBLE CLASSIFIER FOR SARCASM DETECTION USING SOCIAL NETWORKING DATA"* submitted by Jyoti Godara for the award of degree of Doctor of Philosophy in Computer Science and Engineering, Lovely Professional University, is entirely based on the work carried out by her under my supervision and guidance. The work reported, embodies the original work of the candidate and has not been submitted to any other university or institution for the award of any degree or diploma, according to the best of my knowledge.



Signature of Supervisor

Name: Dr. Isha Batra

Date: 14th Sep,22

Signature of Co- Supervisor

Name: Dr. Rajni Aron

Date: 14th Sep,22

ABSTRACT

The internet has shifted the way people express their thoughts and opinions. It is currently primarily done through blog postings, online forums, product review websites, social media, Facebook, Twitter, Google Plus, and other social media sites are such examples. As a result, social media generates a vast amount of sentiment-rich data in text, audio, video, and other formats. This big data collected plays a significant role in banking, agriculture, education, healthcare, marketing, economy, stocks etc. People nowadays rely heavily on user-generated content on the internet when making decisions. For example, if someone wants to buy a product or use a service, they research it online and then discuss it on social media before deciding. This form of user-generated material can benefit a variety of organizations because it allows them to examine customer ratings and comments to enhance their products. Among all social networking platforms, Twitter is one of the famous micro-blogging services. It is a social media platform that allows friends, family, and coworkers to communicate and stay in touch via short, frequent messages. These messages, which users can read or post containing the maximum length of 148 characters, are called Tweets. It can include images, videos, links, and text. These messages appear on your profile, are forwarded to your followers, and can be found via a Twitter search. Sentiment analysis studies people's or society's feelings or thoughts regarding a specific occurrence or subject. It analyses the users' sentiments and classifies the text's polarity. However, it must automate because the data is too large for a single person to analyze. Machine learning approaches, linguistic techniques, hybrid techniques, and other sentiment analysis techniques have all been used to analyze text from social network data. Sentiment analysis is divided into several stages: data collection, pre-processing, sentiment detection, and data classification. Various issues have been identified during the sentiment analysis process, such as domain dependence, thwarted articulations, order dependence, an explicit negation of opinion, identification of the subjective parts of the text, etc. One of the significant issues identified is sarcasm. It can be viewed as an expression in which individuals speak or write something entirely

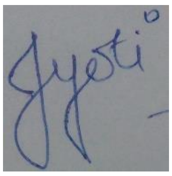
at odds with their intended meaning. Due to its obscurity, sarcasm is incredibly difficult to spot. Sarcasm can sometimes take the shape of irony. Criticism is one of the most frequent ways to utilize sarcasm. Most commonly, it is employed to express one's feelings or opinions, particularly on social networking websites like Twitter and Facebook. It is one of the most crucial parts of sentiment analysis as, if wrongly classified, it will degrade the system's overall performance. As per the study of other researchers, it was observed that the accuracy of sentiment analysis could be improved by rigorous analysis and interpretation of sarcasm sentences. As a result, the primary goal of this research is to design a technique that can best detect sarcasm and improve the entire sentiment analysis process. To achieve this, the fundamental research is carried out in the form of these objectives 1) To study and analyze various sarcasm detection techniques for sentiment analysis. 2) To design and implement an ensemble classification method for sarcasm detection. 3) To test and validate the proposed method and compare the result with existing approaches. To achieve the first objective, which is to study and analyze various sarcasm detection techniques for sentiment analysis, we have reviewed literature related to sentiment analysis, Twitter sentiment analysis, sarcasm detection, multiple strategies to perform the detection and various ensemble learning techniques. Ensemble learning methods use multiple classifiers for the training purpose and combine them. As per the previous studies, it was observed that ensemble learning provides better results for the sentiment analysis as compared to others. The way of achieving the second objective which is to design and implement an ensemble classification method for sarcasm detection. Firstly, we have collected social networking dataset in the form of tweets and news headlines from the website of Kaggle. The two datasets are combined to create a multidomain dataset for model training and testing. The combined dataset consists of 67,895 records containing sarcastic and non-sarcastic records. After performing the pre-processing step, data will be cleaned by removing the hyperlinks, redundant data, stop words, etc. Features are extracted using the random forest algorithm. Principal Component Analysis would be applied for the feature reduction phase. Later, K-Mean clustering is applied for similar clustering information. After that, the ensemble classifier was built and implemented, with four ensemble classifiers designed to detect sarcasm. (Ensemble1 -SKD), a mix of Support Vector Machine, K-Nearest Neighbor, and Decision Tree is the first ensemble classification algorithm. The second ensemble classifier (Ensemble2 -SLD) for sarcasm detection combines Support Vector Machine, Logistic Regression, and Decision Tree classifiers. The third ensemble model (Ensemble3 -MLD) combines Multilayer Perceptron, Logistic Regression, and Decision Tree, while the fourth ensemble model (Ensemble4 -SLM) uses Multilayer Perceptron, Support Vector Machine and Logistic Regression. A voting classifier has been applied to test data for producing the output. Spyder, a robust development environment for the Python language with advanced editing, testing and numerical computation environments, was used to implement the code altogether. Pandas library is used for handling the dataset. Scikit learns libraries are used for feature representation, classification, similarity measures and evaluation purposes. The suggested model is written in Python programming language. Most of the existing works focused on ensemble of the features for detecting sarcasm. In the proposed work, random forest algorithm, PCA and K- mean clustering are used in the sequence and further multiple classifiers are ensembled for detection of sarcasm. The model's performance is evaluated using a variety of performance indicators as illustrated by F1-score, precision, recall, and accuracy. Ensemble 2 (a combination of Support Vector Machine, Logistic Regression, and Decision Tree) outperformed the other three ensemble classifiers, according to the results. On the dataset, Ensemble 2 outperformed Adaboost, Decision Tree, Random Forest, and K-Nearest Neighbor classifier based on accuracy, precision, recall and f1-score. The accuracy of the proposed approach improved by 4.31 percent, 3.76 percent, 76.15 percent and 6.38 percent, the precision of the proposed approach improved by 3.37 percent, 2.22 percent, 29.57 percent and 4.54 percent, the recall of the proposed approach improved by 3.33 percent, 3.33 percent, 75.47 percent and 5.68 percent, F1-score of proposed approach improved by 4.49 percent, 3.33 percent, 132 percent and 5.68 percent respectively, when compared to Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor classifier. Also, in comparison with existing ensemble approaches, the proposed method has achieved the highest accuracy among others.

ACKNOWLEDGEMENT

I would like to present my deepest gratitude to **Dr. Isha Batra and Dr. Rajni Aron** for their guidance, advice, understanding and supervision throughout the development of this thesis and study. Despite their busy schedule they have been available at every step, devoting time and energy and the much-needed counsel and advice. This enabled me to sail through the tough times and complete this enormous task.

I would like to thank to the **research project committee members** for their valuable comments and discussions. A special thanks to the management of **Lovely Professional University** for their support in academic concerns and letting me involve in research study. The doctoral programme of LPU has made it possible for me to pursue my dream of research and upgrade my knowledge.

My sincere feeling of gratefulness also goes to my parents and family members who always motivated me in all the endeavors of my life including this research work in LPU. I am also thankful to my mother-in-law Smt. Kamlesh Poonia and my husband Rajesh Poonia for offering full support to me during the entire period of my research work. My special thanks to my daughters Jiya Poonia and Kaira Poonia for giving me joyful and happy moments during the entire journey of my research work. Finally, I would like to thank each person who has directly and indirectly helped and motivated me in this journey.



JYOTI GODARA

TABLE OF CONTENTS

Contents	Page No.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	**DESCRIPTION**
API	APPLICATION PROGRAMMING INTERFACE
URL	UNIFORM RESOURCE LOCATOR
POS	PARTS OF SPEECH
ML	MACHINE LEARNING
AI	ARTIFICIAL INTELLIGENCE
ASD	AUTOMATIC SARCASM DETECTION
NLP	NATURAL LANGUAGE PROCESSING
NN	NEURAL NETWORK
DT	DECISION TREE
RF	RANDOM FOREST
SVM	SUPPORT VECTOR MACHINE
LR	LOGISTIC REGRESSION
KNN	K-NEAREST NEIGHBOR
NB	NAÏVE BAYES
TF-IDF	TERM FREQUENCY- INVERSE DOCUMENT FREQUENY
WMV	WEIGHTED MAJORITY VOTING
CV	CROSS VALIDATION
BOW	BAG OF WORDS
W2VC	WORD2VEC AND CLUSTERING
HDFS	HADOOP DISTRIBUTED FILE SYSTEM
HIVE	HADOOP SCRIPTING LANGUAGE
DM	DATA MINING
DOC2VEC	DOCUMENT VECTORS
DL	DEEP LEARNING
SA	SENTIMENT ANALYSIS
ME	MAXIMUM ENTROPY
DCNN	DEEP CONVOLUTION NEURAL NETWORK
MLP	MULTILEVEL PERCEPTRON
PCA	PRINCIPAL COMPONENT ANALYSIS
FFNN	FEED-FORWARD NEURAL NETWORK

CHAPTER 1

INTRODUCTION

This section builds up the knowledge required to understand the problem statements of the research work by knowing the basics of sentiment analysis, its process, applications and various issues. Study about sarcasm and ensemble learning is also explained in the section. Based on the existing problems, multiple research objectives have been framed to carry out the research.

.

1.1 Introduction

The area of social media analytics has seen a rising interest in sentiment analysis over a while. Misunderstanding, confusion, and inaccuracy in social networking data have increased as the data's volume, rationality, and authenticity have increased. Detecting sarcasm in textual data has become a difficult task, as it has become a new way of expressing the feelings in which people write or say something that isn't exactly what they mean. As a result, researchers have recently expressed an interest in developing various strategies for detecting sarcasm in texts to improve sentiment analysis performance.

1.2 Sentiment Analysis

Sentiment analysis is now more widely used by many people with various interests and inspirations, thanks to massive growth in social media usage over the past few years. Retrieving knowledge from those data is of enormous importance and significance since people worldwide may have diverse ideas about several themes linked to politics, education, tourism, culture, commercial items, or topics of general interest. Knowing their feelings as indicated by their words on various platforms and data on the websites people frequent, their top priority for purchases, etc., became an essential component for gauging public opinion on a particular issue. Today, one of the most used sentiment analysis techniques is classifying a text's polarity [1].

The texts can be categorized as positive, negative way or neutral in nature concerning labelling or number of levels. Still, overall, it refers to the emotions of the text fluctuating from a happy to sad mode. The word ‘sentiment’ represents a subjective and objective subject and a real or imaginable subject that bridges the distinction between a positive or negative subject. Sentiment analysis is an analysis dependent on the spread of rumors or gossip. Sentiment analysis is an analytical scheme based on text analysis. Sentiment analysis involves determining the subjectivity of belief, and the outcome of a review or tweet. Sentiment analysis is concerned with classifying an individual's opinion into different classes according to data size and document type. The schemes applied for sentiment analysis are many and depend on a variety of methods of natural language processing and machine learning techniques for extracting sufficient features and classifying text into suitable polarity labels [2].

1.2.1 Need of Sentiment Analysis

Text classification is a necessary activity that revolves around the interpretation of human language and emotions. Textual or spoken data reflects these sentiments. These expressions are extremely complex since they contain a wide range of emotions. The use of symbolic linguistic indicators such as punctuation (amazing!!!), emojis, wordplay (greattttttttttt for incredible), innovative spellings (multi-day for now), and usage of slang over internet (GMG for "Goodness My God") has increased the complexity for analysis of the social media content. As a result, automation to detect the expressed attitudes is required. It bridges the gap between what the user has written and what he wishes to express through textual data. For e-commercial links and websites like Amazon.com and Epinion.com [171], several customer-generated product reviews and services are used to construct a marketing plan. It has the potential to influence customers' purchase decisions and service subscriptions 172]. For example, if a person wishes to visit a particular region, rather than asking a friend or relative, he searches online for visitor reviews before making any selections. When it comes to business regulation, if a customer is willing to buy a particular product, he will first read all of the customer reviews before deciding whether or not to buy it. As a result, we can conclude that the internet contains a vast amount of material that could be thoroughly investigated [173]. As a result, analyzing and forecasting the polarity of sentiment is critical to comprehending social phenomena and societal trends [171].

1.2.2 Sentiment Analysis process

Analyzing sentiments is a difficult job. The pipeline of sentiment analysis includes five tasks viz collection of data, preparation of text, detection of text, classification of text and finally output presentation. The Figure 1.1 represents the general phases of performing sentiment analysis.

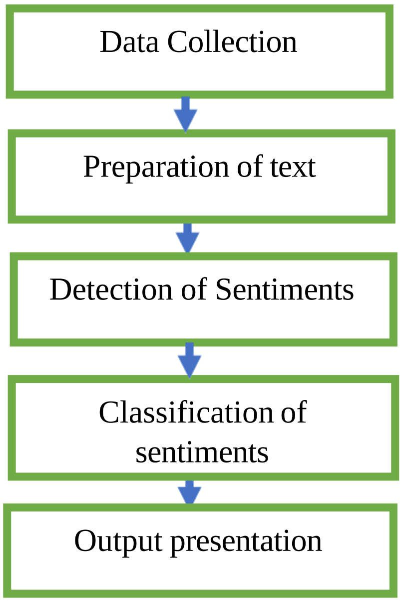


Figure 1.1: Process of Sentiment Analysis

All steps included in the above figure have been explained below:

i. Data collection: The first task in sentiment analysis is to gather textual data, which is collected by the user-generated content comprised of blogs, forum boards, and social networks. The datasets collected are unsystematic. Various vocabularies, slang, the context of writing etc., are expressed in different ways in these data. It is probably not possible to analyze this data manually. Hence, text analytics and natural language processing are applied for feature extraction and classification.

ii. Text preparation: This task is concerned with the cleaning of the extracted data before analysis. This process identifies and eliminates the non-textual and redundant contents.

iii. Sentiment detection: The purpose of this task is to inspect the extracted sentences of the reviews and opinions. This process retains statements containing subjective kinds of expressions (opinions, beliefs and views) but discards those sentences conveying the objective data (facts, factual information) [3].

iv. Classification of Sentiment: Here, the classification of subjective type of sentences is carried out into positive, negative, good, bad; like, dislike etc. however, there are several points on which classification depends.

v. Presentation of output: Converting amorphous text into useful information is the major purpose of sentiment analysis. After analysis, the outcomes can be displayed on bar chart, pie chart, graphs and line graphs. In addition to this, time can be reviewed and a sentiment timeline with the selected value (percentages, frequency and averages) is constructed with time by displaying time in the graph format.

1.2.3 Sentiment Analysis levels

The sentiment analysis pipeline comprises of three levels to class the sentiment analysis approaches. These levels are document, sentence and feature or aspect level [4] as represented in Figure 1.2.

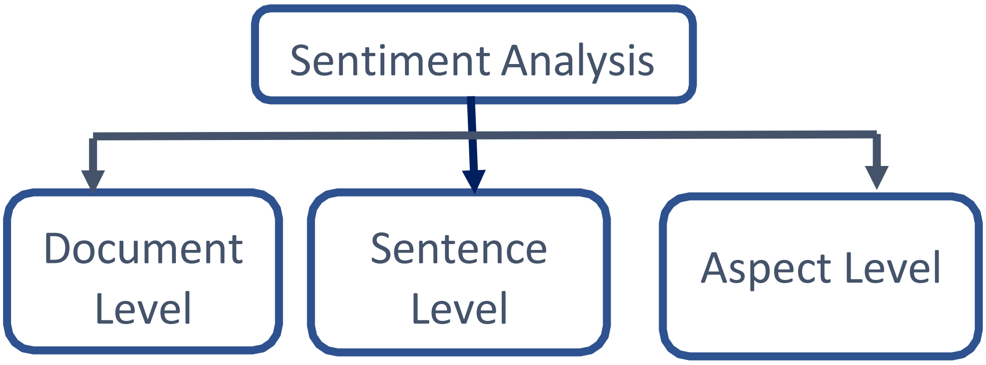


Figure 1.2: Sentiment Analysis levels

i. Document level: Document level task aims at classifying whether an overall document conveys a negative or positive sentiment. After a product is reviewed, for example, the system determines if the report expresses a completely good or negative impression about the product. This level of analysis is founded on the assumption that every document expresses a viewpoint on a subject (for example, a product). As a result, it should not be used in documents that evaluate or compare numerous objects.

ii. Sentence level: This level of analysis aims at determining whether the opinion conveyed in the sentence is neutral, negative or positive [5]. Two ways are used to conduct sentence level analysis. One method is to consider the analysis as a simple three-way classification task, with positive, negative, and neutral labels. Another strategy is to first investigate subjectivity in the phrase by separating opinionated and non-opinionated texts, then labelling those subjective sentences with one of the two labels (positive or negative). The disadvantage of sentence level analysis is that each sentence is semantically and syntactically connected to other parts of the text. Thus, this task needs both local as well as global circumstantial information.

iii. Aspect-level: Sentiment analysis at the document and sentence levels use fixed sentiment polarity rested upon the entire document/sentence instead of the subjects in the document/sentence. Clearly, this is not appropriate in several cases. Sentiment analysis at the aspect level is a sub-function of sentiment analysis, and its goal is to perform sentiment recognition and classification at aspect level [6]. Aspect-level sentiment analysis attempts to anticipate the sentiment polarities of each explicit aspect word in a sentence.

1.3 Applications of sentiment analysis

Sentiment analysis can be applied to many areas. Some of them are discussed as follows:

- Applications using websites' reviews: On social media, there is a massive collection of analysis and opinions on just about anything. This includes product reviews, political opinions, and comments on various services, among other things. As a result, an opinion scrutiny method is required for the extraction of opinions on a specific item or entity. This is essential in order to achieve automation in the expression of opinion or scoring of the specified object, entity, or other entity. It will meet both the needs of the customers and the needs of the retailer.

- Sub-component technique: The recommender scheme could benefit from the opinion forecaster scheme as well. The best scheme will not recommend items that receive a lot of negative feedback or receive a tiny amount of scoring. Certain users are subjected to nasty language and other unenthusiastic rudiments during online discussions. These signals can be easily identified with the identification of excessively negative beliefs, and appropriate action can be taken against them.
- Industry intellect: People nowadays read evaluations about a product before purchasing it, which are available on many social media platforms. In several businesses, the social media estimation also determines the victory or failure of certain things. As a result, it is reasonable to conclude that the scrutiny of opinions has a significant role in businesses. For the improvement of their businesses, industries also want to collect opinions through internet assessments. This positively impacts their reputation and is also beneficial to customers.
- Across realm: New investigators in sociology and a variety of other fields, including therapeutics and sports instruction, benefit from opinion analysis, which reveals trends in human attitudes, particularly on a societal level.
- Smart homes: Smart homes are regarded as the machines of the future. In the coming years, all houses will be systematized, and people will be able to operate any part of the house via remote technology.
- Business intelligence: Customers who are distributed throughout the globe can be challenging to study; however, their comments and thoughts can be reviewed on the company's internet forum. [174].
- Forensic investigation: It can also be useful in the legal investigation of fraud and criminal systems. [175].

1.4 Twitter Sentiment Analysis

These days, Twitter is one of the most influential social networking microblogging services. It provides its services globally as a medium of information sharing. Thus, extracting people's opinions from tweets on different topics measures the impact of various events or classifies sentiments became a topic of huge interest. With over 413 million monthly active users and over 500 million tweets each day, Twitter is the world's most popular social media platform and has now turned out to be a treasure house for businesses and individuals having a vast social media following in preserving and growing their influence and status. Sentiment analysis provides these businesses potential to monitor various social media sites online. The process of sentiment analysis automatically detects whether a part of the text holds sentimental or opinionated content and can further find the text's polarity [7]. Twitter sentiment categorization aims to determine if a tweet's sentiment polarity is positive, negative, or neutral. Tweets are generally made up of partial, noisy and unstructured sentences, asymmetrical words, misspelt words and non-dictionary words. Below mentioned properties are generally used for performing Twitter sentiment analysis-

- Length: A tweet can be up to 140 characters long.
- Data availability: The Twitter API (Application Programming Interface) makes gathering millions of tweets for training relatively simple.
- Language model: Twitter users post tweets using a variety of media, including their smartphones. Misspellings and slang are significantly more common in Tweet messages than in other areas.
- Domain: The tweets posted on twitter are of small size. The tweets can be posted about many topics that adapt itself a certain topic, different from other sites [8].

1.4.1 General Process of Twitter Sentiment Analysis

Due to its many applications, tweets are currently one of the most talked-about topics. One method for creating campaign strategies for political elections is based on the polarization analysis of Twitter users' opinions about political parties and politicians. Due to Twitter sentiment analysis's effectiveness and timeliness, businesses use it to monitor consumer attitudes toward their brands and products.

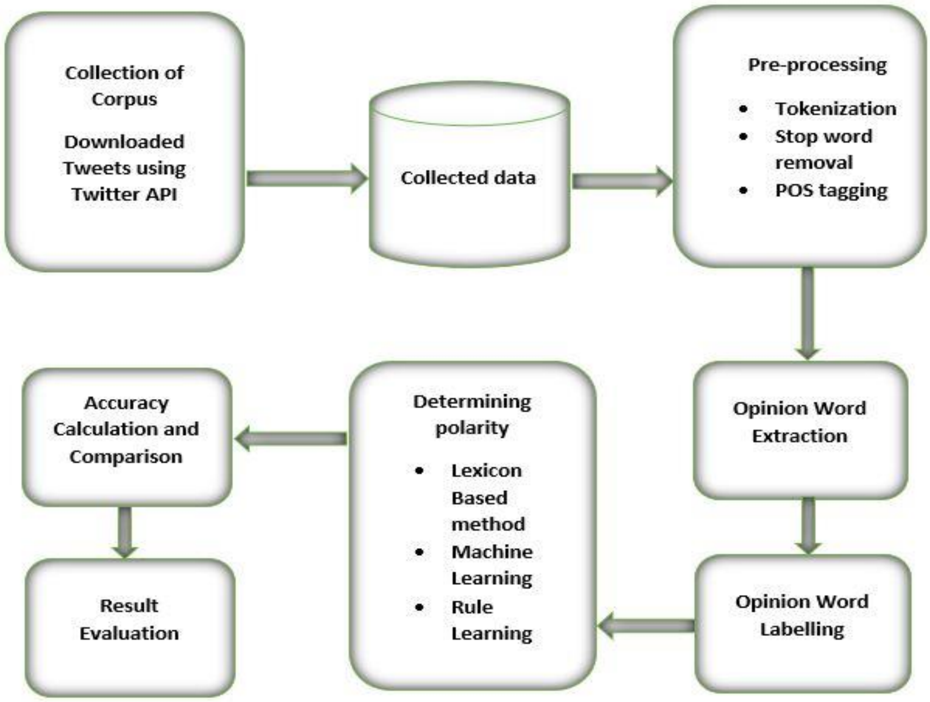


Figure 1.3: Twitter Sentiment Analysis Process

Figure 1. 3 shreds of evidence the general process of Twitter sentiment analysis. The steps involved in Twitter sentiment analysis process have been discussed as follows:

a. Corpus collection: One of the critical difficulties in Twitter's sentiment analysis is gathering labelled datasets. The text posts were gathered using the Twitter API. Then, a dataset of three kinds [9]—positive feelings, negative emotions, and a group of objective texts—is created by combining these tweet posts (no emotion).

b. Data Pre-processing: A major shortcoming of getting data sets from Twitter is the noise within the data. Twitter data (tweets) can be simple text, mentions of users (@user), and references to URLs (Uniform Resource Locator) or content tags, also called hashtags (#). Before feature extraction and categorization, twitter data is pre-processed in this step. Many steps are practised to eradicate noise from the Twitter dataset, which includes duplicates, retweets, points, tweets with a single URL and tweets written in other languages. These noises do not contribute to the accuracy of data classification and are thus eliminated. In addition, other fundamental pre-processing techniques, including tokenization, stop word removal, and POS (Parts of Speech) tagging, are used to convert the text input to lower case:

- Tokenization: Tokens denote individual words or terms. The purpose of tokenization is to split a thread of text into tokens.
- Stop Word Removal: All human languages consist of plenty of Stop words. By removing these words, the low-level information is removed from the text to bring more attention to important information. Removing stop words certainly decreases the dataset dimensionality and thus lessens the training time owing to the smaller number of tokens used in training.
- POS tagging: Adjectives, adverbs, and verbs, as well as specific groups of nouns, are good indicators of subjectivity and emotion. Parsing or dependency trees can be used to produce syntactic dependency patterns. [10].

c. Opinion Word Extraction: There are various unique elements of the Twitter language paradigm. The feature space is thought to be reduced by some of these features. To start the procedure, all bigrams and unigrams above a predetermined threshold are removed from the corpus. As an illustration, all bigrams and unigrams with frequencies higher than five are eliminated as candidate features. Bigrams and Unigrams are typically chosen in a word- and phrase-level sentiment analysis. Trigrams can also be used to simply stretch it. The frequency of each candidate trait found in a tweet is then determined. In light of this, the word "frequency for every tweet" is used to build the feature vector shown below:

$$(\{word\ 1: frequncy1, word2: frequency2 \ldots\}, "polarity")$$

d. Opinion Word Labelling: A polarity is then assigned to the tweet based on the aggregated words' verification against a vocabulary of positive and negative words that contains two files. +1 will be given to any good words or hashtags that are included in the tweet, and -1 will be given to any negative ones.

e. Polarity Determining: The polarity of a tweet is established in the final step. Using either of the two approaches—a lexicon-based technique or machine learning—the effect of pre-processing on sentiment categorization is estimated [11].

1.4.2 Twitter Sentiment Classification Algorithms

In the field of sentiment analysis, researchers frequently presumptively believe that a single holder's consistent sentiment polarity toward an object permeates the entire document being studied. The range of reviews is an excellent illustration of how the presumption is true. This is accurate for tweet statistics as well, given that tweets are typically shorter. For a person to comprehend complex information in a tweet is not natural.

Approaches can largely be classified into three classes: Approaches based on lexicons, machine learning, and rules.

i. Lexicon-based approach: This method uses a dictionary that already has labelled dictionaries. The tokenize creates tokens from the input text. Then, all recently acquired tokens are compared to the lexicon in the dictionary. In the event of a successful match, the score is added to the input score's full set of scores. Think about the term "dramatic." The overall text score is raised if this term has a good dictionary match. In the second scenario, the term is either marked as negative, or the score is reduced. Even though this strategy sounds amateurish, it has some good possibilities. The idea behind vocabulary-based classification algorithms is that a document's spirit is determined by its essential elements (words or phrases). Simple word counts, document scoring with thresholding, and majority voting are fundamental strategies. [12]. Lexicon-based approaches use some predefined lists of words to relate each word in the list to a particular emotion. Besides this, these approaches have two main categories.

- Dictionary-based approaches: These approaches use the lexicon dictionary to locate positive opinion words and negative opinion words. Dictionary-based techniques use lexical databases such as WordNet to increase a physically generated seed set. Automatic development will detect pairwise word relationships and produce a dictionary of appropriate magnitude. Though the dictionary-based approaches can have an abundance of semantic words, those words are generally independent in context and domain.
- Corpus-based approaches: These approaches use a large corpus of words, and depending on the syntactic patterns, one can obtain opinion words in the context. The corpus-based approaches aid in fixing the issue of obtaining opinion

words with context-based orientation [13].

ii. Machine Learning (ML) based approach: Machine Learning is an AI (Artificial Intelligence) subfield that deals with algorithms. These algorithms facilitate learning to the computer. In general, this means that an algorithm is given a set of data and is asked to infer knowledge about the data's attributes. This information aids in the prediction of other data that may be encountered in the future. The potential to predict the hidden data becomes easy as most of the non-random data consists of patterns which facilitate the machines to general. For generalization, a model is trained through a computer to determine the significant aspects of the data.

Different machine learning algorithms, whether supervised or unsupervised, applied to sentiment analysis extract expressive information from structured and unstructured textual data to assist decision-makers. Supervised algorithms have been shown to be more efficient in deciding the polarity of emotions; however, they need plenty of labelled data which is difficult to obtain. In contrast, unsupervised algorithms, although not greater, are still beneficial because they can operate without labelled data.

iii. Rule-based classifiers: The data space is modelled using a rule set in a rule-based classifier. The one on the left represents a normalized location on the feature set, while the one on the right is labelled the class. The phrases revolve around the concept of presence. Because it is not informative with limited data, the term absence is rarely employed. There are numerous criteria for creating rules, and the training phase constructs all of the rules using these criteria. Support and confidence are the two most usually used criteria. The support training data set contains a number of examples that are relevant to the rule. The conditional likelihood that the right side of the rule will be satisfied if the left side is met is known as confidence. [20]. VADER is a well-known rule-based model with multiple lexical features. It aims to perform sentiment analysis in micro-blog data and provides actual and wide-ranging results in contrast to standard models.

1.5 Issues of sentiment analysis

In this section, a few challenges of sentiment analysis are presented:

- Identifying subjective parts of the text: Subjective portions reflect emotional content. In one example, the same term can be subjective, while in another, it can be objective. It makes it difficult to distinguish between the subjective and objective sections of the text.
- Domain dependence: A same sentence or word might have different meanings in different domains. For example, the word "unique" has a favourable connotation in the film industry, but it has a negative connotation when applied to a vehicle.
- Thwarted articulations: The general polarity of the document is decided by a segment of the content in a few sentences.
- Explicit negation of opinion: Instead of employing the basic no, not, never, and so on, feelings can be negated from numerous perspectives. Such negations are difficult to spot.
- Order dependence: The importance of discourse study in sentiment analysis cannot be overstated.
- Entity recognition: The content of a specific drug should be isolated, and then the polarity toward it should be examined.
- Building a classifier for emotional and target tweets: It's important to distinguish between tweets with and without sentiments when categorizing them.
- Implementing sentiment investigation to Facebook messages: Because of numerous restrictions on the Facebook diagram programming interface and security arrangements in obtaining information, little work on sentiment investigation on Facebook information has been done for the most part.
- Sarcasm: In sentiment analysis, understanding the meaning of a statement in a certain context in order to classify the text on the basis of polarity is a major challenge. Sentiment analysis produces excellent results in theoretical language because it conveys the intended meaning. Nonetheless, the intrinsically symbolic use of figurative language signifies something other than a specific context, making sentiment research a difficult task. Sarcasm is defined as "a type of assessment in which people convey their bad feelings through the use of positive or reinforced positive terms in the content." [176].

1.6 Sarcasm Detection

Microblogging sites give a public forum for ordinary people to express their thoughts, ideas, and opinions on a range of topics and happenings. Sarcasm is a more sophisticated form of irony found on social media and blogging sites, where trolling and/or condemnation of others are prominent. Sarcasm and irony differ slightly from one another. The word "sarcasm" is typically used to convey verbal irony. In the fields of psychology, semantics, and cognitive science, sarcasm has generated a lot of scientific interest. Automatic sarcasm detection is a great asset for reputation management and opinion mining. Therefore, the NLP (natural language processing) group has given ASD (Automatic Sarcasm Detection) a lot of attention. [21].

Managing text on social networks is a difficult task. It stands out because it uses slanted vocabulary and is informal. People also utilize unstructured content defensively as a means of expression. We've noticed that much of the material on social networking sites is misspelt and contains slang and acronyms. The use of figurative language is done so short on Twitter due to the character constraint of 140, which creates another problem. People are free to use sarcastic language to communicate their thoughts in order to achieve their communication goals. There is no specific structure there for creating sarcastic sentences. As a result, the main objective of the research on sarcasm detection is to identify the traits that allow readers to tell satirical writings apart from other types of texts. It's a sort of negation that doesn't have an explicit negation marker. The sarcastic expression "Waking up at 4 a.m. with a headache is enjoyable." [177] is equivalent to the non-sarcastic sentence, "Waking up at 4 a.m. with a headache is not fun.". Performance suffers as a result of these sarcastic views. As a result, detecting a sarcastic message is critical for improving sentiment analysis performance [161]. It will also aid in the removal of the text's deliberate ambiguity [177]. From a business standpoint, it is critical to comprehend product reviews and movie popularity because they may suffer if placed in the incorrect category [142]. Sarcasm can be used to shame someone online, to be disrespectful to others, or to make fun of them [129]. For improved data classification, spam filtering and manufacturing product market analysis employ sarcasm detection. [181].

1.6.1 Sarcasm Detection from Twitter data

A binary text classification function can be used to represent the process of sarcasm detection in tweets. A crucial mechanism, sarcasm detection in text classification has ramifications for a wide range of fields, including safety, marketing, and fitness. Methods for sarcasm detection may assist businesses in analyzing customer opinions of their items. It uses those businesses as leverage to highlight the calibre of their output. Classifying tweets, which contain latent information in the message that a person shares with others, is particularly important in sentiment analysis [22]. In addition, it is possible to determine sarcasm from a tweet's structure. Sarcasm detection using machine learning methods can be accomplished successfully.

Building an effective classification model depends on many aspects. The main aspects are the attributes used along with the sovereign features observed in the learning algorithm, which can be easily combined within the class. The following stages illustrate the sarcasm detection process based on machine learning.

a. Data Collection: Obtaining an appropriate dataset is the first step toward sarcasm detection. Dataset plays an important role in any data mining study. For both sarcastic and non-sarcastic collections, data is typically obtained using the Twitter Streaming API. Each Tweet derived using the API contains comprehensive information about users, including user identity, URL, username, user name, account information, and text of tweets [23]. The text in tweets is the key text data analytics as it includes behavioural, emotional and other types of information and ideas. This data is utilized to create a collection of features that allows Twitter data to be effectively categorized.

b. Data pre-processing: The noise in the data is a significant drawback of using Twitter data sets. Tweets on Twitter can include plain text, mentions of other users (@user), URL references, and content tags, which are also known as hashtags (#). Pre-processing of satirical and non-satirical data is done in this step in order to prepare them for the following tasks. This stage involves a number of algorithms to remove noise from the satire dataset, which contains retweets, duplicates, numerals, tweets in many languages, and tweets with a single URL. These distractions are eliminated because they don't help with improving the data classification accuracy. Many basic pre-processing methods, such as tokenization, stop word removal, spell check stemming, and lemmatizing, as well as POS tagging, are utilized when text input is changed to lower case.

- Tokenization: It is the process that breaks the sentences into smaller parts called tokens, which can be phrases, words or symbols which are valuable in and of themselves. Tokenization also eliminates blank whitespace characters that appear in text documents. Each token is a group of characters which is found in a document that come together to form a useful semantic unit that can be used later during analysis. As a result, the token output can be used as an intake for additional analysis [24]. Tokenization work can be done with NLP Toolkit.
- Removal of Stop Words: All human languages consist of plenty of Stop words. By removing these words, the low-level information is removed from the text to bring more attention to important information. Removing stop words certainly decreases the dataset dimensionality and thus lessens the training time owing to the smaller number of tokens used in training.
- Spell correction: It is a procedure of verifying the spelling of text to correct misspelt text. A common tool to correct any misspelt words is the PyEnchant spell checker python package.
- Stemming: Stemming is to restore derived words to their original form or to obtain the root word named stem by eliminating prefixes and suffixes from the word. Stemming decreases the number of keyword spaces and makes classification better when a keyword is derived from the keywords of a different category.
- Lemmatizing: A derived word sometimes becomes meaningless when prefixes and suffixes are removed from it [25]. Lemmatization is a generalization method that reduces a word's inflexion to a dictionary form by morphological and lexical analysis. As a result, the Lemmatizer fills in the missing characters in the stemmed word to make it comprehensible. This procedure reduces the term to its simplest versions. Lemmatization does not create the word stem, unlike stemming. Instead, it replaces the input word's suffix with a new word to produce its generalized form. The term 'to walk,' for example, can be linked to the word 'walking,' which can then be lemmatized to 'to walk.'
- POS tagging: The POS Tagger software scans text documents and assigns parts

of speech to all tokens on the basis of their definitions. Tagger assigns verbs, nouns, adverbs, adjectives, conjunctions, interjections, and other components of speech. Fine-grained POS tagging is required for a wide variety of computational science applications. Noun tagging, for example, can take many forms, including singular nouns, possessive nouns, and multiple nouns. Different notations are used by POS Tagger [26]. Singular common nouns are denoted by NN, plural common nouns are denoted by NNs, and singular proper nouns are denoted by NP. POS Tagger, on the other hand, tags items using stochastic and rule-based algorithms.

c. Feature engineering: Feature extraction is important in determining the result of a machine learning operation. Feature engineering can be considered a crucial process in the phase of text classification. The performance of the classification, both considering qualitatively and quantitatively is contingent upon the chosen characteristics. The feature engineering phase aims to extract features with discriminatory ability from the processed data so as to separate sarcastic and non-sarcastic text. Sarcasm has many applications as it is a sophisticated form of speech. Interpreting the data, the interpreters conclude that these motives mostly fall into three categories, but not entirely: sarcasm in the form of intelligence, sarcasm in the form of whispering, and sarcasm in the form of avoidance.

- Sarcasm in the form of intelligence: The application of sarcasm in the form of intelligence means being funny. To make it easier to detect, the individual employs certain forms of speech, exaggerates, or has a tone that is different from what he typically does. In social networks, tone of voice is transformed into dedicated forms of writing, for example, using capitalized words, exclamations and question marks along with some satirical emoticons [27].
- Sarcasm in the form of whispering: Satire in the form of a whisper is used to demonstrate how angry a person is. Hence, it attempts to display the level of trouble based on exaggeration or employs very positive expressions to define a negative case.
- Sarcasm in the form of avoidance: It denotes a person who uses sarcasm because he or she does not want to give a clear response. The person in such a situation makes use of complex sentences, unusual words, and some rare expressions.

d. Classification: The categorization models, sometimes referred to as machine learning algorithms are the focus of this step. Machine learning examines the learning algorithm before making predictions based on the data. Following the feature extraction phase is generally referred to as the model training step. Using features obtained from the dataset, this stage creates machine learning algorithms. The developed models are also applied to the classification of sarcasm. In other words, sarcasm and non-sarcasm are classified in tweets using machine learning algorithms. The best classifier is chosen by comparing a large number of classifiers using various sarcastic prediction tests. The most used classifiers for sarcasm prediction are random forest, decision tree, support vector machine, logistic regression and naïve Bayes etc.

i. Decision tree (DT): Being a learning algorithm, the Decision Tree uses a tree-shaped model for decision making. This model has no parameter; however, it is quite easy to manage the interaction of features. In an asymmetrical scenario, the classifier is focused on a rule that represents the DT obtained from a disorganized class. For example, classification using a sorting algorithm [28] is dependent on the feature value. Edges, leaf and decision nodes (representing each element to be categorized) and branch nodes make up the tree (representing the value which the node can do). Instance categorization starts with the root node and is based on the value of its feature. The fundamental flaw of this classifier is overfitting, which occurs as a result of its capacity to fit all regions of the data as well as noise, resulting in poor performance. Overfitting can be avoided by employing a multi-classifier model like the random forest.

ii. Random forest (RF): Unlike single classifiers, Ensemble classifiers have recently drawn more attention because of their strength and accuracy to noise. A random forest [29] is a strong ensemble that incorporates many decision trees. The idea of combining numerous classifiers offers a random forest with more characteristics that set it distinct from typical tree classifier models. RF classifiers, like a single decision tree classifier with outliers or noise that could impact a model's overall performance, provide randomization to address such problems. Both the data and the characteristics are random in Random Forest. The same concepts as in bootstrapping and bagging classifiers are used in this classifier. This is accomplished by enabling the trees to develop from different training data subsets collected by bootstrap aggregation, so enhancing their diversity.

iii. Support vector machine: A support vector machine (SVM) is one of the binary linear classifiers. It uses a larger size space to build a collection of hyperplanes because it is a non-probabilistic supervised learning technique. The main aim here is to divide the data into several classes using training data. However, the target value is predicted by constructing the model using the training data. This data contains only the features of the test data. SVM is one of the most popular employed text classification algorithms. It selects the best hyperplane for the appropriate classification of problem cases.

iv. Logistic regression (LR): Logistic Regression is one of the linear predictive models. It defines an event's probability as a linear function of a set of predictor variables. The decision boundaries are generally created using a linear function of the features in this technique. The goal of logistic regression is to help find document class labels by supplementing the probability function. The parameters are chosen in order to obtain the highest conditional probability. Despite the positive findings of LR, the majority of the time, the class formed is outside the variable (0–1) and does not suit the probability range [30].

v. K-nearest neighbor (KNN): KNN is a type of machine learning model which is built on examples. The k-nearest neighbour of each sample determines the identity of the class label for that instance in this algorithm. As a result, the majority voting notion is used to establish the class label in the nearby example.

vi. Naïve Bayes (NB): Naïve Bayes is a type of probabilistic classification model. This classifier is completely based on the concept of the Bayes theorem. This classifier makes assumptions about the features about their strengths. It can work on a lesser amount of training data for the parameter computation for carrying out the prediction task [99]. It is considered the major benefit of the approach. Due to the feature's independence, it computes the variance of the feature merely instead of computing the overall complete covariance matrix. For a provided textual review data 'd' and for class 'c' positive or negative), the calculation of the conditional probability for every class provided a review is P(c|d).

It is expressed as:

$$P(c|d) = \frac{P\,(d|c) * P(\,c\,)}{P\,(d)} \tag{1.1}$$

1.6.2 Common Features of Sarcasm Detection

Deriving data features is a critical step in the text mining process since it helps classification algorithms reach their conclusions. When categorizing text messages, one might employ specific social media post characteristics as a vital factor for sarcasm detection. As a result, creating a data set with the right attributes will significantly increase machine learning's overall productivity. Using different text mining techniques can have various properties. Each classification system uses the following features as key attributes to identify sarcastic tweets:

a) Sentiment-related feature: Whisper is the widely used sarcasm type available on social media. In a whisper, composers of sarcastic accents use positive emotion to define a negative case [31]. Sarcasm, by the way, takes advantage of the contradicting emotion that may be found in the portrayal of a bad instance through positive emotion.
b) Pragmatic features: Practical characteristics are represented using symbolic and metaphorical texts. (e.g., smiles and other kinds of emoticons). These are most common in tweets, particularly because of the limited length of tweets. Practical features are a strong sign of sarcasm detection on Twitter. Hence, many researchers have derived these features to use them in the sarcasm classification operation.
c) Frequency-related features: These are the most commonly utilized characteristics in a document or text corpus. It shows the significance of a particular word in a document or collection. Extraction of frequency-related characteristics is a crucial task. These characteristics can be used in a variety of ways to characterize sarcasm.
d) Term Frequency- Inverse Document frequency: TF-IDF is a numerical statistic representing the significance of a word (period) for a document in the corpus. In TF-IDF, a comparison must be made between the frequency of a word in a document against its number in other documents. TF-IDF is commonly employed to prevent filtering of words in text summarization and classification

applications [32]. It can also be used to proportionally increase the count of times a word seems in a document.

e) Hashtag features: Users sometimes use hashtags in their tweets to convey their emotions. Sometimes, the user expresses emotional content using hashtags. These are also used to illustrate the true purpose of conveying the message. For example, in a tweet, "Thank you so much for always helping me, #I hate you." In this statement, the hashtag "# I hate you" suggests that the user is expressing thanks for not really wanting but hating it so much for not helping when needed. Hashtag features can be positive or negative hashtags. The said expression is a tweet with a negative hashtag.

f) Lexical features: Lexical features are frequently used in text mining. To represent the range of polarity, these attributes include distinct words, phrases, noun phrases, or named objects associated with a score. The use of these attributes for emotion-mining can help decide the level of emotion in a text.

g) Stemmed features: The notion behind stemming is that words with the same root have similar meanings. The stemming process aims to effectively reduce the words in the world list. One can use stemmers to combine words which not only decreases the size of indexing files but also increases the recovery performance [33].

h) Ambiguity: Using a word with several meanings increases the likelihood of it being used in an ambiguous manner. To detect sarcasm, a number of variables must be computed, including the maximum number of synsets associated with a word, the average number of synsets over all words, and the synset gap (for example, the difference between two earlier features).

i) Synonyms: It's a method for determining attributes from words with similar meanings. This method appears to be effective in identifying sarcasm when expressing a particular viewpoint in a variety of ways. As a result, in the sarcasm detection task, identifying synonyms is accomplished by retrieving a list of synonyms per word and then computing the highest/lowest number of synonyms with a frequency larger than the one in the tweet (in all terms in the tweet).

1.7 Ensemble Learning

A classifier ensemble is just a group of classifiers where the results of each are added together to produce a final conclusion. The approach's ultimate goal is to combine the output of different models, also known as base classifiers, into a conclusion that outperforms each of the individual baseline classifiers. Generating a number of base classifiers is the initial stage in building a classifier ensemble. One option is to create N, multiple classification models, using N different learning techniques and a single training data set. A different approach is to divide the training data set into N pieces and add one learning algorithm to each of them. It's crucial to choose a method that enables the development of numerous classifiers over the course of the learning phase. It has additionally been demonstrated that good basis classifier selection would influence overall classification accuracy rather than merely integrating all of the base classifiers into a single ensemble. The basic classifier can be chosen either statically or dynamically. In the static approach, the same subset of basic classifiers is applied to all study samples. Each new instance is chosen separately in the dynamic strategy.

To reach a final output judgement, the outputs of the baseline classifiers must then be integrated. The most crucial issues to be resolved during this phase are what sorts of knowledge will be blended and the blending mechanism to be employed. In response to an undefined pattern, both base classifiers produce judgments that are then passed back to a combination function. Different methods of blending make use of various base classifier outputs, such as class mark or class probability distribution. Using forecasts as a list of attributes to train a mix feature is another method for meta-learning. The best ensemble solution is built using a method that combines an acceptable mixture scheme and a comprehensive list of the basis classifiers.

When analyzing supervised learning algorithms, the expression "bias, noise and variance decomposition of error" is widely used. The learning algorithm error of a classifier can be split into three categories, according to the review: bias, noise, and variance. These properties are derived when a model's performance is conditioned by varied training data. The term "data noise" refers to an error that occurs independently of the learning procedure. Calculating the average error of a learner prepared with several sets of training data is how bias against a certain input is defined. Variance is a measure of how varied a learner's forecasts are when provided with diverse learning data.

1.7.1 Ensemble Generation

Making a list of various baseline classifiers is the first stage in establishing the classifier ensemble. One strategy would be to choose from N different learning methods. According to the method, each baseline classifier is created using a different learning algorithm but the same set of training data. As a result, N distinct classification models are available for selection. The following stage is integrating their findings to determine the final outcome.

The second option is to build a series of baseline classifiers using a single learning strategy and many training sets. The main problem with this strategy is quickly converting the initial data set into a series of other training data sets. The real data set was divided into N-subsets using a variety of techniques, including random sorting and clustering. Altering the way that data is distributed is the alternative. Each baseline classifier is constructed using the same learning algorithm and the various subsets once the subsets have been developed.

Partitioning of training data set: Bagging is the most popular method for separating multiple training sets from a single data source. Using bootstrap methods, the training sets are selected at random k times from the original data selection. Using this style, some occurrences are likely to appear more than once in some training datasets, while others might never do so. As a result, K-training datasets are constructed, each with the same amount of data (k different classifier classes). The most important advantage of bagging is that it allows for the independent and combined practice of several band configurations, which shortens the preparation period.

Manipulation of data distribution: When employing the same learning paradigm, boosting is a strategy that uses multiple training data sets. Boosting is a form of iterative procedure in which the distribution of the training range is constantly altered in response to the classifier's accuracy. Until a base, the classifier has been developed and applied to the ensemble, both cases that are successfully or mistakenly labelled as weight growth or loss are reweighted. Each base classifier's prediction receives a weighted vote in the final prediction. The weights are based on each classifier's accuracy rather than how well it performed throughout training. AdaBoost is a popular and powerful booster.

Partitioning of the attribute space: The second option is to compile a list of the base classifiers using different feature subspaces and the initial training data set. A popular technique for selecting ensemble members from random subspaces is Random Forest. A classification scheme made up of numerous distinct trees is known as a random forest. The same process is used to render individual bags to create each tree. The input parameters determine the size of the ensemble and the number of variables utilized to determine the break at a tree node. Each tree is built using a single bootstrap sample. Each node's decision-making variables are selected at random. A voting method is used to reach a decision after the formation of a tree body. The Random Forest can be called a variation of the bagging strategy to some degree.

1.7.2 Majority Voting

As a result of bagging, boosting, or some other technique for constructing an ensemble, a selection of basic classifiers is produced. During the creation of a classifier ensemble, the effects of every classification model are merged. The MV method is the simplest technique for combining base-level classifiers. The class with the most votes is chosen as the final classification using this protocol, which gathers the votes of all classifiers. We assume that every model votes for the class that has the highest likelihood when we employ probabilistic classifiers. An improved version of this technique is weighted majority voting (WMV). In this process, various weights are assigned to the base classifiers. Each classifier's weight reveals how important it is to the final result.

1.8 Research Gaps

A detailed literature survey was conducted on the sentiment analysis. Following research gaps have been identified.

- **Gap 1:** There is limited literature specifically aimed at detecting sarcasm in textual data. **It is encouraged to conduct an in-depth and complete study of the area.**
- **Gap 2:** The majority of current studies can determine whether a sentence is positive or negative in polarity, but it is insufficient to define any sentence as sarcastic or not. **As a result, a solution to this research issue is required.**
- **Gap 3:** All of the methods available have used feature-based approaches on manually annotated or unannotated data, resulting in low to average textual data accuracy. **It is required to focus on the overall context of the text, which is one of the vital factors for detecting sarcasm.**
- **Gap 4:** Various machine learning and lexicon-based techniques are used for detecting sarcastic text. Still, there is a scope for further improvement in accuracy. **An ensemble approach is proposed to address the issue.**

1.9 Problem Statements

When it comes to academic and business endeavors, electronic documents are an important aspect of receiving and transmitting the information. There's no point in having access to web data if we can't extract it and apply it to our businesses. As a result, state of the art in sentiment analysis has identified the following problem statements, which can be regarded as part of the research work.

Problem statement 1: Different formats and types of data are being collected by various social networking sites. How can we convert it into a structured form of data?

Problem statement 2: As there are present different types of sentiments in the text. How to identify the correct type of sentiment expressed in the text?

Problem statement 3: Detection of sarcasm is a challenging task as it affects the performance of the system.

Problem statement 4: What are the different performance parameters available to find the effectiveness of the sentiment analysis system?

Problem statement 5: Is there any effective way to improve the sentiment analysis system's performance?

Therefore, a strategy to detect sarcasm in the textual data is required to enhance the performance of the sentiment analysis system in terms of the performance parameters such as accuracy, precision, recall and f-score. This research work focuses on designing an ensemble classifier for the system.

1.10 Motivation

In the field of text mining, sentiment analysis is a common topic of study. This study provides a quick overview of recent developments in this area. Several recently proposed algorithm enhancements and numerous Sentiment Analysis applications are also reviewed and described.

The following facts are extremely motivating:

• The importance of comprehending different sentiment analysis concepts.

• Understanding the notion of sarcasm and various ways to detection, as well as building an ensemble approach to detect sarcasm, are required.

As a result, detecting sarcasm in the text would improve the sentiment analysis model's efficiency. Analyzing existing strategies and understanding the work of other researchers, on the other hand, will aid in the development of a new methodology that will improve on existing techniques for detecting sarcasm.

1.11 Objectives

Looking at the current scenario, where people are exchanging their thoughts, views, reviews etc., on the social network. It has become an important field to be looked upon by researchers, as the textual data which is available can be used for various types of studies. Analyzing the sentiments in the textual data has been one of the booms in the field of analytics. The presence of sarcasm in the text affects the performance of the system. Thus, in this research, the main objective is to study the various techniques to detect sarcasm and design an ensemble classier to detect sarcasm and improve the performance of the sentiment analysis system.

To achieve these goals, the research work is divided into multiple research objectives as mentioned below:

1. To study and analyze various sarcasm detection techniques for sentiment analysis.
2. To design and implement an ensemble classification method for sarcasm detection.
3. To test and validate the proposed method and compare the result with existing approaches.

1.12 Research Methodology

The purpose of this research is to recognize sarcastic text while conducting sentiment analysis. Detection of Sarcasm in the textual data will be done by designing and implementing an ensemble classifier.

Various stages are:

1. Social networking data will be collected from multiple domains.
2. This data will be pre-processed to convert into a structured format.
3. Features will be extracted from the pre-processed data.
4. The training dataset will be trained using a variety of classifiers.
5. The ensemble classifier will be generated later using the majority voting.

6. The ensemble classifier will be trained on the training dataset, and results will be predicted on the test data set.

   Workflow is represented in Figure 1.4

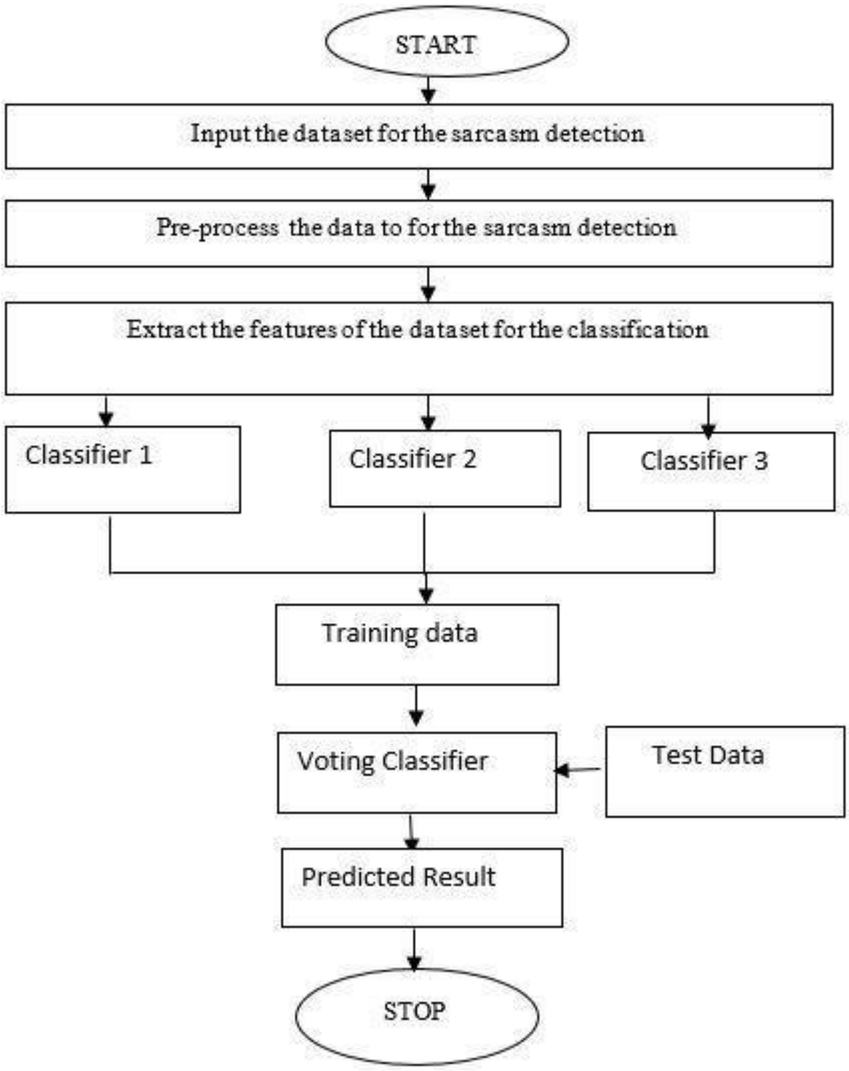


Figure 1.4: Methodology Workflow

1.13 Organization of thesis

Chapter 2 presents the review of literature which is required to know about the existing state of the art for the chosen problem statement. This chapter first highlights the existing research studies related to sentiment analysis. Later research related to sarcasm and ensemble learning is also presented.

Chapter 3 presents the sarcasm detection technique based on machine learning algorithms. For sarcasm detection, four different ensemble classifiers are designed.

In chapter 4, we have compared the performance of existing classifiers with the proposed approach on the data set for the detection of sarcasm. Various evaluation parameters considered were accuracy, precision, recall and f-measure.

Chapter 5 concludes the entire research work done. This chapter also helps the novice by proposing the future aspects of the research.

CHAPTER 2

LITERATURE REVIEW

This section explains the current information related to sentiment analysis by conducting a theoretical and methodological study. Firstly, a detailed review is being undertaken on the sentiment analysis. Later, a significant issue of sentiment analysis, sarcasm, has been studied intensely and described in detail. Finally, the whole literature review has been summarized at the end.

2.1 Literature survey related to sentiment analysis

Sentiment analysis is a type of text mining that identifies and extracts subjective information from data and recognizes the emotional tone hidden beneath the text's body. It is a very popular way used by various organizations and businesses to understand and categorize the opinions about the product, service or idea by monitoring their online conversations. Thus, researchers are building a great interest in this area. In this section, a detailed review related to sentiment analysis is provided.

In [34], the author suggested a pre-processed data model on the basis of NLP (natural language processing) for filtering the tweets. Thereafter, the motion of BoW (Bag of Words) was integrated with TF-IDF (Term Frequency-Inverse Document Frequency) with the objective of analyzing the sentiments. These methods were effective in accurately identifying tweets as positive or negative. The accuracy with which the sentiments were analyzed was maximized using the TF-IDF vectorizer. The simulation outcomes depicted that the suggested model was effective and provided an accuracy of around 85.25% while analyzing the sentiments.

The author in [35] introduced W2VC (word2vec and clustering) based text representation technique in order to analyze the sentiments of Twitter. The SVM (Support Vector Machine) algorithm was implemented to classify the sentiments. Two diverse datasets containing Turkish Twitter feeds were utilized to perform experiments. And the experimental outcomes demonstrated the applicability of the introduced technique with regard to time and performance. Moreover, this technique was capable of mitigating the feature space.

In [36] author projected an iterative algorithm recognized as SentiDiff for predicting the sentiment polarities which were expressed in Twitter messages. The inter-relationships of textual information in Twitter messages with sentiment diffusion patterns were taken into consideration by this algorithm, which aided in improving Twitter's SA (sentiment analysis). A real-world dataset was applied in the experimentation. The experimental outcomes validated the supremacy of the projected algorithm over existing algorithms and provided the PR-AUC up to8.38% while analyzing the sentiments on Twitter.

The author has developed a model to analyze the sentiment with the help of R software [37]. The Twitter API was used to examine the sentiments of Twitter users. In this model, the data was gathered from Twitter and pre-processed later on. Afterwards, the sentiment of the user was analyzed using a lexicon-based technique. The acronym dictionary was created to replace the acronym word, and the emoticons were detected in the tweet. The document level, as well as aspect level, was analyzed to make the decision. The future work would aim to conduct a comparative analysis, make an attempt to utilize ML (machine learning) techniques and construct a hybrid mechanism to analyze the sentiments.

In [38] author has presented the outcomes attained while analyzing the sentiments on Twitter that UK energy customers had posted. The functions taken from two sentiment lexica were integrated with the optimization of the accuracy of outcomes. The primary lexicon helped in extracting the sentiment-bearing terms and negative sentiments due to its efficacy in detecting these. The remaining data were classified using a second lexicon. The experimental outcomes indicated that the presented approach led to augmenting the accuracy of the outcomes in comparison with the common practice of using only one lexicon.

The author has designed a model for visualizing the raw tweets with scalability and efficacy in [39]. This model focused on obtaining the sentiments of people and visualizing them for better understanding. The tweets were taken in real-time using Spring XD. After that, the conversion of raw tweets was done into HDFS (Hadoop Distributed File System). The tweets were refined and labelled using HIVE (Hadoop Scripting Language) with regard to their respective sentiments. Finally, HIVE was used to run simulations on the created model. This approach proved effective in classifying positive, negative, and neutral sentiments. The designed model generated optimal outcomes while analyzing the sentiments.

In [40] author has established an SVM (Support Vector Machine) algorithm in which the n-gram-based internal attributes were integrated with an external sentiment vector for enhancing the traditional n-gram classification algorithm. An ML (machine learning) classifier was deployed to classify the sentiments on Twitter as various textual attributes were utilized along with n-grams of Twitter data in this classifier. Three diverse weighting techniques were adopted to understand the effect of weighting on accuracy. In addition, additional information was obtained from a sentiment score vector of tweets to enhance the efficiency of the established algorithm.

In [41] author intended an NB (Naive Bayes) algorithm to classify the sentiment of Twitter data with the outcomes of ML (Machine Learning). Though this algorithm consumed much time, it assisted in performing estimation inefficient way. This algorithm was implemented to extract and pre-process the tweets and classify them into three sentiments later on. The intended algorithm was useful to recognize the efficient mobile operating systems in accordance with the sentiments of social media users. The outcomes of the evaluation revealed that the intended algorithm was adaptable.

The author has formulated a popular ML (machine learning) technique to classify the text of Twitter posts in positive or negative sentiment [42]. Moreover, a trained technique was constructed through tweets to which labels were assigned. This technique made the deployment of external lexicons for detecting the tweets as subjective and objective. The attributes were filtered by means of TF-IDF (Term

Frequency-Inverse Document Frequency). With the goal of analyzing the impact of public mood toward gunexpected events, the tweets were mined, filtered, and processed using Twitter Streaming API and some of the official World Cup hashtags. Such a technique for analyzing the sentiments assisted in utilizing the Twitter data to extract the patterns on the basis of opinionated texts.

The author devised a novel mechanism to analyze the sentiment on the basis of Domain Specific Ontology in [43]. Oman tourism ontology was produced in accordance with the ConceptNet. POS (Part of Speech) tagger was implemented to recognize the entities from the tweets and to conduct a comparative analysis of them. Moreover, an integrated sentiment lexicon technique was employed to determine the sentiment of the extracted entities. In the end, the incorporation of the semantic orientations of attributes available in a particular domain was done. The conceptual semantic was considered as an attribute and incorporated with an ML (machine learning) algorithm for improving performance while analyzing the sentiments of Oman tourism.

In [47] author has suggested a phrase-level analysis of tweets in which Doc2vec (document vectors) were utilized. The ensemble methods, namely AdaBoost, were comprised of the classifiers, which were further generated by one classification algorithm and yielded an accuracy of up to 84.5%. The accuracy obtained from these algorithms was found to be higher. The airline industry took advantage of these accuracies to execute the investigation in accordance with customer satisfaction. The number of tweets was maximized to develop an effective framework that provided superior accuracy. The suggested approach was useful for airline companies as it efficient analyzed the Twitter data.

The author introduced the Word2Vec model to classify the labelled data in English, and Turkish Twitter fed and discovered the impact of feeds on the Word2Vec model [48]. Two datasets such as English and Turkish, were utilized in this model. The tweets having roots or without roots were extracted using BoW (Bag of Words) and Word2Vec models. Python was applied to implement the introduced model. A comparative analysis was conducted on the introduced model to prove its efficacy.

The author in [50] intended an innovative framework to analyze the social networks' sentiment on the basis of TSS (Twitter sentiment score) [50]. The intended TSS framework generated a novel baseline correlation technique to provide higher predictive accuracy as well as alleviate the computation burden. This framework was adaptable to make the decision without any deployment of information related to historical data. The future stock market trend was predicted earlier using this approach, and obtained accuracy was calculated at 67.22%. The intended framework yielded an accuracy of around97.87% with the help of LR (logistic regression) and LDA (linear discriminant analysis) to predict the upward trend of the future market.

The author designed a CNN- LSTM (Convolutional Neural Network-Long Short-Term Memory) based DL (deep learning) technique with a pre-trained embedding technique that learned the way of extracting the attributes in an automatic manner so that the sentiments of reviews or opinions were analyzed and classified into two polarities: positive and negative [52]. An effectual DL model was deployed along with tuned hyperparameters on CNN layers prior to implementing the Bi-LSTM (Bi-directional LSTM), which had long-range dependencies. The designed approach yielded an accuracy of around 81.20% in comparison with conventional techniques. The future work would emphasize discovering other sentiment twitter datasets and hybrid DL models to analyze the sentiments.

In [54] author recommended the SegAnalysis model for segmenting the tweets, detecting the event and analyzing the sentiments. The POS (part of speech) tagger and the currently posted tweets of users have segmented the tweet in a batch mode. This process led to preserving the named entities and computing their stickiness score. The events were detected using NB (Naïve Bayes) and online clustering. These events were aided in enhancing the awareness related to circumstances and decision support. On the basis of a tweet's sentiment score, the SA (sentiment analysis) procedure was used to classify the tweets as positive, negative, or neutral. The recommended model was further expanded for handling the events related to multiple clusters.

In [55] author established an approach called VADER so that the sentiments were analyzed. A technique was utilized in text mining to clean several unusual things. Moreover, data considered as material in the analysis was saved so that the data was read easily. The sentiments were analyzed to map and count the tweets posted by users. The outcomes revealed that the established approach achieved predictions with regard to actual outcomes of the US presidential election.

The author in [56] devised a technique in which tweets were pre-processed, and attributes were extracted to generate an effective attribute. Thereafter, these attributes were scored and balanced under various classes. The sentiments were analyzed and classified from the tweets using SoftMax, SVR (Support Vector Regression), DT (Decision Tree) and RF (Random Forest). A Twitter dataset obtained from NLTK corpora resources was utilized to simulate the devised technique. The outcomes exhibited that DT had generated superior outcomes in comparison with other algorithms.

In order to examine Twitter data collection, the author in [57] devised a SA (sentiment analysis) technique. This technique used the NB (Naive Bayes) algorithm to categorize the text as positive or negative polarity. The dataset used various linguistic and NLP (natural language processing) pre-processing approaches. These methods concentrated on examining the impact on the quality of classifying the big data. The accuracy of the constructed technique was enhanced using these methods. The experimental outcomes indicated that the created technique led to enhance the efficacy of analyzing the sentiment by 5% with the help of NLP and linguistic processing and offered an accuracy of around 73%.

The author recommended the Jen-Ton model in [58] for analyzing online data from the social networking site Twitter. The main purpose of this model was to collect, pre-process, and assess the sentiment of tweets, as well as the sentiment of a product.

To increase the accuracy of the aspect-based sentiment analysis utilizing clustering with a Genetic Algorithm, researchers applied IMS (Imputation of Missing Sentiment), ASFuL (Aspect based Sentiment Analysis using Fuzzy Logic), and ASTA-CGA. The proposed approach aided in improving the company outcome, with public opinion playing a vital role in its development.

The author [60] proposed a method for classifying the sentiment of tweets in an automatic manner through the integration of ML (machine learning) classification algorithm with a lexicon-based approach. SentiWordNet, NB (naive Bayes) and HHM (hidden Markov model) were comprised in this approach. The majority voting principle was implemented on the result obtained from these algorithms to determine the tweets as positive or negative. The sentiment classification algorithm was utilized to discover the political sentiment from real-time tweets. Hence, the introduced ensemble approach was capable of enhancing the accuracy while analyzing the sentiments.

In [61] author developed the word2vec model to analyze the sentiments from the real-time Twitter data of the 2019 election. An RF (random forest) algorithm was deployed to classify the sentiments. The developed approach assisted in maximizing the accuracy of analyzing the sentiments in contrast to conventional techniques. The Word2vec model considered the semantics of words in a text to augment the quality of attributes, due to which the accuracy of ML and SA was enhanced. The developed model achieved an accuracy of around 86.8%.

The author in [62] emphasized computing the sentiments posted in tweets and comparing this sentiment with polling data to find correlations shared among them later on. A lexicon-based technique and NB (Naive Bayes) algorithms were suggested with the purpose of calculating the sentiment of political tweets whose collection was done 100 days before the election. The labels were assigned to the tweets not only manual but also in an automatic way on the basis of hashtag content. The outcomes proved the reliability of the suggested approach over the existing technique.

SVM (support vector machine) and TF-IDF(Term Frequency-Inverse Document Frequency) were developed by the author in [66]. The data crawling procedure was utilized to extract the data from Twitter with the keywords gojekindonesia and grabID. Three classes were included in this dataset: positive, negative and neutral. Five diverse scenarios were considered to classify the sentiments. Moreover, four kernels were utilized to execute the classification procedure. The training phase employed 90% of the data, and the rest of the data was used in testing data. The devised technique provided an accuracy of 80%.

In [67] author recommended the Hadoop framework for analyzing the sentiment of Twitter data successfully from an enormous volume of data. This framework assisted in the analysis of Twitter data with Map Reduce. The Hive ODBC driver was adopted for creating a graph in Excel and proving the real orientation of the data. The tool flume was applicable and useful for maintaining the log and the data set in, much alike scoop. The future work would focus on extending the analysis up to a more prolific linguistic analysis.

The author of [69] predicted that DICE (Deep Intelligent Contextual Embedding) would increase tweet quality by removing noise from contexts. Then, for polysemy in context, semantics, syntax, and sentiment interpretation of words in a tweet, four embeddings were combined. For evaluating the emotion of a tweet, the BiLSTM (Bi-directional Long Short-Term Memory) network used a projected model. When assessing the sentiment from airline-related tweets, the testing results demonstrated the superiority of the projected model over many methods.

In [72], the authors focused on storing Twitter Streaming Data in Hadoop HDFS using Flume and then extracting them using Apache Hive, as mentioned. Thereafter, the sentiment was decoded in this data on Apache Mahout using ML (Machine Learning) classifiers. An innovative hybrid approach was developed in which NB (Naïve Bayes) was integrated with DT (Decision Tree) to improve the SA (sentiment analysis) of streaming Twitter data. The accuracy of the developed technique was calculated at 86.44% and found to be higher in contrast to the NB algorithm.

In [73] author developed SVM (Support Vector Machine) with the purpose of analyzing the sentiment. The data was initially saved using the Twitter API and keywords. Following that, the data was collected and pre-processed before extracting the attributes for each tweet. A feature list gathered all the attained attributes. Finally, the feature list was converted into a binary feature vector using the Term Frequency-Inverse Document Frequency technique. The test outcomes indicated that the developed approach provided an accuracy of around 98%. Moreover, the SVM technique proved a promising tool for classifying the sentiment of JNE Twitter data.

The author of [74] proposed a method for reducing textual input to a CNN (convolutional neural network) based on the MSP (morphological sentence pattern) model. Instead of using aspect-based prediction, this method used broad sentiment prediction. The input was trimmed to the most significant words according to sentiments in order to transmit the pre-processed data to the CNN (Convolutional

Neural Network) for the training phase. The output obtained from the designed method was optimal for analyzing the sentiment analysis and also useful to enhance the performance.

In [75] author projected a new SA (sentiment analysis) on the basis of DL (deep learning) mechanism for extracting the sentiment from social media. The collection of data was done to generate a dataset. When these special terms were processed, a semantic dataset was created to conduct further research. The extracted information proved valuable for various future applications. Diverse social media platforms were utilized to acquire the experimental data. The future work would emphasize constructing a model based on a projected model for analyzing the crawled intelligence from our review dataset.

In [76] author emphasized implementing the SVM (Support Vector Machine) algorithm for classifying the sentiment and texts for smartphones so that diverse datasets were analyzed to classify the sentiments and texts. Moreover, the training as well as testing phase, was executed on diverse datasets. The presented algorithm was utilized to find the polarity of the ambiguous tweets. The experiments were conducted on the implemented algorithm with regard to three metrics - precision, recall and F-measure. The results validated the implemented algorithm offered higher accuracy.

The author presented an approach to analyzing the sentiments in Twitter data. First, API twitter was considered with every candidate named on Jakarta Governor Election to gather the data [77]. This data was utilized as an input in the pre-processing stage. The second stage focused on extracting the attributes for which a list was created. The conversion of this list was done into a feature vector in binary form, and TF-IDF (Term Frequency-Inverse Document Frequency) was utilized to convert it again. The dataset was divided into two sections: training and testing. The presented approach was tested using k-fold cross validation. The testing outcomes revealed that the accuracy of the presented approach was computed at 74%.

In [78] author constructed a novel system in order to enhance the efficiency of analyzing the sentiments of twitter. The twitter sentiments were analyzed on twitter-sanders-apple 2 datasets. In particular, noises were included in the raw tweets whose removal was required for classifying the sentiments. At last, a hybrid classification algorithm was put forward in which ANFIS (Adaptive Neuro-Fuzzy Inference System) was integrated with GA (Genetic Algorithm) with the objective of classifying the Twitter sentiment as positive and negative. The outcomes of experiments confirmed that the constructed system was capable of boosting the accuracy of analyzing the sentiment by 5.5-6% as compared to the traditional technique.

In [79] author investigated two classifiers known as BPNN (Back Propagation Neural Network) and SVM (Support Vector Machine). The sentiments of tweets were classified in order to acquire promising accuracy in the classification outcomes. The predictive labels were compared with the actual labels using the resulting confusion matrix. In comparison to other algorithms, the SVM method fared better in the experiments.

Using predictive and descriptive approaches, the author of [80] attempted to analyze the sentiment of DKI Jakarta's gubernatorial election in 2017 on social media Twitter. The username of every candidate was utilized as the search query to collect the dataset from Twitter. ML (machine learning) techniques, namely MNB (Multinomial Naive Bayes) and SVM (Support Vector Machine) were utilized in the predictive approach for classifying the dataset. The time series graphs and word clouds were adopted in a descriptive approach for achieving deeper insights into the dataset and discovering the link between the sentiment of Twitter with the result of the election itself.

In [82] author formulated a SmartSenti application to analyze the sentiments on Twitter posts related to tourism centres. For this, ML (machine learning) and TL (transfer learning) techniques were utilized. The collection of tweets on a touristic place in Turkey was done. When the data was pre-processed, and attributes were extracted, this method was used to classify the tweets as positive or negative. The formulated application was adopted to share the outcomes in a visual manner with users. The formulated approach was applicable for enhancing any services of tourism centres for more positive reviews.

In [83] author introduced word embedding models, namely Word2Vec and Glove, for detecting the sentiment polarity in tweets using DL (deep learning) techniques. The RNN (Recurrent Neural Network) model was implemented with LSTM (Long-Short

Term Memory) to analyze the sentiments to handle the long-term dependencies. To achieve this, memory was established in a network model to perform prediction and visualization. The outcomes demonstrated that the introduced model provided higher accuracy, and its reliability was proved for analyzing the sentiments. Moreover, the Bi-LSTM (Bidirectional- Long-Short Term Memory) was established to enhance this performance.

In [84] author designed a technique to classify the sentiments on Urdu news tweets. Initially, the data was pre-processed in which hashtag and stop words were eliminated. Subsequently, all kinds of words and POS (part of speech) tags were recognized to construct the feature vector. The DT (decision tree) was utilized as a classifier. The experimental outcomes of the designed technique indicated that the designed technique was effective for analyzing the sentiments with regard to accuracy.

In [85] author described that the technique of selecting attributes of each score word was implemented to examine the attitudes expressed in Twitter tweets. The effective attributes were selected using NB (Naive Bayes) algorithm in order to train and test the features of words and also to compute the sentiment polarity of every tweet. The presented algorithm was evaluated concerning diverse metrics such as accuracy and precision and its comparison was done with various ML (machine learning) techniques.

The author in [87] aimed to extract the sentiment from Twitter at which user posted their views and opinion. The sentiments of tweets were analyzed to assist business intelligence. As a result, a Hadoop model was developed to process the movie data set gathered from Twitter in the form of comments, reviews and feedback. The outcomes obtained by analyzing the sentiments were presented in diverse sections and classifying the tweets as positive, negative and neutral sentiments. The projected model generated a fast-downloading technique to analyze the tweets on Twitter effectively.

In [88] author developed a novel model to analyze the sentiment called SLCABG planned on the basis of the sentiment lexicon. In his approach, the CNN (Convolutional Neural Network) was integrated with attention-based BiGRU (Bidirectional Gated Recurrent Unit). The sentiment attributes were improved using the sentiment lexicon. Thereafter, CNN and GRU (Gated Recurrent Unit) were adopted for extracting the major sentiment attributes and context features in the reviews. In the end, the weighted sentiments were classified. The experimental outcomes exhibited that the developed model was adaptable for enhancing the efficiency of analyzing the sentiments.

In [114] author proposed a novel technique that mechanically allotted different scores to all sentiments within a tweet. The sentiments containing maximum scores were used by the proposed approach. The added components were presented in this work. This work analyzed the feasibility of quantification. In this work, a manually labelled data set was used. The algorithmic analysis' results were cross-checked against the human explanation. The proposed method was extremely practicable, scoring 45.9% on the F1 scale.

In [115] author proposed a lexicon-based approach through which sentiment analysis was performed on BBC news articles. The outcomes showed that there were a higher number of positive articles in the business and sports categories and negative articles in the entertainment and technology-based categories.

In [116] author integrated sentiment analysis into a machine learning technique on the basis of SVM classifier. In addition, more reliable and realistic sentiment indexes were constructed in this work by considering the day-of-week effect. According to the obtained experimental results, it was feasible to enhance the prediction accuracy of the SSE 50 Index movement direction up to 89.93 percent. After integrating sentiment variables, the accuracy rate increased by 18.6 percent. In the meanwhile, the proposed approach assisted shareholders in making more efficient decisions.

The author developed a thorough sentiment dictionary in [117]. This dictionary contains the basic sentiment words, field sentiment words, and polysemic sentiment words. Following the addition of these terms, the sentiment analysis accuracy improved. A naive Bayesian classifier was used to determine the text field containing the polysemic sentiment word. The results of the tests showed that the sentiment analysis technique based on a comprehensive sentiment dictionary was viable and accurate.

In [120] author proposed a novel word embeddings technique. Large twitter corpora-based unsupervised learning was used to provide this technique. This method used latent contextual semantic associations and co-occurrence statistical features among in tweeted words. The recommended model's performance was compared to that of the baseline model. A word n-grams model based on five Twitter data sets served as the baseline model. The suggested model beat the other existing model in terms of accuracy and F1-measure to categorize tweet sentiments, according to the results.

The author of [121] suggested a lexicon-enhanced LSTM model. Initially, the sentiment lexicon was utilized to supplement the pre-training of a word sentiment classifier in this model. After that, the sentiment embeddings of words were obtained by this classifier. Non-lexicon terms were incorporated in these embeddings. By integrating the sentiment embedding and its word embedding, it was possible to represent words more precisely. In addition, a novel technique for locating the attention vector in wide-ranging sentiment analysis without a target was given in this paper. This phenomenon may improve the LSTM's capacity to obtain comprehensive sentiment information. The recommended models outperformed the other models offered, according to the findings.

In [122] author stated that mining of opinion was done by sentiment analysis at word, sentence, and document levels. This phenomenon provided sentiment polarities and strengthened the editorials. The fact that sentiment Chinese words described the users' opinions was well-known. However, because of the ambiguity of Chinese lettering, traditional machine learning algorithms were unable to accurately characterize the viewpoint of documents. To tackle this problem, this study suggested using a multi-strategy sentiment analysis technique with semantic fuzziness. The results showed that the suggested technique might achieve a good efficiency rate.

The author in [172] presented a survey of published literature during 2002- 2015 on the various aspects of sentiment analysis. Different machine learning algorithms, natural language processing techniques and applications of sentiment analysis were discussed in detail. It was found that online discussions and political discussions often contain irony or sarcastic sentences, which needs more computational approaches to deal with.

In [170] author presented a comprehensive study and has re-implemented different approaches for Arabic sentiment analysis. It was observed that the transformer-based language model was best with F-scores of 0.69, 0.76 and 0.92 on SemEval, ASTD, and ArSAS benchmark datasets. It was observed that sentiment analysis was highly affected by sarcasm.

In [180] author has detected and analyzed sentiments and emotion expressed by people from the text written in their Twitter posts and have used it for the generation of recommendation for the user based on their Twitter activity. The use of sarcastic text in the text will majorly contribute to the text's sentiment.

The primary goal of [173] is to investigate and analyze the use of sentiment analysis on Twitter using soft computing approaches; secondly, in contrast to previous reviews, we take a systematic approach to identifying, gathering empirical evidence, interpreting results, critically analyzing, and integrating the findings of all relevant, high-quality studies to address specific research questions pertaining to the defined research domain. Emotion analysis, sarcasm recognition, rumor detection, and irony identification have all been mentioned as viable study fields.

In [179], the author describes a novel way to use aspect-level sentiment detection that focuses on the item's features. The work studied the identification of words altering polarity in the presence of context and its effect on the overall evaluation of the product as well as the particular aspect, and the results were impressive. Future work will be done with the goal of using cutting-edge technologies to address more difficult problems such as spam and fake news, negations, sarcasm, and so on.

In [178] author presents a comprehensive systematic literature review to discuss both technical aspects of opinion mining and sentiment analysis and non-technical aspects in the form of application areas. Sarcasm detection can also be considered a future direction of research.

In [142] author has discussed that from a business perspective, detection of sarcasm would be very crucial in understanding that if product reviews, movie popularity, and social comments are placed in the wrong category, they may suffer.

Spam filtering and manufacturing product market analysis use detection of sarcasm for better classification of data as discussed in [181].

[161] discussed that sarcastic opinions lead to performance degradation. As a result, in order to increase sentiment analysis performance, there is a need to detect whether a post is sarcastic or not.

In [145] author considered the IMDb movie review dataset and converted text reviews into numeric matrices using a count vectorizer and TF-IDF, which was then given input to machine learning algorithms for classification. Naïve Bayes, Maximum Entropy, Support Vector Machine, and Stochastic Gradient Descent achieved 86.23 %, 88.48%, 88.94 % and 85.11% accuracy. It was stated that the use of hybrid techniques and different methods of feature selection could improve the accuracy further.

It is observed according to the study of sentiment analysis that the presence of sarcasm in the text can lead to performance degradation during the analysis. Thus, it is a major challenge to detect sarcasm. The following section discusses studies related to the detection of sarcasm.

2.2 Literature survey related to sarcasm detection

Analyzing sentiments from the textual data is not an easy task if the data consist of a few sarcastic comments in between. Thus, it has become a very important area to first identify and detect the sarcastic content in the text and thus improve the overall performance of the system. In this area, a detailed review has been conducted of various methods of detecting sarcasm in textual data.

The author developed a neural network (M_3N_2) based sarcasm detection framework that was multi-modal, multi-interactive, and multinomial [89]. The researcher selected Twitter, Image, Text in Image and Image Caption as inputs for the new framework because the brain's observation of sarcasm required several modalities. Especially, a multi-hop process was applied per modality interaction to pull out modal information repeatedly with GA to obtain multi-dimensional information. In addition, a two-hierarchical design promoting self-attention with attention pooling was used in this study to incorporate multi-modal semantic information from various levels. The results of the experiments showed that the developed architecture outperformed other existing models in detecting sarcasm and had good generalization abilities in multi-modal emotion scrutiny and emotion identification.

In [90] author presented a sarcasm detection method in Indonesian Twitter posts, mostly on a number of important issues, for example, politics, public figures and the service sector. The presented approach employed two feature extraction techniques known as interjection and punctuation. These techniques were subsequently applied in two diverse weighting and classification algorithmic approaches. The tested outcomes demonstrated that the unification of different feature extraction techniques such as tf-idf, k-Nearest Neighbor etc., achieved competitive performance in sarcasm detection.

In [101] author focused on enhancing the efficiency of the technique of detecting sarcasm. A bootstrap technique was utilized for acquiring a longer role pair list using a role pair extraction technique. Furthermore, the weighting techniques were implemented for every role pair. Thereafter, a comparative analysis was conducted on the presented approach against the traditional techniques. The analysis demonstrated that the presented approach was capable of enhancing the role pairs. The outcomes of experiments indicated that the presented approach with topic similarity became effective for mitigating the influence of diverse kinds of noise pairs.

The author introduced a procedure of extracting the features so that the sarcasm was detected with the help of bilingual texts and public comments, which were posted on economic posts on Facebook in [102]. The idiosyncratic attribute was deployed with the objective of capturing the peculiar and odd comments. The introduced procedure was computed using a non-linear SVM (Support Vector Machine) algorithm. This algorithm was useful for classifying the texts as sarcastic or normal with regard to recognized attributes. The outcomes revealed that the introduced procedure performed well by integrating syntactic, pragmatic and prosodic attributes and provided the F-measure score up to 0.852.

In [103], a model named IWAN (Incongruity-Aware Attention Network) in order to detect sarcasm was developed. To achieve this, a scoring system was implemented to consider the word-level incongruity among modalities. The larger weights were assigned to words with incongruent modalities using this system. The created model was quantified using the MUStARD dataset. The experimental outcomes exhibited that the developed model was efficient with regard to interpretability and performed well as compared to the existing technique. The developed model had the potential for capturing the incongruity among modalities and offered a scheme for probing the incongruity among modalities in sarcasm.

The author has designed a new framework in DL (deep learning) in [104] to detect the sarcasm for which common sense knowledge was incorporated. A pre-trained

COMET system was deployed to produce appropriate common-sense knowledge. Moreover, two techniques of selecting knowledge were compared for computing the impact of that knowledge on performance. Finally, the text, as well as knowledge, was modelled using a knowledge-text integration module. The results of experimentation on three datasets confirmed that the designed framework was applicable.

The author investigated a FL (fuzzy logic) system to detect sarcasm with the help of social information related to replies, historical tweets and likes, etc. and for categorizing the outcomes of the same text as having sarcasm and normal [107]. The value of recall was maximized by adding a degree of importance to social information that was utilized in the computation. The outcomes proved that the investigated system was applicable for enhancing the precision while classifying the text and boosting the accuracy. The future work would aim to augment the recall for which a dataset having an undecidable class would be discovered.

The authors of [123] present a comprehensive discussion of existing sarcasm detection algorithms. Feature Engineering-based methodologies, Deep Learning-based techniques, and Big Data-based approaches were loosely split into three categories.

The author of [125] presented a description of prior sarcasm detection work, a generic architecture for sarcasm detection, multiple types of sarcasm, various sarcasm identification techniques, and some sarcasm detection challenges.

The semantic gap was examined to calculate the polarity differences by producing the hybrid lexicon, according to [127].

In [129], the author suggested an algorithm to automate the task of detecting public shaming on Twitter from the perspective of victims, focusing on two aspects: events and shamers. Abusive, comparison, passing judgement, religious/ethnic, sarcasm/joke, and whataboutery are the six sorts of shameful tweets, and each tweet is categorized into one of these categories or as non-shaming. It relieves moderators of the burden of deciding on a threshold. Large datasets were not suited for the suggested work.

By focusing on two types of feature words, the author proposed a method for detecting sarcasm on Twitter [132]. The first are words transformed by features, and the second are terms that indicate a role. It produced useful findings, but more research into the value of feature words and the use of a weighting method for classifiers is needed.

The author described a supervised machine learning-based strategy for detecting sarcasm on Facebook that focused on both the content of postings (e.g., text, image) and the engagement of users with such posts in [133]. The findings revealed that supervised learning algorithms, particularly ensemble learning algorithms, are well suited to social media applications. The impact of spam messages and duplicated data on the accuracy of sarcasm detection could be studied in the future.

The author of [135] developed a new self-deprecating sarcasm detection method. For sarcasm detection, the suggested method integrated rule-based and machine learning methods. By selecting 11 features, three distinct classification models were trained: decision tree, nave Bayes, and bagging. Six of the features were self-deprecating, while the other five were exaggerated. The recommended strategy was evaluated using a Twitter dataset of 107536 tweets. The proposed strategy was compared to some state-of-the-art techniques in order to identify sarcasm in this study.

The author of [136] developed a system that organizes postings based on emotions and sentiments, as well as detecting sarcastic posts if they exist. The method creates a prototype that aids in inferring the post's emotions, which include anger, surprise, happiness, fear, grief, trust, anticipation, and disgust, each with three semantic levels.

In [137] author implemented twelve different classification algorithms on four sorts of datasets. These dataset types were called Set1, Set2, Set3 and Set4. In order to verify the accuracy of all algorithms in different conditions, the split ratio of the datasets was changed accordingly. In this work, a behavioural approach has been implemented for detecting sarcasm in the Twitter dataset. However, in set 4, which offered ideal accuracy in all three split ratio scenarios, there was a gradient improvement. These cases were 50:50, 25:75, and 10:90, and the accuracy rates were 85.14 percent, 85.71 percent, and 85.03 percent, respectively.

The author indicated that a lot of studies was done on the dataset, which included the # sarcasm tag in [140]. The structural and sentiment traits were mostly utilized by these studies. However, in certain circumstances, the # sarcastic tags were missing. To transcribe these types of cases, emotional and semantic similarity elements were required. In this paper, a unique approach for classifying sarcastic and non-sarcastic utterances was proposed. In this study, structural, emotional, and semantic similarity elements were used in conjunction with the MLP-BP technique.

In [141] author stated that it was possible to use the sarcasm detection approach for improving sentiment analysis. In contrast to text, only a few researchers were conducted on sarcasm through multimodalities in the sentiment analysis domain. According to the current studies, the use of multi modalities is on the rise, notably on social media platforms. For those studies that were solely focused on text analysis, it was useful to employ a large amount of online available multi-modal data.

In [143] author made a comparison between two conditions called ‘egocentric’ and ‘allocentric’. In the first condition, the preferential analysis was ironic just from the viewpoint of the candidate. In the second condition, the sarcastic analysis was prominent from the viewpoints of both the addressee and the candidate. In order to deal with the second question, the comparison of both conditions was performed after adding prominent prosodic cues. The obtained outcomes revealed that the perspective shifting was egocentrically secured.

The author examined the ability of children with ADHD to distinguish between sarcasm and sincerity in [144]. In this study, twenty-two children with a medical diagnosis of ADHD were compared to twenty-two children of the same age. The Awareness of Social Inference Test's Social Inference–Minimal Test was used to match the verbal IQ of maturing youngsters. The attained achievements were discussed, with challenges in the perception of challenging social signs and non-literal language as indicators of youngsters with a medical diagnosis of ADHD being discussed. Because of its potential for identifying social and expressive information, pragmatic language expertise was also highlighted in this study.

In [151,152], the author proposes a modified solution for the K-means clustering system by reducing the number of features using Principal Component Analysis and finds that the modified algorithm takes significantly less time than the K-mean algorithm when applied to a large number of data sets.

The author has suggested a method for sentiment analysis that categorizes text emotions as positive or negative [156]. Principal component analysis was used as a feature reduction tool, and a back propagation neural network classifier was used for the classification of data. In digital camera reviews, the datasets were downloaded from www. amazonreviews.com. Cross-validation was done ten times. Results were tested using the receiver operating characteristics (ROC) curve. When the performance of BPN and PCA+BPN is contrasted, it was discovered that the ROC curve for PCA + BPN was closer to the perfect point (0,1) compared to the BPN-based model.

One of the most difficult tasks of emotion analysis is detecting sarcasm. The author in [157] investigated emotion analysis sarcasm in tweets about a single subject utilizing features such as interjection and unigram features in this article. They used a Support Vector Machine with polynomial kernels to detect sarcastic sentences and compare them. It was discovered that using the interjection and unigram functionality on tweets with SVM increased sentiment analysis accuracy by 91%.

Since there is no static form for sarcasm in the data stream, as a result, utilizing Machine Intelligence to forecast sarcasm in Twitter (or every other semi-structured knowledge format) is challenging. As opposed to other heuristics that use pattern match or context-dependent, this is a more challenging yet thorough assignment. In the paper [158], the author demonstrated how various digital technologies could be utilized to combat societal issues and constructs which impede free speech. It is shown by the usage of the classification schemes for description and tweets classification. It was accomplished using a hyperbolic feature set. The project's potential analysis will involve resolving semantic uncertainty by utilizing a radical Recurrent Neural Network paradigm. Feed the network with the functionality and metadata created by the current model to accomplish this. Bidirectional LSTMs can be considered for context identification, and the VADER library can be used to perform a comprehensive emotion search.

In [159], a method for detecting sarcasm in bilingual texts that uses a variety of feature extraction categories and NLP is presented. The method extracts functionality from bilingual or interpreted corpora. Pragmatic, lexical, syntactic, idiosyncratic, and prosodic NLP characteristics were all listed. To test the feature groups, a non-linear SVM was used for classification purposes for the sarcasm detection (used on their own and in combination). The proposed model outperformed the others as compared to a baseline function.

The author provided a review of previous sarcasm detection work, an architecture for detecting sarcasm, various types of sarcasm, various sarcasm detection techniques, and certain sarcasm detection challenges in [160]. The complexity present in sarcasm renders things a more difficult task and raises the chances of finding jobs. The bulk of study into sarcasm detection is done in English. Future analysis should focus on detecting sarcasm in other languages. New datasets, features set, and consideration of different types of sarcasm, among other aspects, were proposed for future research.

The author outlined the most important work on ensemble approaches that has been done to date in [163]. With consideration of the numerous refinements to these three tactics presented by other studies, the major focus was on the three most well-known approaches: bagging, boosting, and stacking. Bagging approaches aim to increase variety among basic classifiers while boosting techniques aim to adapt it to more stable classifiers. One of the ways to generate a classifier ensemble was clustering algorithms. The design of novel classifier ensemble methods based on clustering analysis was highlighted as a promising new research problem in ensemble learning.

Table 2.1 explains the comparative analysis of various sarcasm detection used so far. Findings from the existing research include: the dataset used, the technique applied, major findings and future research scope, which leads to the researchers to solve existing problems.

Table 2.1: Findings from existing research techniques related to sarcasm detection

Citation	Data set	Technique	Findings	Limitations/Future Scope
[91]	Self-designed twitter dataset containing 2000 tweets	Recurrent Neural Network	The author studied various features of sarcasm such as semantic, syntactic, lexical attributes etc. Also considered contextual nature of sarcasm. Two neural network layers are used, each containing 256 LSTM cells. This architecture had the potential to extract features required for a machine learning approach automatically. The model gave 91 % accuracy.	Increasing the neural layers will increase computation, and decreasing it will affect the accuracy. Availability of contextual text would be a challenge for every domain.
[92]	Facebook Data	The hybrid approach of emoticon-based and content based	The author presented a new sentiment-based sarcasm detection framework. The proposed framework used various algorithms, libraries, and techniques at the emotion recognition stage and identified sarcastic posts using obtained results. This framework also considered hashtags and emoticons as a significant feature set of Facebook posts to detect sarcasm. The unified strategy adopted by this framework proved better than single approaches in terms of accuracy. Precision recorded was 88.57%.	A full-fledged database is required containing all the positive and negative situation phrases in English literature. The next stage of research will focus on detecting sarcasm in images and videos.
[93]	Arabic Tweets containing 350 sarcastic and non-sarcastic tweets	Naïve Bayes multinomial text classifier	The author put forward a Weka classification framework for detecting sarcastic Arabic tweets. The tweets in the Arabic language were gathered physically from different Saudi trending hashtags. The proposed classification system achieved 0.659 precision, 0.710 recall and 0.676 f-score. These results were significantly high, especially with respect to Arabic text.	The use of emojis and other features are not considered.

[94]	The Twitter dataset containing 107536 tweets	Combined rule-based and machine learning approach	The author applied a new self-deprecatory strategy for detecting sarcasm in tweets. Machine learning approaches were used for feature extraction and classification, while rule-based techniques were used to identify tweets about the candidate. In this study, 11 features were discovered that may be utilized to train three different classifier models: Naive Bayes, Decision Trees, and Bagging. Six of the features were self-deprecating, while the other five were exaggerated. The proposed approach was also compared with some standard methods of sarcasm detection in this work and performed with the better results.	Detection of other categories of sarcasm was not considered, such as brooding, raging etc.
[95]	Two Instagram datasets. One is Silver Standard Dataset containing 20K posts, and the other is a Gold Standard dataset containing 1600 posts	Recurrent Neural Network	The author presented a deep learning-based efficient framework by combining both textual and visual information for detecting multi-modal sarcasm. The strategy applied in this work was premised on RNNs (Recurrent Neural Networks) intended to use interactions between input modalities for forecasting. The new framework classified sarcasm by learning an illustration based on attention scores. The outcomes of the tests suggested that the use of all modalities with gating provided the most optimal interpretation	It has used only image and text modularity while other modularity such as conversation has not been considered.
[96]	1000 tweets containing 700 sarcastic and 300 non-sarcastic tweets	Support Vector Machine with RBF kernel, Decision Tree	The author proposed a novel sarcasm detection framework using various features to describe sarcasm in text, such as lexical, pragmatic, context inconsistency, subject and emotion. This enabled the framework to identify sarcastic tweets, even without context inconsistency in them. To model the new framework, SVMs (Support vector machines) and DTs (Decision Trees) models were employed in this work, which provided satisfactory results having an accuracy of 79.4 % and 74.1 % respectively.	Accuracy to detect sarcasm in the text can be further improved.

[99]	Internet Argument Corpus containing quotes and responses having 4692 instances of a sarcastic and non-sarcastic category	Long Short-Term Memory	A new multi-dimensional question answering (MQA) network for detecting sarcastic tweets was set forth to represent the introduced rich semantic information to interpret sarcasm ambiguity as well as generated conversation context information. Bi-LSTM (Bidirectional Long Short-Term Memory) and deep memory question answer network premised on the attention method for sarcasm exploration. The output of tests clearly proved the superiority of the presented model over other standard methodologies, and then more illustrations also confirmed the development and efficiency of the new architecture in sarcasm detection.	Features like hashtags and emojis were not considered.
[106]	Real-time Twitter dataset	NLP and corpus-based approach	The projected technique assisted in classifying the tweets as sarcastic and normal. Generally, the sarcastic statements were detected using the projected technique. The POS (Part-of-Speech) tags were employed to utilize the action words. This method has generated promising outcomes.	It works well for a small corpus but not for a large corpus
[108]	Internet Augment Corpus and	Convolutional Neural Network	To capture the properties of sarcasm expressions, the author constructed a	It does not resolve the problem of sparsity.

	Twitter dataset		multi-level memory network that used sentiment semantics. The sentiment semantics were captured using a first-level memory network, while the difference between the sentiment semantics and the scenario in each sentence was captured using a second-level memory network. In addition, when local information was absent, the memory network was improved using an enhanced CNN (convolutional neural network). Experiments proved that the network that was built was successful.	
[110]	Japanese product review	Rule-based approach	Initially, the technique focused on analyzing the sarcastic sentences in product reviews and classifying the sentences into eight classes with respect to evaluation parameters. Subsequently, the classification rules were produced for every class and the sentences containing sarcasm were extracted using them. The judgment processes, in which 3 phases were included, were utilized on the basis of rules for eight classes, boosting rules and rejection rules. The experiments were conducted to compare the suggested technique with the traditional technique. The results of the experiments showed that the suggested technique was adaptable.	Dataset was very small thus chances of improvement of accuracy will be there on a large dataset.
[27]	Twitter dataset with #sarcasm	Pattern-based approach	Four sets of attributes were put forward, which were able to cover diverse kinds of sarcasm. The tweets were classified as sarcastic or normal using these attributes. The accuracy acquired from the formulated method was calculated at 83.1%, and the precision was 91.1%. This method emphasized analyzing the significance of every set and computing its added value to classify the tweets.	Accuracy would have improved if large training data were used.
[112]	Sarcastic book	Combinatio n n of	Initially, machine learning technique was utilized for classifying the sarcastic.	Computation cost has been increased.

	snippets, sarcastic tweets, sarcastic Reddit comments	machine learning and deep learning model	sentences so that it was determined whether the sarcastic sentence had a target or not. After that, the target was extracted through a deep learning model generated from the ABSA (Aspect-Based Sentiment Analysis). The outcomes indicated the devised supremacy of the devised approach over the existing techniques, and this technique provided an enhancement of 18% on the Reddit data set.	
[126]	Tweets containing product reviews from amazon	Guess-based - Naive Bayes	The author developed a fresh Naive Bayes approach for detecting sarcasm in the Amazon Alexa dataset. SentiWordNet and TextBlob were used to extract essential characteristics from the dataset, which were then utilized for training the model. The test dataset is put to the test using a Gauss-based naive Bayes approach as well as three baseline methods: Naive Bayes, decision tree, and support vector machine. The proposed strategy achieved 70.96, which outperforms the other classifiers.	A major focus was on capital words only.
[111]	Twitter dataset	Pattern-based approach	Random forest, SVM, KNN classifiers were compared and observed that Random Forest classifier has achieved the highest accuracy of 81% as compared to others.	All patterns were not covered in the extracted patterns. In a future neural network, pattern-based approach and genetic algorithm can be combined for more accuracy.
[138]	1500 tweets collected from Twitter	Neural network	The author explored Neural Network classifiers for detecting sarcasm within tweets. This research proposes two distinct context-added neural models for detecting sarcasm based on a convolutional neural network. The proposed model was able to decode sarcastic clues from content-based information. In terms of detecting sarcasm, the model performed admirably by achieving accuracy of 62.05 %.	The availability of history-based tweets would not be suitable for all domains.

[30]	Twitter dataset	Parsing-based lexicon generation algorithm (PBLGA), IWS (Interjection_word_start)	The author recommended two novel techniques of sarcasm detection within tweets. The first technique was called the parsing-based lexicon generation algorithm (PBLGA). The next approach was based on the incidence of the interjection word. In order to identify sarcasm, these two methods were combined and analyzed to the best existing methods. The first strategy yielded precision, recall, and f – score of 0.89, 0.81, and 0.84, respectively. However, in a Twitter text with a sarcastic #(hashtag), the following strategy obtained precision, recall, and f - score of 0.85, 0.96, and 0.90, respectively.	Future work can be done for sarcasm detection on audio clips and images.
[142]	Twitter dataset	Naïve Bayes and fuzzy clustering	The author considered different text-independent feature sets, including function words and part of speech n-grams. Several feature sets were tested using Nave Bayes and fuzzy clustering techniques. The results of the tests showed that it was beneficial to include several aspects that captured the microblog writers' writing style to detect sarcasm. In this study, the recommended strategy had a 65 percent accuracy rate.	Features including the author's writing style were included, which might not be available in every domain.

2.3 Literature survey related to ensemble learning

After studying the various sarcastic detection techniques in the previous section, a few more literature has been studied related to the ensemble learning method of sarcasm detection.

In [151,152], the author proposes a modified solution for the K-means clustering system by reducing the number of features using Principal Component Analysis and finds that the modified algorithm takes significantly less time than the K-mean algorithm when applied to a large number of data sets.

The author suggested a method for sentiment analysis that categorizes text emotions as positive or negative [156]. Principal component analysis was used as a feature reduction tool, and a back propagation neural network classifier was used for the classification of data. In digital camera reviews, the datasets were downloaded from www. amazonreviews.com. Cross-validation was done ten times. Results were tested using the receiver operating characteristics (ROC) curve. When the performance of BPN and PCA+BPN is contrasted, it was discovered that the ROC curve for PCA + BPN was closer to the perfect point (0,1) compared to the BPN-based model.

One of the most difficult tasks of emotion analysis is detecting sarcasm. The author in [157] investigated emotion analysis sarcasm in tweets about a single subject utilizing features such as interjection and unigram features in this article. They used a Support Vector Machine with polynomial kernels to detect sarcastic sentences and compare them. It was discovered that using the interjection and unigram functionality on tweets with SVM increased sentiment analysis accuracy by 91%.

Since there is no static form for sarcasm in the data stream, as a result, utilizing Machine Intelligence to forecast sarcasm in Twitter (or every other semi-structured knowledge format) is challenging. This is a more challenging yet thorough assignment than other heuristics that use pattern match or context dependent. In the paper [158], the author demonstrated how various digital technologies could be utilized to combat societal issues and constructs which impede free speech. It is shown by the usage of the classification schemes for description and tweets classification. It was accomplished using a hyperbolic feature set. The project's potential analysis will involve resolving semantic uncertainty by utilizing a radical Recurrent Neural Network paradigm. Feed the network with the functionality and metadata created by the current model to accomplish this. Bidirectional LSTMs can be considered for context identification, and the VADER library can be used to perform a comprehensive emotion search.

In [159], a method for detecting sarcasm in bilingual texts that uses a variety of feature extraction categories and NLP is presented. The method extracts functionality from bilingual or interpreted corpora. Pragmatic, lexical, syntactic, idiosyncratic, and prosodic NLP characteristics were all listed. To test the feature groups, a nonlinear SVM was used for classification purposes for sarcasm detection (used on their own and in combination). The proposed model outperformed the others as compared to a baseline function.

In [160] author provided a review of previous sarcasm detection work, an architecture for detecting sarcasm, various types of sarcasm, various sarcasm detection techniques, and certain sarcasm detection challenges. The complexity present in sarcasm renders things a more difficult task and raises the chances of finding jobs. The bulk of study into sarcasm detection is done in English. Future analysis should focus on detecting sarcasm in other languages. New datasets, features set, and consideration of different types of sarcasm, among other aspects, were proposed for future research.

The author outlined the most important work on ensemble approaches that has been done to date in [163]. The main focus was on the three most well-known approaches: bagging, boosting, and stacking, with a review of the numerous enhancements to these three strategies offered by other studies. The bagging approach focuses on increasing diversity among base classifiers, whereas boosting techniques focus on adapting it to more stable classifiers. Clustering algorithms were used as one of the methods for constructing a classifier ensemble. As a promising new research issue in ensemble learning, the creation of novel classifier ensemble methods based on clustering analysis was emphasized.

In [170], The author presented an ensemble classification technique, which has been found to increase Twitter sentiment classification performance. The suggested ensemble classification system was compared to a number of classic sentiment analysis approaches as well as the most widely used majority voting ensemble classification system. Different base learners, such as the Nave Bayes, Random Forest classifier, SVMs, and Logistic Regression, make up the ensemble classification system. According to the findings, the proposed ensemble classifier outperforms stand-alone classifiers and the widely used majority vote ensemble classifier.

Table 2.2 explains the comparative analysis of various ensemble learning techniques for the sentiment analysis used so far. Findings from the existing research include: dataset used, the technique applied, major findings and future research scope, which leads the researchers to solve existing problems.

Table 2.2: Finding from existing research related to ensemble learning

Citation	Dataset	Technique	Major Findings	Limitations/ Future scope
[147]	Twitter data	Multinomial Naïve Bayes, Logistic Regression, Random Forest	The author proposed a classifier ensemble using lexicons, emoticons, a bag of terms, and attribute hashing. Random Forest, Multinomial Naive Bayes, logistic regression and Support Vector Machine were chosen as baseline classifiers with recorded precision is 81.08 percent.	Works well for small datasets only.
[31]	Twitter dataset	Ensemble method based on Bayesian Model Averaging	By using pragmatic particles and POS marks in the function sets, the author used an ensemble technique to spot sarcasm and cynicism in the document. Ensemble models include Support Vector Machines, Decision Trees, Naive Bayes and Bayesian Networks. Pragmatic particles were found to be better at identifying sarcasm, whereas POS tags were better at identifying irony.	Dataset is highly unbalanced
[150]	500 Product review dataset	Principal Component Analysis	The author applied the feature reduction approach – Principal Component Analysis to a Twitter dataset of product feedback and tested it with Support Vector Machine and Nave Bayes. The use of PCA resulted in an improvement in precision.	Dataset considered is too small.
[153]	Twitter dataset of movie review	Principal Component Analysis with Random Forest Classifier	The author proposed using a feature reduction method known as PCA to improve the classifier's performance on tweets (Principal Component Analysis). It was discovered that the proposed random forest tree-based feature reduction increased the classifier's accuracy, precision and recall. The highest precision was 81.45%.	Sarcasm is not detected.

[146]	The dataset containing 994 samples having sarcasm tag, sentiment tag and eye movement data of seven readers.	GRU-based neural network	A multitask learning-based framework using a deep neural network is proposed by the author. An improvement of 3-4 % was observed when the proposed approach was compared with the state-of-the-art.	Multimodal information can be added in the architecture for further improvement.
[154]	Real-time tweets	Various algorithms	The author introduced Hadoop-based architecture for recording tweets in real-time and manipulating them using a series of algorithms to efficiently recognize satirical sentiments. TCUF, LDC, IWS, PBLGA, TCTDF and PSWAP is six algorithms suggested in this paper for detecting sarcasm in the tweets received from Twitter. Then three algorithms were tested using the Hadoop system and without it. Processing time was found to be decreased by up to 66 percent using the Hadoop system.	The accuracy of the proposed approach is dependent on time-dependent facts.
[155]	Twitter datasets on a variety of topics: Stanford – Sentiment 140 corpus, Health Care Changes, First GOP Discussion Twitter Sentiment Dataset, and Twitter Sentiment Analysis Dataset.	Nave Baye s classifiers, Random Forest classifiers, Support Vector Machines, and Logistic Regression are the most popul ar classification techniques used.	The author suggested an ensemble classification scheme to improve sentiment analysis accuracy for tweets. Nave Bayes classifiers, Random Forest classifiers, Support Vector Machines, and Logistic Regression are most popular classification techniques used. The proposed ensemble classifier outperforms stand-alone classifiers and majority voting ensemble classifiers, according to the results.	Analysis of neutral tweets can be done in future.
[168]	Reddit and Twitter comments	Adaboost classifier	The author presented an ensemble technique for identifying sarcasm in Reddit and Twitter comments. The model uses an Adaboost classifier using the decision tree technique as the basis estimator to learn the sarcasm probability. The ensemble provides F1 scores of 67 percent and 74 percent on	Availability of conversation features for all domains will be challenging.

			the Reddit and Twitter test data, respectively.	
[169]	Github Repository	Hybrid model and weighted average model for Support Vector Machine, Random Forest and Naïve Bayes algorithm	The author proposed two approaches cascading algorithm (Hybrid model) and a hybrid weighted average approach for the detection of sarcasm. The Random Forest algorithm was merged with the Nave Bayesian (NB) and Support Vector Machine (SVM) algorithms. With a cascade accuracy of 90.37 percent and a weighted average accuracy of 67.29 percent, the cascade method was found to be more accurate than the weighted average approach. There was also a comparison of approaches based on algorithm placement, with the random approach placed first in the code, outperforming others with a 90.37 percent accuracy.	Order of placing algorithms is a major issue to be considered.
[118]	Twitter dataset	Genetic Algorithm Based Feature Reduction, A hybrid technique combining lexicon-based and machine learning approaches.	Based on a Genetic Algorithm, the author developed a new feature reduction strategy (GA). This hybrid technique was able to reduce the size of the feature set by up to 42% without sacrificing accuracy. The suggested feature reduction method was compared to more widely used feature reduction methods such as Principal Component Analysis (PCA) and Latent Semantic Analysis (LSA) (LSA). In comparison to PCA and LSA-based techniques, the proposed approach enhanced accuracy by 15.4 percent and 40.2 percent, respectively.	Problem of scalability
[128]	Amazon product review datasetand Cornell movie review dataset	Ensemble classifier with base learners: SVM, NB, and GLM	For document sentiment classification, the author suggested an algorithm that detects and selects appropriate features with the goal of decreasing high-dimensional feature space while taking into account semantics, sentiment clue, and word order. Idioms, metaphors, irony, and sarcasm are excluded from the suggested algorithm.	The proposed approach does not fit the sarcastic text.

[131]	Arabic tweets	Ensemble model of surface and deep features	The author attempted to extract sentiment-specific word embeddings from Arabic tweets. The sentiments of Arabic Tweets were eventually classified using these word embeddings. In this study, a unique feature ensemble model comprising a surface and deep features were suggested. Physically, the surface features were obtained. Generic word embeddings and sentiment-specific word embeddings were the deep features. The efficacy of the surface and deep features ensemble was put to the test in a number of ways. Pooling functions, embeddings dimension, and cross-dataset models were also investigated in these studies. The accuracy rate was 80.47 percent.	The model may not work well for the large dataset.
[148]	Facebook data	Ensemble approach	The author proposed a supervised machine learning method for detecting sarcasm in Facebook messages. To determine whether or not a post is satirical, a combination of numeric, text, and pictures is used. Support vector machine along with linear kernel, two of the ensemble algorithms – Adaboost along decision tree classifier and Random Forest along with multi-layer perceptron and Gaussian Nave Bayes – and five machine learning algorithms were included. Both ensemble learning approaches have a high level of precision (>90%).	The collection of multimodal data sets will be a challenge in other domains.
[165]	Ten public sentiment analysis datasets	Ensemble approach	For sentiment analysis, the author assessed ensemble approaches (Bagging, Boosting, and Random Subspace). Random Subspace has superior comparison results when using DT, KNN, and SVM as the base learner, but it has the lowest comparative results when using NB as the base learner. Boosting has poor accuracy unless when using DT as the base learner. The lack of interpretability of the findings gained by ensembles is a key	Only the bag of words feature is used. Ensemble learning needs a lot of computation time.

			disadvantage of ensemble learning methods: the knowledge obtained by ensembles is difficult for humans to understand. The findings show that ensemble learning methods are a feasible option for sentiment categorization.	
[166]	Corpus of textual data collected from various online sources	Ensemble of knowledge-based and statistical machine learning classification methods	The author proposed a sentiment analysis system that uses an ensemble of classifiers to automatically recognize emotions in text, combining knowledge-based and statistical machine learning classification methods. It is built around three core classifiers: a naive Bayes learner, a maximum entropy learner, and a knowledge-based tool, all of which are integrated using a majority voting method. The findings reveal that the ensemble schema does a fantastic job of recognizing the subjective polarity of the sentences.	Sarcasm is not detected.
[167]	Twitter dataset	Ensemble learning using Random Forest, Naive Bayes, Support Vector Machine, K-Nearest Neighbor, Gradient Boosting, AdaBoost, Logistic Regression, and Decision Tree.	Ensemble learning is used to locate and choose the optimal set of features for detecting sarcasm accurately. The results show that the ensemble-based feature selection method achieves an accuracy of 92.7%.	Availability of past history tweets of the user would be a challenge.
[113]	News Headlines	Convolutional Neural network with GloVe embeddings	Accuracy of 86.13% was achieved.	False labelled results can raise an issue
[97]	Datasets were collected from SEMEVAL Task 11	Ensemble of features	The author divided the work of sarcasm detection from tweets into two stages. The first stage was concerned with extracting features associated with	Ensemble of classifiers was not performed.

	containing 16000 tweets		emotion and punctuation. Afterwards, the chi-square test is conducted to choose the most fitting features. The next stage involved extracting the most useful 200 tf-idf features, which were later integrated with sentiment and punctuation-related features to figure out sarcastic Twitter feeds. In the first stage, the Support Vector Machine technique was utilized to achieve high accuracy of 74.59 percent, whereas the next stage witnessed the highest accuracy of 83.53%, achieved using the voting classifier.	
[98]	IMDb Moview dataset	Four different classifiers were used- SVM, KNN, Maximum entropy and Naïve Bayes	To extract features from machine learning and lexicon-based feature extraction approaches, a hybrid feature selection method was applied. Maximum Entropy has performed better than others with an accuracy of 83.93%.	Sarcasm is not detected
[149]	Twitter and News Headlines dataset	A hybrid approach using CNN and LSTM	The author has used CNN for feature extraction, and LSTM is trained and tested on those features. The proposed sarcasm detection method beat a variety of machine learning methods, including random forest, support vector classifier, additional tree classifier, and decision tree, with a 91.60 percent accuracy.	Voting classifiers are not used.
[124]	Reddit and Twitter dataset	Ensemble of Adaboost classifier with decision tree algorithm	The author constructed an ensemble classifier using four component models- LSTM, CNN-LSTM, SVM and MLP. The features used were conversational context. The ensemble approach yields the best F1 score of 66.7% and 74% on the Reddit and Twitter datasets.	The imbalance of Reddit and Twitter datasets is a challenge which leads to lower performance.

2.4 Summary

This chapter clarifies the existing state of the art of sentiment analysis. To do so, the first general process and methods of sentiment analysis are discussed in detail. It was seen that sarcasm is one of the major issues during the process of sentiment analysis. Thus, literature related to sarcasm is studied in detail. Various techniques used by the researchers to detect sarcasm in the text have been described in this chapter. It was observed that various machine learning and lexicon-based approaches used still have a scope for improvement by proposing the ensemble techniques. Finally, the ensemble learning method of detecting sarcasm in the text has been studied and described. It was found that most of the studies have been conducted on small datasets, and some of them have included a few features which are not available for every domain and thus would be a challenge. The use of multimodal data and voting classifiers was also missing. A further improvement in the accuracy is envisioned by the proposed ensemble approach.

The next chapter proposes an ensemble approach for sarcasm detection in order to address the issues identified in problem formulation and achieve the objectives of this research work.

CHAPTER 3

DESIGNING AND IMPLEMENTING ENSEMBLE CLASSIFIER FOR SARCASM DETECTION

This section represents the sarcasm detection technique based on machine learning algorithms. The four ensemble classifiers were created to identify sarcasm. The ensemble classifiers are a combination of multiple classifiers. SVM, KNN, and decision tree are combined in the first ensemble classifier. For sarcasm detection, the second ensemble classifier, SVM combines logistic regression and decision tree classifiers. MLP, logistic regression and decision tree are integrated into the third ensemble model, while MLP, logistic regression, and SVM are coupled in the last.

3.1 Introduction

In this chapter, the ensemble approach for detecting sarcasm is designed and discussed. Ensemble learning is a technique for constructing many base classifiers from which a new classifier is produced that outperforms any of its constituent classifiers.

3.2 Need for Ensemble Classifier

According to the literature studied, the presence of sarcasm in the text leads to degradation of the performance of sentiment analysis. Thus, there is a need to detect sarcasm. Various machine learning algorithms and lexicon-based methods have been applied to detect sarcasm. Still, it was observed as per the study that the use of ensemble classifiers might have improved the accuracy of detecting the sarcasm in the text and further improved the overall sentiment analysis process. Thus, the main reasons behind applying ensemble learning are summarized below:

- To obtain the best possible classification accuracy.
- Reduces the chances of misclassification of the data.

3.3 Proposed algorithm

Algorithm

K: Number of Clusters

C: Number of Centroids

Step 1. Input Dataset for the sentiment analysis.

Step 2. Pre-process dataset to remove missing and redundant values.

Step 3. Feature Extraction.

3.1 Repeat until tree creation

3.1.1 From the training data, create a bootstrap sample Z of size N.

3.1.1.1 By recursively repeating, grow a random-forest tree Tb from the bootstrapped data.

3.1.1.2 The following procedures should be repeated for each tree's terminal node until the minimum node size of n min is reached.

3.1.1.3 Select m variables at random from the p variables.

3.1.1.4 Pick the best variable/split-point among the m.

3.1.1.5 Split the node into two daughter nodes

Step 4. Apply K-mean

4.1. Choose the number of clusters(K) and obtain the data points

4.2. Place the centroids c_1, c_2, ... c_k randomly

4.3. Repeat steps 4.1 and 4.2 until convergence or until the end of a fixed number of iterations

4.4. Repeat until clustered data

4.4.1 Find the nearest centroid (c_1, c_2 .... c_k)

4.4.2 Assign the point to that cluster

4.5. for each cluster j = 1..k

4.5.1 New centroid = mean of all points assigned to that cluster

4.6. End

Step 5. Apply PCA Algorithm

5.1. Maximizes Variance of Projected Data

5.2. Minimizes the Mean Squared Distance Between Data Point and Projections

5.3.Given Data points in a D-Dimensional Space, Project them into a Lower Dimensional Space While Preserving as Much Information as Possible.

5.4. Find the best planar approximation of 3D

data. Step 6. Classification Step

6.1. Apply Support Vector Machine, K Nearest Neighbor and Decision tree algorithms to train the first ensemble.

6.1. Apply Support Vector Machine, Logistic Regression and Decision tree algorithms to train the second ensemble.

6.1. Apply Multiple Layer Perceptron, Logistic Regression and decision tree algorithms to train the third ensemble.

6.1. Apply Multiple Layer Perceptron, Logistic Regression and Support Vector Machine algorithms for the training of the fourth ensemble.

6.2 Apply the voting process for the prediction.

3.4 Methodology

This research work presents an ensemble classifier for detecting sarcasm. Pre-processing, feature reduction, grouping, and classification are all steps in the sarcasm detection process. The PCA technique is used to minimize the characteristics. K-mean clustering is used to group together similar and dissimilar types of data. Various voting classifier models are created to classify the data. The SKD, SLD, MLD, and SLM voting ensemble classification models are created. Support Vector Machine, K-Nearest Neighbor, and Decision Tree classifiers are merged in the first model using a voting method. Support Vector Machine, Logistic Regression, and Decision Tree are combined in the second model. Multilayer Perceptron, Logistic Regression, and Decision Tress are combined in the third model. The final model combines Multilayer Perceptron, Logistic Regression, and Support Vector Machine into a single model.

Various steps are explained below:

1) **Dataset Collection**

Data is collected from multiple domains to comprehend the efficacy of the model. Two datasets, including 'Tweet' datasets [162] and 'News headline' datasets [164], are combined to form a single dataset as considered in [149]. There were no challenges faced during dataset collection as we have considered the same dataset which was earlier used in [149].

The Tweet dataset contains 81,408 randomly generated tweets that have been categorized as figurative (including sarcasm and irony), regular, sarcasm, and irony [162]. The tweets labelled sarcasm, and regular are extracted from the tweets dataset, resulting in 39,276 tweets. Sarcasm is identified in 20,681 of the tweets, whereas regular is labelled in 18,595. The news headlines dataset includes 26,709 headlines obtained from two separate news websites: 'The Onion' and 'HuffPost' [164]. There are 14,985 sarcastic headlines and 13,634 non-sarcastic headlines among the news headlines.

These two datasets (Tweets and News Headlines) are combined to create a multi-domain dataset for model training and testing.

When the datasets are combined into a single dataset, it yields 65,985 records for model training and testing, as described in Table 3.1, which is sufficient for performance evaluation. The number of sarcastic and non-sarcastic records in the dataset is shown in Figure 3.1. Sample of the dataset is also presented in Fig 3.2 in which class 1 represents sarcastic and class 0 represents non sarcastic sentences.

Table 3.1: Details of dataset

Dataset	No of rows
Tweets	39,276
News Headlines	26,709
Total	65,985

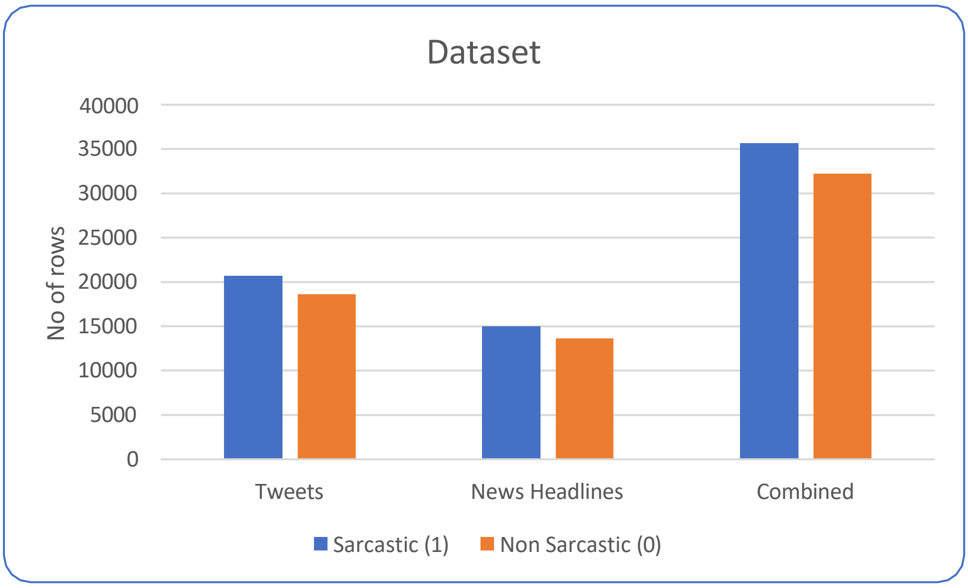


Fig 3.1: No of sarcastic and non-sarcastic records in the dataset

data	class
New #quote : It's both a blessing and a curse to feel so much. #mondayinspiration #late #quote #emotion #feel #emp http://t.co/5zLIMGgH70	0
#AugustPhotoChallenge #Day15 - #Late #Nails I need a #manicure really bad. @ Raleigh-Durham	0
This morning wasn't good, any time your alarm becomes part of your dream can only mean one thing! #Late	0
I love it when at family dinners people spend the whole dinner trying to get me to sign up for their fitness class. #sarcasm	1
Finally! I found out how to make the Kylie Jenner lip challenge work! #late http://t.co/S7A2mg0pJv	0
When your devices run out of batteries, let nature charge thee. It can see you clearly. It even smiles. #late http://t.co/1gQVPn4uN4	0
Thanks for letting me know about our money issues before I plan a night out with friends for my birthday, parents	1
The theory of everything is killer. #late	0
Using Skype for some of my coaching sessions has turned out to be very productive. #productivity #late #technology	0
Goodnight people - video day tomorrow :) #French #late #holidays	0
At the populating slides stage #theconf #late #iot	0
When NO One Asks you while you Do The Bad Thing. . . Then therez no wrong in making it . ... #Realized #late .. #Indian_Mentality	0
Good morning! Thank you God for another day. #LATE	0
We always #figure out #life no #matter what, but we figure it #late or too #soon, that why we end up #alone	0
I saw this at SFCC. It used to be 90% but so many people showed up that they had to change it. #inflation #late #ev	0
Looking for a #Part-Time #Late Night Team Member - Harbour Landing #jobs http://t.co/Bm4gOS5uBw	0
Sick on my last day...great just what I needed	1
Finally off work. #late #night #tyson #tysontheactor #dayjob #follow #chasingdreams https://t.co/a1K8Oziuzg	0
Best not to fly DUB LGW if you are in a hurry. Flights now have to bus to the main terminal because of UK immigration #late	0
Okay wait wait, who is in one of the good rows at AWP and wants to split a table? #late #AWP #dontwanttobeintheboondocks	0
The whole doctors office is losing their shit bc they got new computers.. Isn't technology supposed to make things easier??	0
I havent tweeted in ages #late	0
Day 229: Coming back content is exactly why we need to restart ourselves from our everyday routines. #late #365Days	0
Facebook have built a drone, and it doesn't look creepy at all http://t.co/SOWvSa6vtU	1
So today's challenge getting the 3 year old to nursery!! #late #lazylizzie	0
Girls, don't do this when you're late! http://t.co/pY52pIhPc2 #nailpolish #late #hurry	0
When you decide to grab a coffee before work and you get the new guy... #late	0
Opps didn't know I worked today all well #lookinghettoasfuck #late	0
Alright flight crew where the hell are you? #JetBlue #delays #iwannagohome #JFK #late	0
I'm going to miss my best friend going off to college. I love you, hope it's a great year #late @daniijacqueline http://t.co/o4REhfbCXR	0
It had been a long time since last time i had to run to catch the bus #late	0
Best part of having everyone at work love me is that they all cover for me #late	0
Unfriended was too fucking good. #late	0
After so long, i finally gave into Twitter... #late	0
Can I get ready for work in 10? #late	0
When your mind says get ready for work but your body says stay in bed. #late night	0
Can I get ready for work in 10? #late	0
It's a beautiful Monday when you get a call from someone's angry wife. #journalism #sarcasm	1
When your mind says get ready for work but your body says stay in bed. #late night	0
Late dinner + Alone #alone #dinner #late	0
2 decades on earth! Thank you, Lord! #late http://t.co/lYGMvt64E4	0
Guess my sleep schedule has caught up with me. #late	0
Time to learn how to tweet #late	0
People can't drive in the rain. #late •	0
You know it is going to be a great week when you work 13 hours on a Monday #Sarcasm	1
Make up primer is the best! I just discovered it. It's definitely a make up bag staple now. #late #beauty #makeup	0
Shouldn't have snoozed my alarm for the fourth time... #Late	0
Sometimes we have to do things we are not ready for ... Before it's too #late	0
What is going on my phone wants to work half of the time can't get my messages on fb or nothing ugggh.....#late	0
: Slowly and silently walk in to the office. #late	0
Thank you for taking the time to answer my texts... Means a lot #sarcasm	1
Late night in the garden. A bit glamorous. #garden #roses #late #night https://t.co/dIezFYn9or	0

Hey awesome people check out my #MovieReview of #Dope!!! #DopeMovie #Drugs https://t.co/XAaIBDzqlE	0
Twitter is actually free...hmmm okay, thanks. #late	0
What You're Really Telling #People When You're 'Just' a Minute or Two #Late http://t.co/j0q48ZeAtM #Entrepreneur	0
I love when random people who have never interacted with me start following me. #sarcasm #wtf	1
Teach your #kids #photography and they will never have #money for #drugs @ Photo Fusion Studio https://t.co/YzIqu595ri	0
Now the express train will be a local train. That way it will be even more #late #unbelievable @MBTA #subpar	0
Better late then never is the best policy when getting to my yoga. It's noon and I'm jumping in. #yoga #meditation #spirit #prayer #late	0
Signs that a Loved One is Abusing #Drugs - http://t.co/pvDoklDzQZ #Addiction #Treatment #Rehab #Recovery	0
Tea, such a good way to start the day @ 11:56!!!!! #LATE #SUMMER #MORNINGS !!!!!!!	0
Watching the Athletics has become like WWE. It's fiction. I'm waiting for bolt to put Gatlin through the table #TLC #Drugs #Doping	0
Thanks to the extremely hot weather and my allergies, I feel like dying. Absolutely lovely day isn't it? #Sarcasm	1
Get up early, go to bed late - what an unhealthy lifestyle #lifestyle #healthy #not #late	0
Saturday Night !! Been outside ? = very nice Enjoy yourself + Be Safe We are #Downtown #SanAntonio Popped up and	0
Ain't it Funny how #drugs are celebrated for improving health, but frowned upon when it comes to improving performance #WorldChampionships	0
When you got your life together but the cab driver don't #late	0
we will never #understand...till its too #late	0
Thank you sun, it's always been my dream to faint because of heat exhaustion Ã°Å¸Ëœâ€™Ã°Å¸Å'Â¾ #sarcasm #takeiteas	1
A little late in the day funny for our WOW, "Humor"...#WOW #WordoftheWeek #Humor http://t.co/Htq6M5MYf2	0
So happy I finally got to watch the #GOPDebate¸ #beinformed	0
Guns in banks to protect our money but no guns in schools to protect our children? #Late term abortion? maybe? https://t.co/9zsGYk32Da	0
There are two kinds of politicians those trying to get an investigation started and those trying to get one stopped. - Joe Moore #humor	0
Nobody's perfect. Except fat girls. #sarcasm	1
10 types of men I would never date http://t.co/UuVGIjiN9n #humor #offbeat #funny #fun http://t.co/VXIX0khK8V http://t.co/0TwN11C5Iv	0
Be Careful Who You Make Angry #humor 9 http://t.co/i0ABIHfMyE	0
when you're supposed to be home before dark and your automatic lights come on #late	0
My dad got jumped by a black man and I just stood there taking a video http://t.co/Urx9AdEdVx #meme #humor	0
I love when a change in plans by someone else completely alters the plan you had #sarcasm	1
India - where bus drivers honk even inside an airport. ;) #India #humor #honking #driver #fun	0
Will the iPhone influence human evolution? - http://t.co/rx8DMgnd5J #news #tech #football http://t.co/tyHwRs3TxD	0
Living an hour away from school instead of a couple minutes away really sucks. #late	0
What's the point of having a train schedule when it's always late. #amtrak #train #late #why	0
If you lie to people to get their money, that's fraud. If you lie to tehm to get their votes, that's politics. - Joe Moore #quotes #humor	0
You never realize how short a month is until you pay alimony. - Joe Moore #quotes #humor #quote	0
Today's cartoon, British Version. #cartoon #humor #Zombies http://t.co/IeMb0ECUcm	0
Children are natural mimics: they act like their parents in spite of every attempt to teach them good manners. #Parenting #Humor #TeamJesus	0
BREAKING NEWS...Donald Trump doesn't think Donald Trump's book is the best of all time. #Politics #GOP #USA http://t.co/QWKuc3Lg1r	0
Take the Dog shopping! It'll be #Fun they said http://t.co/iahcSsXgMu #humor	0
Donald Trump and Hillary Clinton revealed to be distant cousins http://t.co/Qc4wuvxAav #politics #tech #military	0
Just waking up is not cool #late	0
"A boss with no #humor is like a job that's no fun." #figuremeout #TeamFollowBack #FOLLOWBACK #FOLLOWME #iFollowBack	0
I'm okay now, I've just taken my happy pill #humor #drugs #pills #bipolardisorder #bipolar #depression #read #IARTG http://t.co/O50HMvxS9d	0
The power of idiots. #humor http://t.co/39w9tfc8dw	0
When You Know Your Screwed' - http://t.co/x22XGm8MG9 - #funny #humor #humour http://t.co/H3ttg3ACJk	0
Why is my weather app promoting smoking all day? #humor http://t.co/l80Xhkt8ml	0
I love summer allergies #sarcasm #sickofsneezing	1
We'll never tell :) #wisdomoftheday #humor http://t.co/quyOSgsmbE	0
No matter how bad a child is, he is still good for a tax deduction. - Joe Moore #quotes #humor #quote	0

Fig 3.2 Sample dataset

2) Pre- Processing of the dataset

After the collection of data, cleaning is done. It helps to remove the noise from the data, such as mentioned below:

- Removal of hashtags: The hashtags are usually used for monitoring and are removed before the tests on the tweets are conducted.
- Removal of unnecessary items: Hyperlinks, punctuation, and author names may be present in a tweet, but they are eliminated because they are unrelated to the experimental environment.
- Removal of unwanted tweets: Redundancy may exist in the corpus due to similar tweets being retweeted. As a result, they must be deleted in order to avoid skewed results. Tweets with fewer than three words are also eliminated.
- Tokenization: Tokens are symbols, phrases, words, or other valuable pieces that can be broken down from a series of words.
- Stop word removal: Stop words are terms that are removed from the text before or after it is pre-processed. These are controlled by humans and are frequently ineffective for text classification, such as a, an, and the.
- Stemming: Stemming is the process of returning developed words to their basic form, as in Attending—stemmed to—attend.
- Lemmatizing: Stemming frequently renders a term meaningless since stemmers only remove the affixes and do not add the missing characters that constitute the root of a complete meaningful word. Lemmatizer is in charge of this. It not only removes the affixes, but also adds the missing characters to the root, yielding a semantically complete word, such as stemming: decided–decid; Lemmatizing: decided–decide.

Figure 3.3 represents the pre-processed dataset.

cleaned_text	class
new quote it 's both a blessing and a curse to feel so much. mondayinspiration late quote emotion feel emp	0
augustphotochallenge day15 late nails i need a manicure really bad.	0
this morning was n't good any time your alarm becomes part of your dream can only mean one thing late	0
i love it when at family dinners people spend the whole dinner trying to get me to sign up for their fitness class.	1
finally i found out how to make the kylie jenner lip challenge work late	0
when your devices run out of batteries let nature charge thee. it can see you clearly. it even smiles. late	0
thanks for letting me know about our money issues before i plan a night out with friends for my birthday parents	1
the theory of everything is killer. late	0
using skype for some of my coaching sessions has turned out to be very productive. productivity late technology	0
goodnight people video day tomorrow french late holidays	0
at the populating slides stage theconf late iot	0
when no one asks you while you do the bad thing. then therez no wrong in making it realized late indian_mentality	0
good morning thank you god for another day. late	0
we always figure out life no matter what but we figure it late or too soon that why we end up alone	0
i saw this at sfcc. it used to be but so many people showed up that they had to change it. inflation late ev	0
looking for a part-time late night team member harbour landing jobs	0
sick on my last day great just what i needed	1
finally off work. late night tyson tysontheactor dayjob follow chasingdreams	0
best not to fly dub lgw if you are in a hurry. flights now have to bus to the main terminal because of uk immigration	0
okay wait wait who is in one of the good rows at awp and wants to split a table late awp	0
the whole doctors office is losing their shit bc they got new computers.. is n't technology supposed to make things easier annoyed late	0
i havent tweeted in ages late	0
day coming back content is exactly why we need to restart ourselves from our everyday routines. late 365days	0
facebook have built a drone and it does n't look creepy at all sarcasm	1
so today 's challenge getting the year old to nursery late lazylizzie	0
girls do n't do this when you 're late nailpolish late hurry	0
when you decide to grab a coffee before work and you get the new guy late	0
opps did n't know i worked today all well lookingghettoasfuck late	0
alright flight crew where the hell are you jetblue delays iwannagohome jfk late	0
i 'm going to miss my best friend going off to college. i love you hope it 's a great year late	0
it had been a long time since last time i had to run to catch the bus late	0
best part of having everyone at work love me is that they all cover for me late	0
unfriended was too fucking good. late	0
after so long i finally gave into twitter late	0
can i get ready for work in late	0
when your mind says get ready for work but your body says stay in bed. late night	0
can i get ready for work in late	0
it 's a beautiful monday when you get a call from someone 's angry wife. journalism sarcasm	1
late dinner alone alone dinner late	0
decades on earth thank you lord late	0
guess my sleep schedule has caught up with me. late	0
time to learn how to tweet late	0
people ca n't drive in the rain. late	0
you know it is going to be a great week when you work hours on a monday sarcasm	1
make up primer is the best i just discovered it. it 's definitely a make up bag staple now. late beauty makeup	0
should n't have snoozed my alarm for the fourth time late	0
sometimes we have to do things we are not ready for before it 's too late	0
what is going on my phone wants to work half of the time ca n't get my messages on fb or nothing ugggh late	0
slowly and silently walk in to the office. late	0
thank you for taking the time to answer my texts means a lot sarcasm	1
late night in the garden. a bit glamorous. garden roses late night	0
hey awesome people check out my moviereview of dope dopemovie drugs	0
twitter is actually free hmmm okay thanks. late	0
what you 're really telling people when you 're 'just a minute or two late entrepreneur	0
i love when random people who have never interacted with me start following me. sarcasm wtf	1
teach your kids photography and they will never have money for drugs photo fusion studio	0
now the express train will be a local train. that way it will be even more late unbelievable subpar	0

better late then never is the best policy when getting to my yoga. it 's noon and i 'm jumping in. yoga meditation	0
i blame obama and the kardashians for of america 's problems. kidding sarcasm	1
signs that a loved one is abusing drugs addiction treatment rehab recovery	0
tea such a good way to start the day late summer mornings	0
watching the athletics has become like wwe. it 's fiction. i 'm waiting for bolt to put gatlin through the table tlc drugs	0
thanks to the extremely hot weather and my allergies i feel like dying. absolutely lovely day is n't it sarcasm	1
get up early go to bed late what an unhealthy lifestyle lifestyle healthy not late	0
saturday night been outside very nice enjoy yourself be safe we are downtown sanantonio popped up and under lights late local	0
ai n't it funny how drugs are celebrated for improving health but frowned upon when it comes to improving performance worldchampionships	0
when you got your life together but the cab driver do n't late	0
we will never understand till its too late	0
thank you sun it 's always been my dream to faint because of heat exhaustion sarcasm takeiteasy summer2015	1
a little late in the day funny for our wow humor wow wordoftheweek humor	0
so happy i finally got to watch the gopdebate beinformed	0
guns in banks to protect our money but no guns in schools to protect our children late term abortion maybe	0
there are two kinds of politiciansthose trying to get an investigation started and those trying to get one stopped. joe	0
nobody 's perfect. except fat girls. sarcasm	1
types of men i would never date humor offbeat funny fun	0
be careful who you make angry humor	0
when you 're supposed to be home before dark and your automatic lights come on late	0
my dad got jumped by a black man and i just stood there taking a video meme humor	0
i love when a change in plans by someone else completely alters the plan you had sarcasm	1
india where bus drivers honk even inside an airport. india humor honking driver fun	0
will the iphone influence human evolution news tech football	0
living an hour away from school instead of a couple minutes away really sucks. late	0
what 's the point of having a train schedule when it 's always late. amtrak train late why	0
if you lie to people to get their money that 's fraud. if you lie to tehm to get their votes that 's politics. joe moore	0
you never realize how short a month is until you pay alimony. joe moore quotes humor quote	0
today 's cartoon british version. cartoon humor zombies	0
children are natural mimics they act like their parents in spite of every attempt to teach them good manners. parenting humor teamjesus	0
breaking news donald trump does n't think donald trump 's book is the best of all time. politics gop usa	0
take the dog shopping it 'll be fun they said humor	0
donald trump and hillary clinton revealed to be distant cousins politics tech military	0
just waking up is not cool late	0
a boss with no humor is like a job that 's no fun. figuremeout teamfollowback followback followme ifollowback	0
i 'm okay now i 've just taken my happy pill humor drugs pills bipolardisorder bipolar depression read iartg	0
the power of idiots. Humor	0
when you know your screwed funny humor humour	0
why is my weather app promoting smoking all day humor	0
i love summer allergies sarcasm sickofsneezing	1
we 'll never tell wisdomoftheday humor	0
no matter how bad a child is he is still good for a tax deduction. joe moore quotes humor quote	0
once again south sudan is south sudan again peace for all and yea we made it olympic games things getting	0
women are funny. and this is a surprise because .why humor	0
indeed another truth which should be educated to europeans are the facts abt immigration. scary comments around there peace	0
it 's not love until there 's attempted homicide. lovequote humor	0
the pain given by the one who you love the most are not curable. sarcasm truthoflife deltathoughts	1
some1 just said they say working together we can do more but after elections its vukusenzele politics	0
let go of what you want others to be and start working on yourself and what you want to be. hooponopono letgo	0
the twitter sign has a white bird. like a dove of peace	0
we 've decided to stop killing each other and to start helping each other unconditionally instead peace betterinternetstories	0
no classes for today. yey. i really need to watch the mv already. 5hmonstervideo worthitvma late	0
you ca n't tell me i 'm not a cutie late flashbackfriday	0

Fig 3.3 Sample pre-processed dataset

3) **Feature Extraction and Reduction**

In the proposed models, we have used the TFID, which calculates the frequency of the words. The Random Forest will work on the frequency of the words, and most likely frequency of words will take into consideration. Later, the PCA algorithm will take input the words with the frequency as input and remove the words which have the least frequency for further processing. The PCA technique is used to reduce the number of extracted features. The PCA technique assesses the data's effective amount of variance by creating a low-dimensional representation of the data. PCA (Principal Component Analysis) is a type of mathematical technique that transforms a collection of related variables to make a set of linear subsets which are unrelated and depend on a transition resulting in the form of uncorrelated variables. It is defined as an orthogonal linearly transformation which allows projecting the initial set of data to the second projection system with the ultimate objective that the first coordinate is projected by the largest variance, the second largest variance which is having a second coordinate projection, and it is vertical to the existing first component. Primarily, PCA supports a linear transformation expressed as where and for enhancing the data variance in the projected space. Considering a data matrix defined £ = (z" zt , z;}, *x; o* R^d , *z C* A" and r € *d*, a group of p-dimensional vectors of weights characterizing the transformation by *W* -- (wt, wt, ..., w$}, w2 e R^k, that matches every z; vector of X to a

$$t_{k(i)} = W_{/(i)} T_{x\,i} \tag{3.1}$$

An original weight W1 must adhere to the condition as expressed to increase the variance:

$$W_i = \arg argmax_{|w||} = \{\sum (x_i . W)^2\} \tag{3.2}$$

Below is the further extension of the previous condition:

$$W_i = \arg argmax_{\| w \|=1} \{ \ \|X.W\|^2 \}$$

$$= \arg argmax_{\| w \|=1} \{ \ W^T X^T X W \} \tag{3.3}$$

A symmetric grid for, e.g. X^T where X can effectively be calculated by finding the largest eigenvalue from the matrix, as W is considered as the linked eigenvector. If is acquired, the first principal component will be extracted by projecting the initial data matrix onto the space resulting from the existing transformation. Following the reduction of the recently acquired components, the other segments can also be obtained along these lines.

4) **Clustering of Similar Information**

The phase of clustering deploys K Means Clustering for the same kind of information clustering. The cluster formulation means the clustering of similar information. The similar information means the word frequency. The words which have a similar frequency index will be a cluster in one cluster and other in the next. The number of centroids will be equal to number of clusters as we have chosen the number of centroids based on number of clusters**.** The k-means algorithm first chooses K points from the data patterns as the initial clustering center. Second, it computes the distance from each sample to the cluster's center. The classification of sample is performed into the class nearest to the cluster's center. Third, the new clustering center is obtained by computing the average value of every recently created clustering data object. Eventually, all these steps are iterated until there is no change in the clustering center of two adjacent times, which depicts that the change in sampling is complete and the clustering principal function has reached the highest value. To execute the algorithm, the distance among data samples is computed using Euclidean distance, and the clustering performance is estimated using the error square sum criterion function. In a sample set $D = \{x_1, x_2, \ldots . x_m\}$, K-means algorithm splits the clusters into $C = \{C_1, C_2, \ldots . C_k\}$to make the squared error minimum, just like equation shows

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|_2^2 \qquad (3.4)$$

Where $\mu_i = \frac{1}{|c_i|} \sum_{x \in C_i} x$ as per Eq2 denotes the mean vector of Ci cluster.

5) **Classification**

In this step, we have designed four different ensemble models for the detection. In each model, different classification algorithms are applied, which are SVM, KNN, MLP, logistic regression and decision tree.

Support Vector Machine algorithm is emphasized on generating a hyperplane for expanding the margin, the distance from the hyperplane to the nearest data from a class, as shown in Figure 3.4. When the margin is large, and the error is least, this is known as generalization. The initial optimization issue is expressed as:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i \tag{3.5}$$

s.t. $\quad y_i(w \cdot x_i + b) \geq 1 - \xi_i, i = 1,2, \ldots N$

$$\xi_i \geq 0, \qquad i = 1,2, \ldots N$$

In which $w \cdot x_i + b$ denotes a hyperplane with weight parameter$w$ and bias parameter$b$, $C > 0$ denotes a regulation metric that assists in controlling the balance amid least misclassification and highest hyperplane margin, and the slack variable is represented with ξ_i. The slack variable is utilized to perform misclassification at some distances. In case $\xi_i = 0$, this illustrates that ith data is located right at the margin or on the right side of the margin. In case $0 < \xi_i \leq 1$, this represents that ith data is present in the margin on the right side. When $\xi_i > 1$, this implies that $i$thdata is available on the wrong side and misclassified. This issue can be expressed as a dual problem as

$$\min \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N} \alpha_i \tag{3.6}$$

$$s.t. \quad \sum_{i=1}^{N} y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1,2, \ldots, N$$

In which, the Lagrange multiplier is defined withα_i. The weight vector is represented as $w = \sum_{i=0}^{N} \alpha_i y_i x_i$.

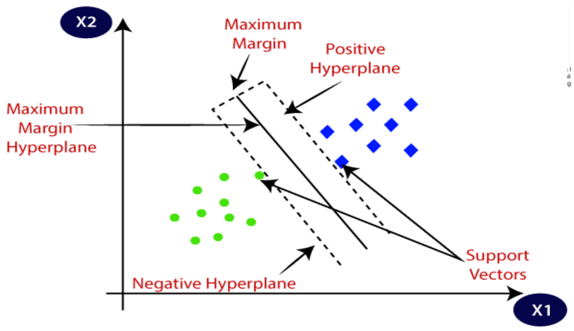


Figure 3.4: SVM Algorithm

K Nearest Neighbor is a simple and principal classification algorithm which assists in recording all the categories in correspondence with the training data, as shown in Figure 3.5. In the case of matching of features of the test object exactly with the features consisted in a training object, the classification is performed. The KNN algorithm is generated on the basis of defined situations. This algorithm emphasizes computing the distance among the nodes as a non-similarity index among nodes to avoid the matching problem among nodes in which Euclidean distance or Manhattan distance is executed as:

:

$$d_{ij} = \sqrt{\sum_{k=1}^{d} (x_{ik} - x_{jk})^2} \tag{3.7}$$

$$d_{ij} = \sum_{k=1}^{d} |x_{ik} - x_{jk}|$$

Simultaneously, K-Nearest Neighbor aims to make the decisions on the basis of dominant categories of k objects instead of on a single object category.

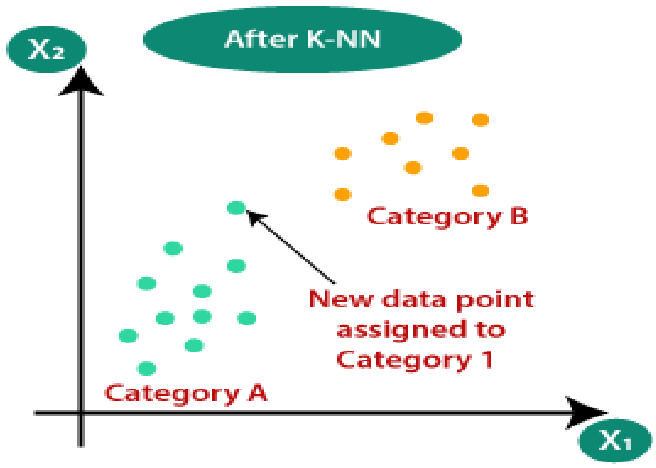


Figure 3.5: KNN Algorithm

Logistic Regression has y as a dependent variable which takes only two values0 and 1. The hypothesis is that the probability p for y=1 is defined in the presence of independent variable x is

$$p = P(y = 1|x) \tag{3.8}$$

Afterwards, odds ratio of the event can be expressed as:

$$odds = \frac{p}{1-p} \tag{3.9}$$

LR model is a linear regression model amid the logarithm In odds and independent variable which generates the odds ratio, such as:

$$In\ odds = \beta_0 + \beta_1 x \tag{3.10}$$

In which β_0 and $\beta_1$denotes the regression coefficients. At the moment, association of probability p with the independent variable is defined as:

$$p = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}} \tag{3.11}$$

This is recognized as the logistic function.

Decision Tree is a common classification algorithm which is adopted in various applications in real-world. This symbolic learning method focuses on correlating the information taken from a training dataset in a stratified structure obtained. The nodes and ramifications are comprised in this dataset. Decision Tree concentrates on alleviating the least squares error for the next split of a node in the tree so that the average of the dependent variable comprised in all training instances covered for unseen instances in a leaf can be predicted. A DT model $T(x; \{R_j\}_{j=1}^{J})$ is capable of partitioning the $x-$ space into Jdisjointregions $\{R_j\}$ and predicting a separate constant value in each one as:

$$x \in R_j \Rightarrow T(x; \{R_j\}_{j=1}^{J}) = \hat{y}_j \tag{3.12}$$

Or equivalently

$$T(x; \{R_j\}_{j=1}^{N}) = \sum_{j=1} \hat{y}_j I(x \in R_j) \tag{3.13}$$

In which,$\hat{y}_j = \frac{1}{a_j} \sum_{i=1}^{a} y_i$denotes the mean of the response y in each region R_j, $y_i \in R_j$, $a_j$represents the size of region R_j. Hence, a tree assist in predicting a constant value $y_j$in each region R_j. The top-down iterative splitting is implemented on the basis of a least squares fitting criterion to construct the trees. In this algorithm, the identities of the predictor variables, that are useful to perform splitting and their corresponding split points, utilize to resolve the regions $\{R_j\}_{j=1}^{J}$ of the partition.

Multi-Layered Perceptron is an effective FFNN (feed-forward neural network) in which common and popular classes of NNs are comprised to process an image and recognize the pattern. A number of subsequent layers having perceptron-type are included in this algorithm, such as an input layer which assists in acquiring the external inputs, a set of hidden layers and one output layer.

Assume $x_l$as the input signals to Multilayer Perceptron, the output value obtained from the jth hidden neuron is defined as:

$$y_{lj} = f\left(\sum_{i=1}^{n} x_{li} w_{ij}\right) \tag{3.14}$$

In which $f$is the activation function and considered as the connection weight from the ith input neuron to the jthhidden neuron. Afterwards, the evaluation of the final output value from the output neuron is done as

$$y^{out} = f\left(\sum_{j=1}^{k} y_{lj} w_{j}\right) \tag{3.15}$$

In which k is utilized to denote the number of hidden neurons and $w_j$defines the connection weight from the jth hidden neuron to the output neuron.

6) Applying Voting Classifier and analyzing performance

Before designing ensemble classifiers, individual classifiers are implemented for sarcasm detection. The classifiers which gave maximum performance are considered in the ensemble process. The classifiers which have a similar classification process are combined for sarcasm detection. Data is apportioned as training set and test set with 80-20 split. The training set is used for training of the model and model is tested on the test set**.** The soft voting has been used in the research work. The soft voting is the bagging process in which voting classifier will bag strong features of classifiers which are involved in the voting process. When the voting classifier bags the strong features of the classifiers, it leads to an increase in performance. The voting classifier would be applied to the test data, and the result would be predicted based on the performance parameters.

3.5 Designing ensemble classification model

Spyder, a robust development environment for the Python language with advanced editing, testing, and numerical computation environments, was used to implement the code altogether. Pandas library is used for handling datasets. Scikit-learn libraries are used for feature representation, classification, similarity measures and evaluation purposes.

The first ensemble classification model is the combination of Support Vector Machine K-nearest neighbor and decision tree. The detailed model is explained in Figure 3.4.

The second ensemble classification model integrates Support Vector Machine, Logistic Regression and Decision Tree. The detailed model is explained in Figure 3.5.

The third ensemble classification model integrates Multilayer perceptron, Logistic Regression and Decision Tree. The detailed model is explained in Figure 3.6.

The fourth ensemble classification model is the combination of MLP, logistic regression and SVM. The detailed model is explained in Figure 3.7.

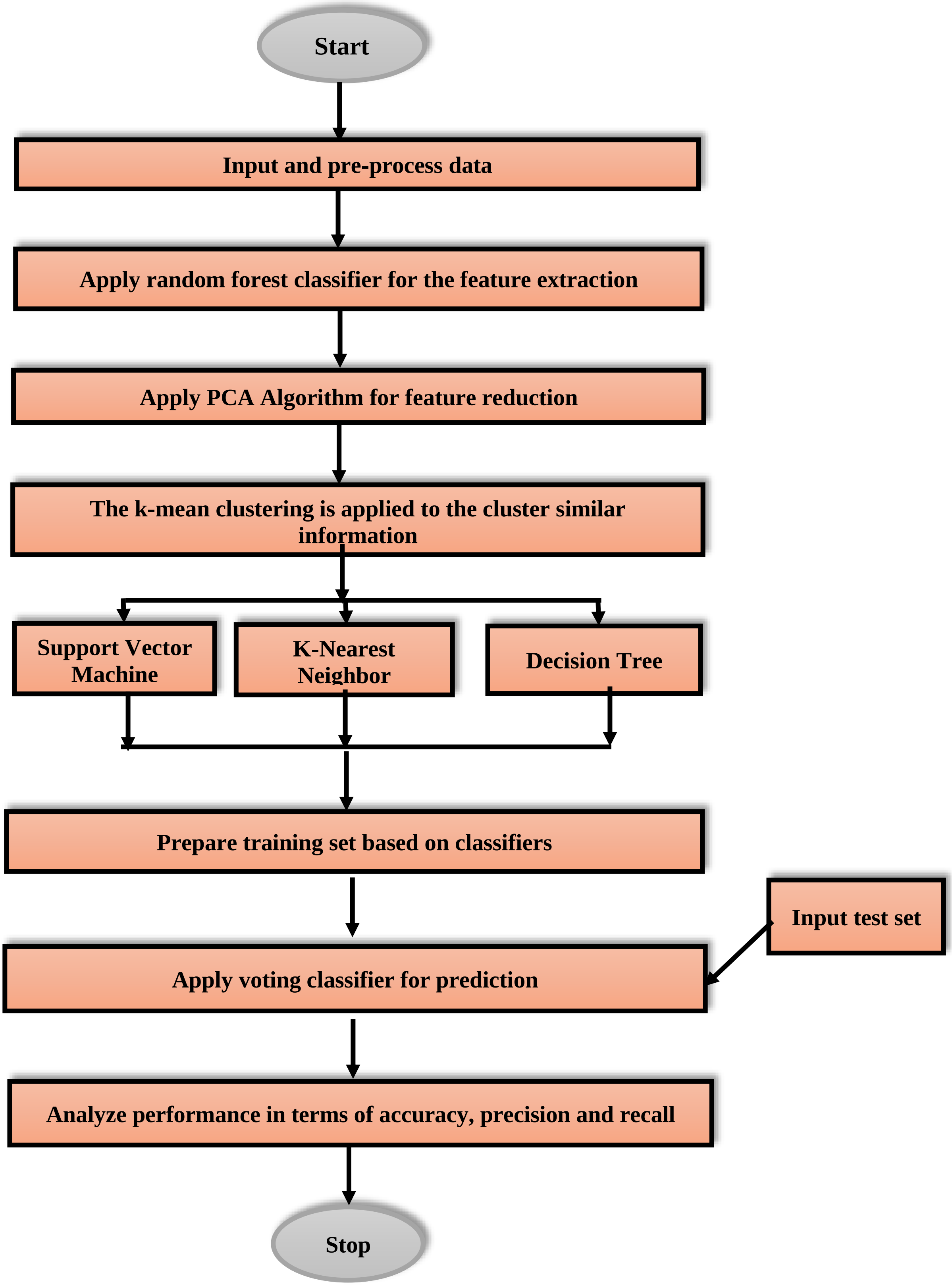


Figure 3.6: Ensemble 1 Classifier (SKD)

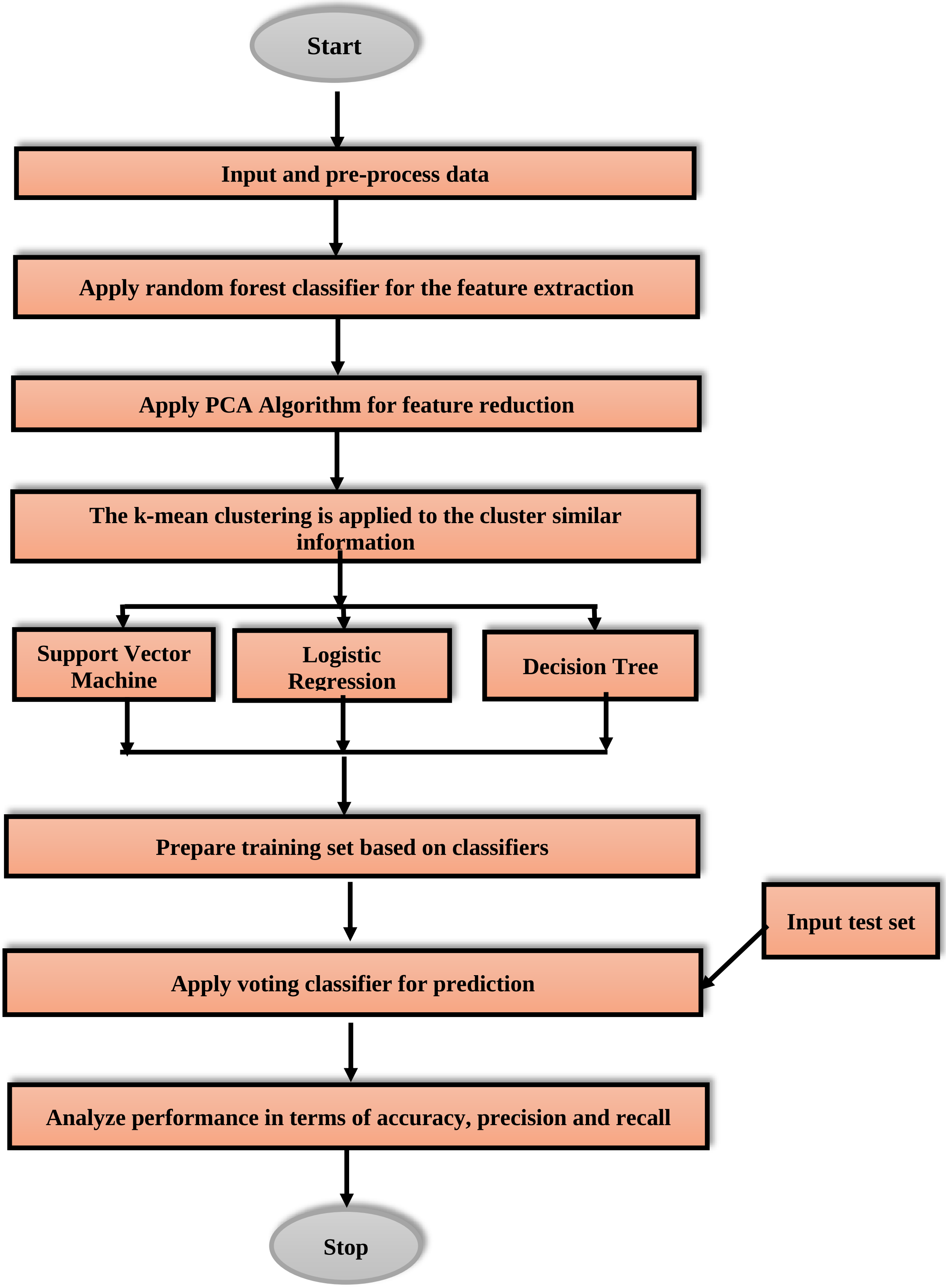


Figure 3.7: Ensemble 2 Classifier (SLD)

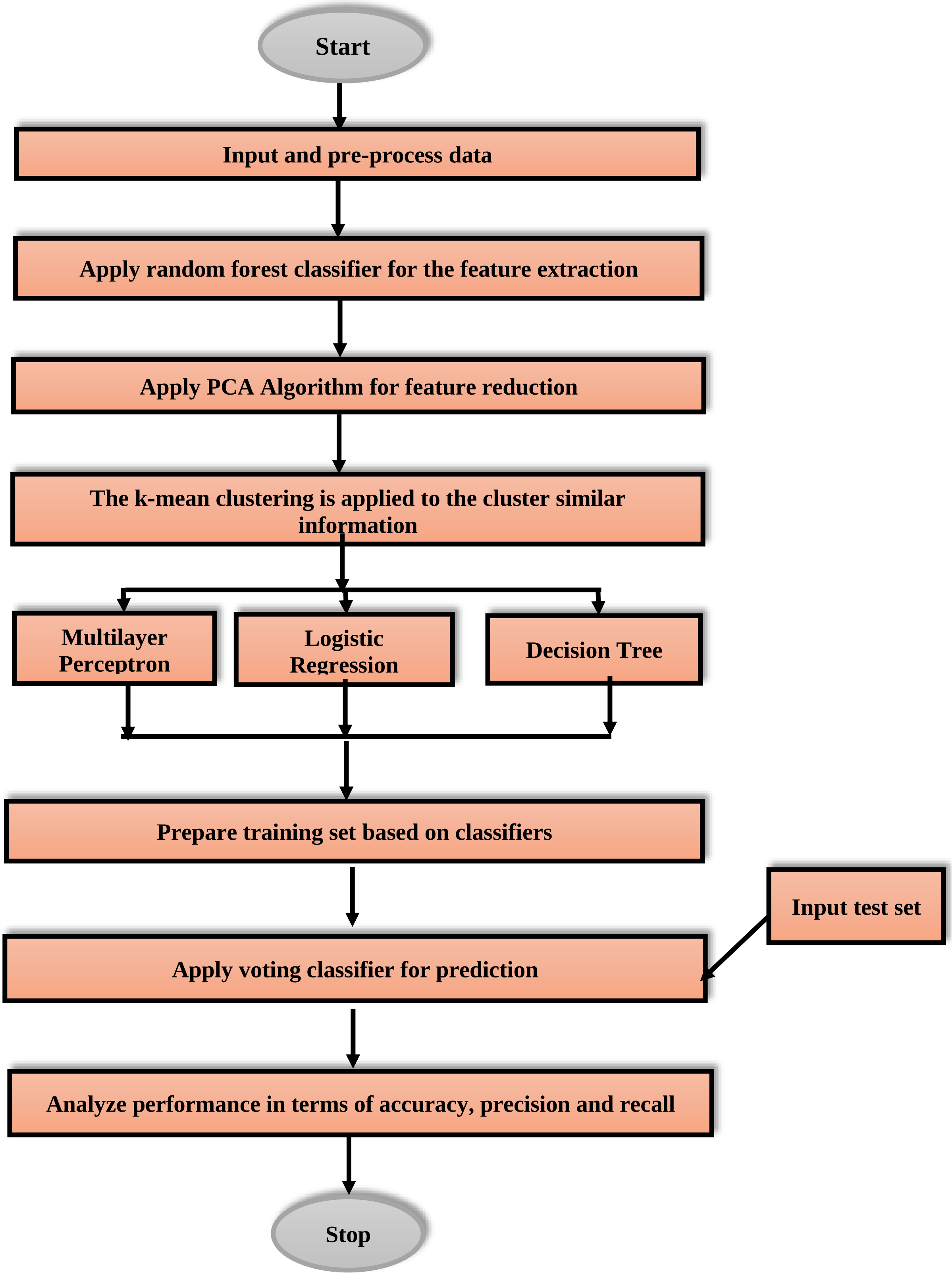


Figure 3.8: Ensemble 3 Classifier (MLD)

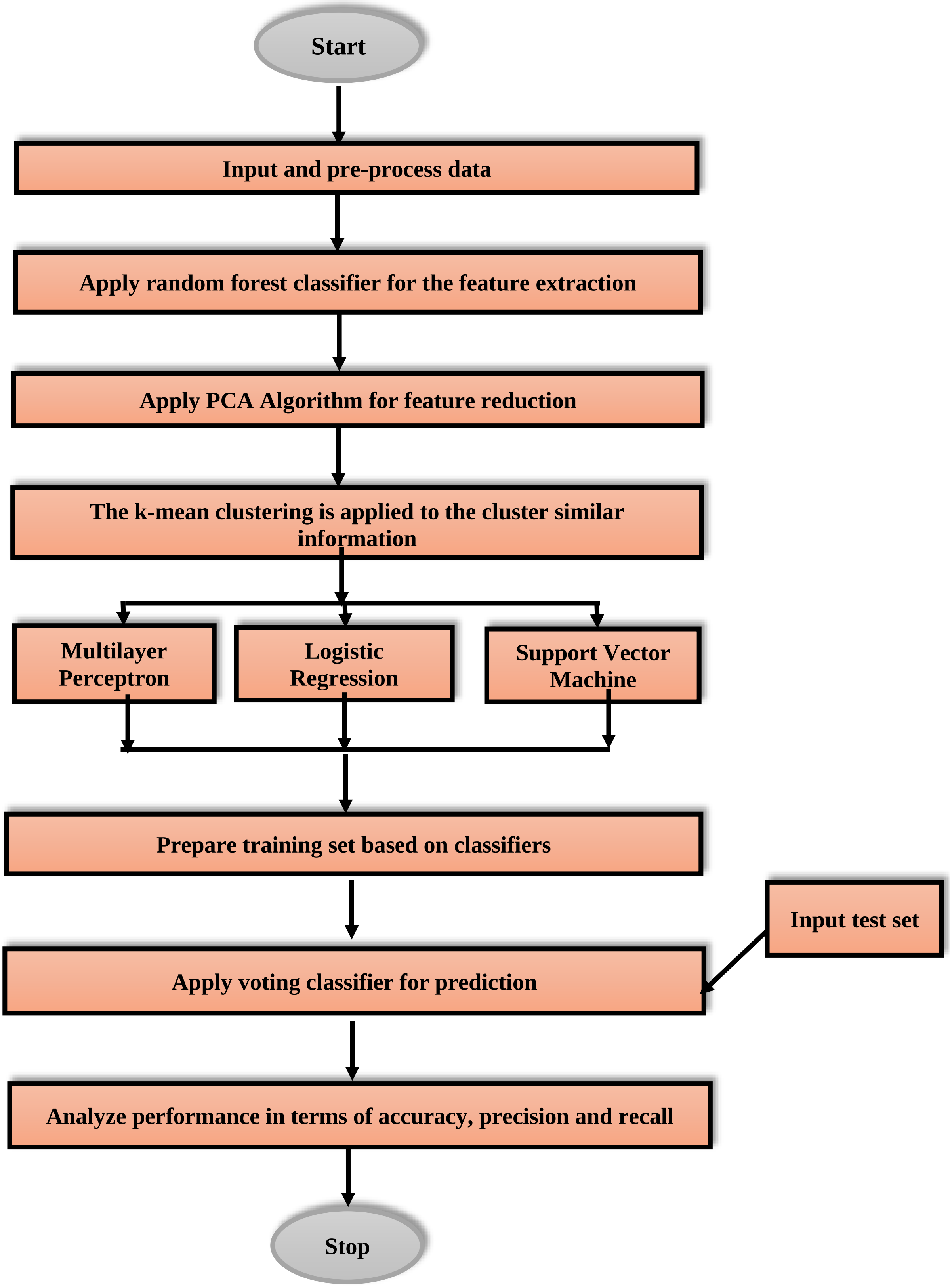


Figure 3.9: Ensemble 4 Classifier (SLM)

3.6 Summary

In this chapter, in order to deal with sarcasm detection for the improvement of sentiment analysis techniques, an approach of ensemble classifier is proposed. All the steps involved were explained in detail, including the details of the datasets used. The TFID model is applied to the data. The TFID model calculates the word frequency. The other dimensionality reduction algorithms can also be applied. The random forest is used in very few cases for feature extraction; it is generally used for classification, which gave uniqueness to our proposed model. The PCA is the most efficient algorithm for the feature reduction which is already proved in the existing research work. When we are applying K-mean clustering, it will cluster words which have a similar frequency. The output of the K-mean clustering will be clustered data; it will be given as input to the classifiers as the target set. The SVM, KNN or DT classifiers are given as input to the voting classifier. When the voting classifier is prepared, then the training set is prepared for the classification. Four ensemble classifiers were constructed, including multiple classifiers.

The next chapter presents the performance of the proposed ensemble approach on the datasets and further compares it to existing algorithms.

CHAPTER 4

PERFORMANCE COMPARISON OF EXISTING SARCASM DETECTION TECHNIQUES TO THE PROPOSED ENSEMBLE TECHNIQUE

This chapter first discusses the various performance parameters for evaluating the process of sentiment analysis in order to detect sarcasm. In terms of accuracy, precision, recall, and f1-score, the proposed ensemble approach is compared to existing sarcasm detection algorithms.

4.1 Performance parameters

The performance of the classifier is described with the help of a Confusion Matrix, as shown in Fig 4.1, by comparing actual results versus the predicted results. True positives, true negatives, false negatives, and false positives are among the terminology used in the parameters for evaluating sentiment analysis. These are the phrases used by a classifier to compare the class labels assigned to documents with the classes to which the objects really belong. Positive phrases that are positive are categorized as such. The classifier does not mark false positives as positive classes, even when they should have been. The classifier accurately labels true negative phrases as being in the negative class. False negative terms are those that are not classified as belonging to the negative class by the classifier but should have been.

		ACTUAL	
		Positive (1)	Negative (0)
PREDICTED	Positive (1)	**True Positives**	**False Positives**
	Negative (0)	**False Negatives**	**True Negatives**

Fig 4.1: Confusion Matrix

Following are the parameters for evaluation of performance

A. **Precision and recall:** Precision and recall are two metrics commonly used to assess performance in text mining and related fields such as information retrieval. These are the parameters that are utilised to determine exactness and completeness.

$$\text{Precision}=\frac{\text{True Positive}}{\text{True Positive}+\text{False Positive}} \tag{4.1}$$

$$\text{Recall}=\frac{\text{True Positive}}{\text{True Positive}+\text{False Negative}} \tag{4.2}$$

B. **F-measure:** The F-Measure is the harmonic mean of precision and recall. F-measure yields a value that strikes a balance between precision and recall.

$$\text{F measure} = \frac{2*\text{recall}*\text{precision}}{\text{precision}+\text{recall}} \tag{4.3}$$

C. **Accuracy**

The most common metric for assessing categorization ability is accuracy. The proportion of correctly identified examples to the total number of instances is defined as accuracy, whereas the proportion of incorrectly classified instances to the total number of instances is described as error rate.

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative} \tag{4.4}$$

4.2 Performance of the proposed approach

In this section, four different proposed ensemble classifiers are compared for detecting sarcasm. Ensemble classifiers are a collection of different classifiers. The first ensemble classifier combines SVM, KNN, and decision tree. The second ensemble classifier combines SVM, logistic regression and decision tree classifiers to detect sarcasm. The third ensemble model combines MLP, logistic regression, and decision tree, while the last combines MLP, logistic regression, and SVM. All four ensemble models are tested on the dataset. Each ensemble classifier's performance is measured in terms of accuracy, precision, recall, and f1-score as per the equations 4.1- 4.4.

The accuracy for SKD, SLD, MLD and SLM is 92.90%, 93.1%, 92.8% and 93% respectively. Precision for SKD, SLD, MLD and SLM is 92%, 92%, 92% and 93 % respectively. Recall for SKD, SLD, MLD and SLM is 93%, 93%, 93% and 93% respectively. F1-score for SKD, SLD, MLD and SLM is 92%, 93%, 92% and 93% respectively. The performance measures have a slight degree of difference as in the methodology; all other steps were the same, and only different classifiers were ensembled, which gives the difference. On comparing all results, the best ensemble was found.

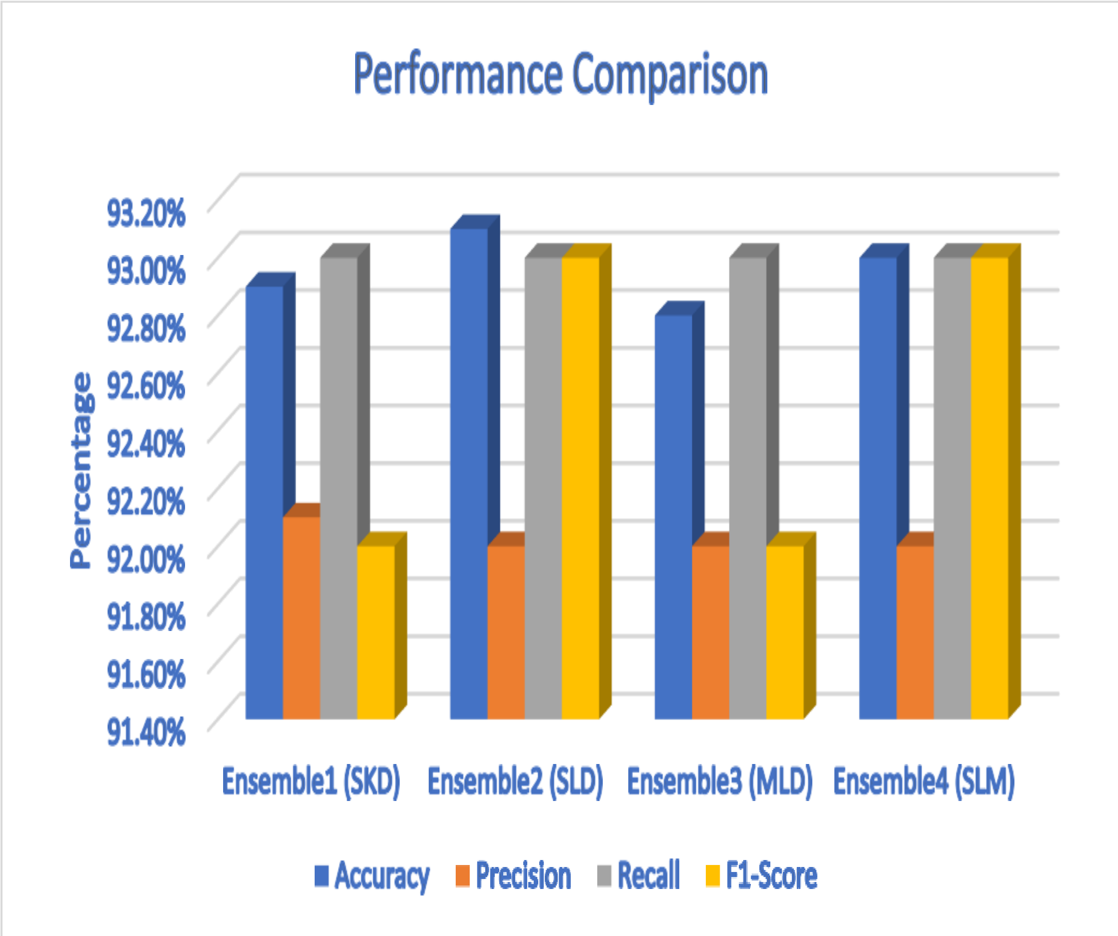


Fig 4.2: Performance Analysis of proposed ensemble classifiers

As shown in Fig 4.2, the ensemble classification model 2 (SLD) gives maximum accuracy of 93.10 percent on the dataset for sarcasm detection.

Figure 4.3 to 4.6 represents the snapshots of the outcomes achieved by the proposed approach.

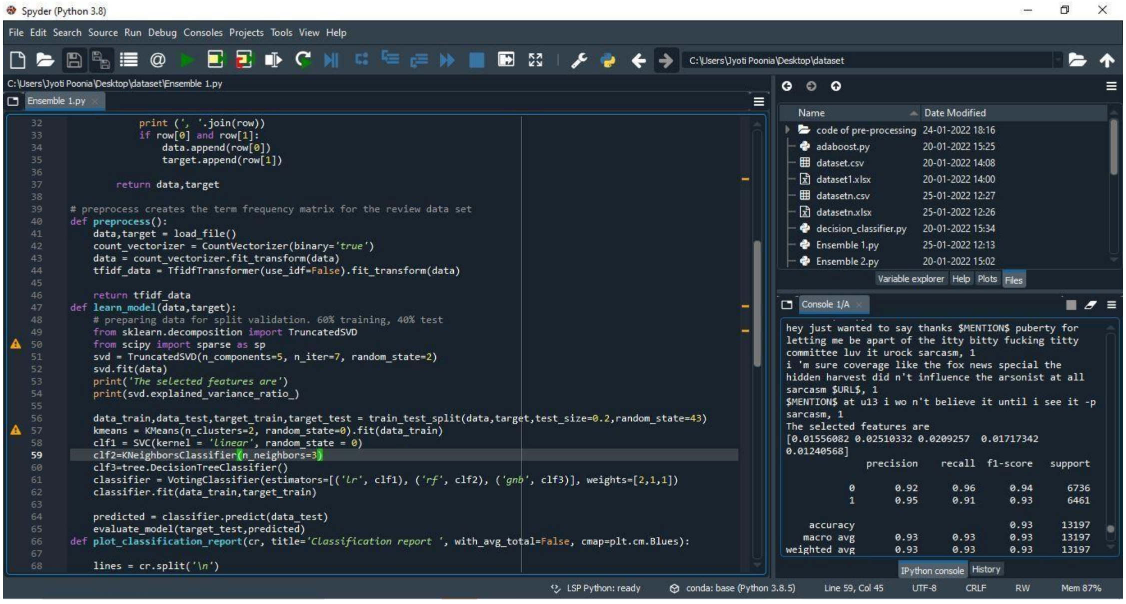


Fig 4.3: Outcome for Ensemble 1(SKD)

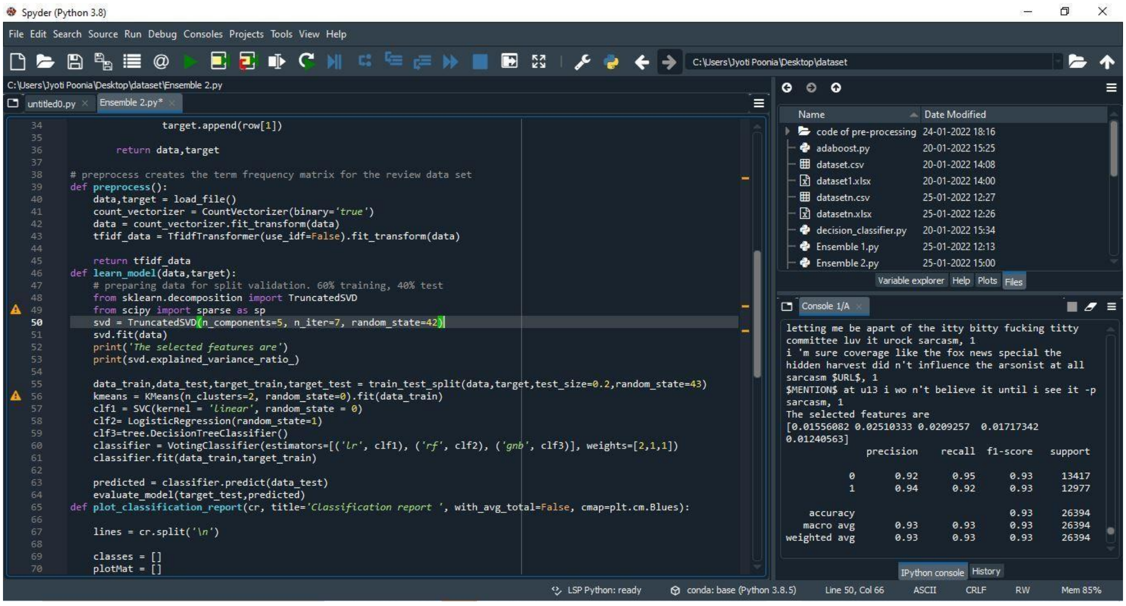


Fig 4.4: Outcome for Ensemble 2(SLD)

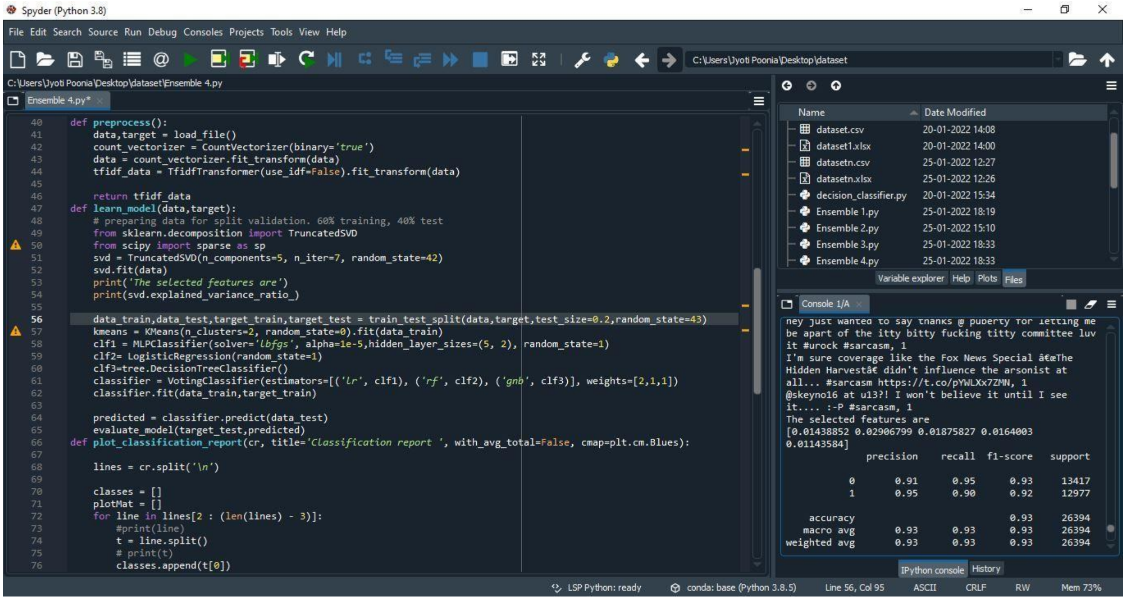


Fig 4.5: Outcome for Ensemble 3(MLD)

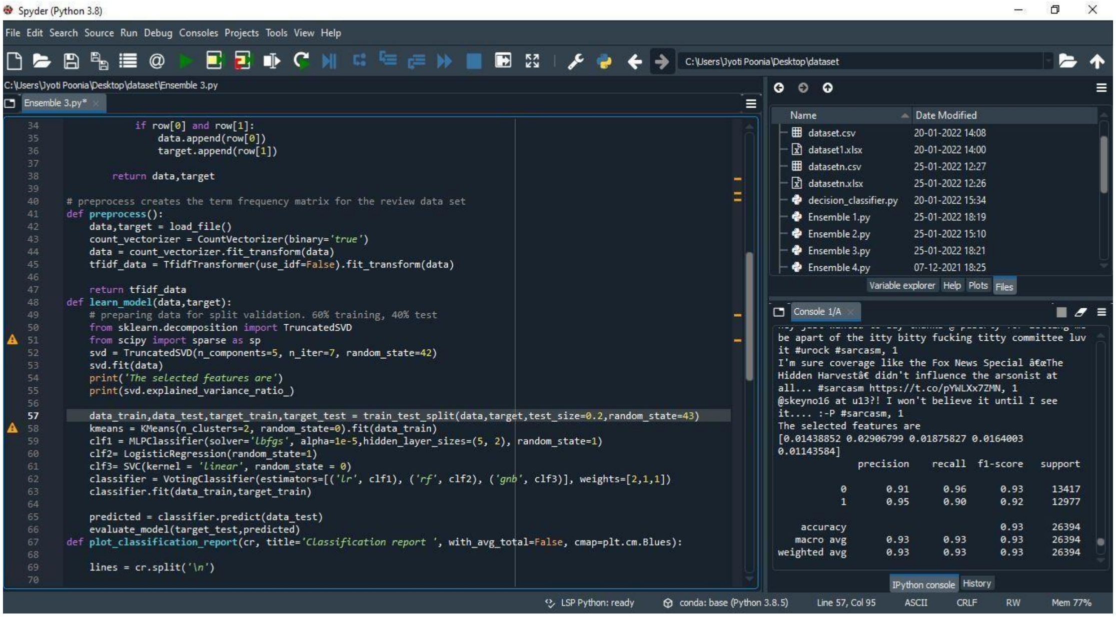


Fig 4.6: Outcome for Ensemble 4(SLM)

4.3 Performance comparison of proposed classifier with traditional approaches

The presence of sarcasm in the text is a great challenge when the sentiments are analyzed. To increase sentiment analysis performance, numerous state-of-the-art algorithms for detecting sarcasm have been developed. On the dataset, we compared the performance of the Adaboost classifier, Decision Tree classifier, Random Forest classifier, and K- Nearest Neighbor classifier to the suggested ensemble classifier in this section.

The accuracy of Adaboost, Decision Tree, Random Forest, K-Nearest Neighbor and proposed approach on the dataset is 89.25 %, 89.72 %, 53.42 %, 87.51 % and 93 %. The precision of Adaboost, Decision Tree, Random Forest, K-Nearest Neighbor and proposed approach is 89%, 90%, 71%, 88% and 93%. The recall of Adaboost, Decision Tree, Random Forest, K-Nearest Neighbor and the proposed approach is 90%, 90%, 53%, 88% and 93%. F1 Score of Adaboost, Decision Tree, Random Forest, K-Nearest Neighbor and proposed approach are 89%, 90%, 40%, 88% and 93%.

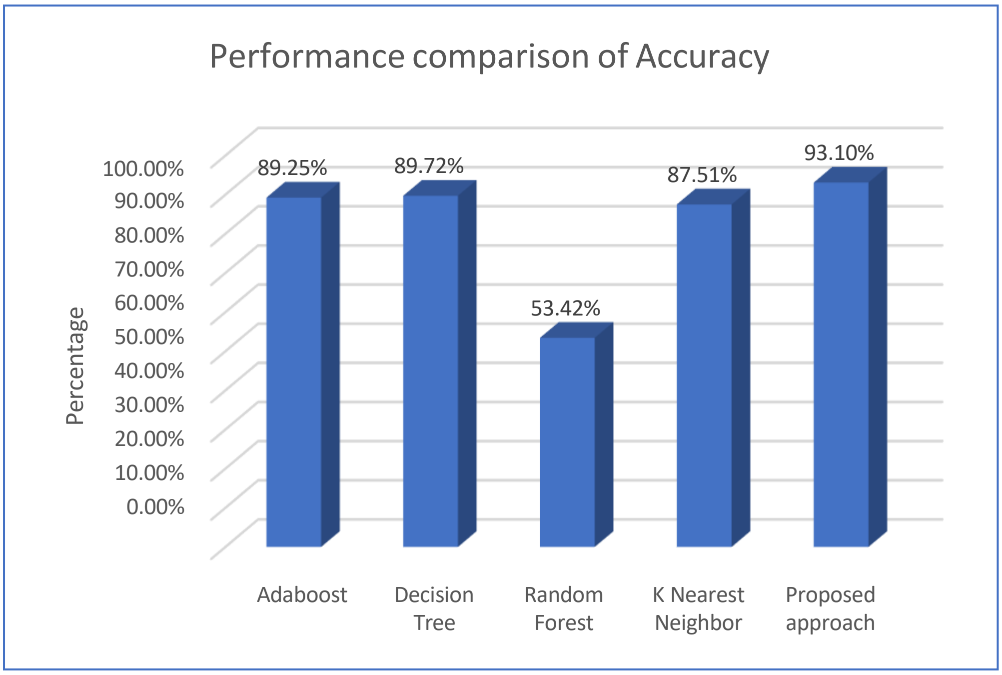


Fig 4.7: Accuracy Comparison

Proposed ensemble classifiers have outperformed other classifiers, as demonstrated in Figure 4.7. The accuracy of the proposed approach improved by 4.31 percent, 3.76 percent, 76.15 percent and 6.38 percent, respectively, when compared to Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor classifier.

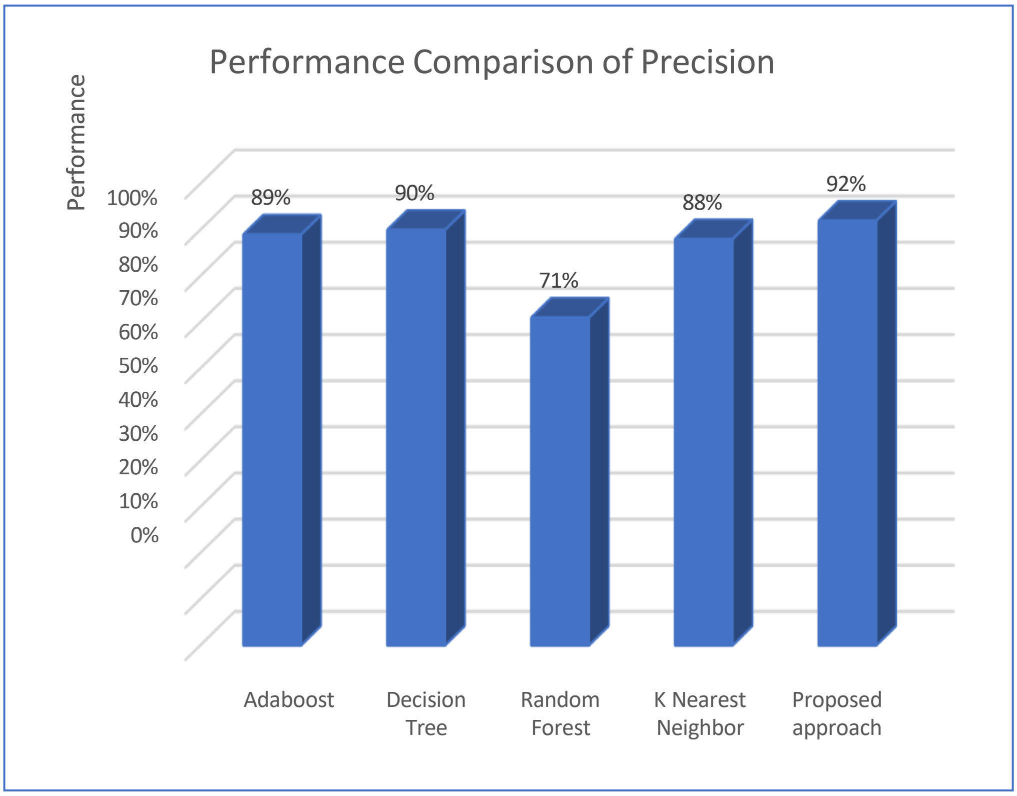


Fig 4.8: Precision Comparison

As shown in Figure 4.8, the precision of proposed approach improved by 3.37 percent, 2.22 percent, 29.57 percent and 4.54 percent, respectively, compared to Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor classifier.

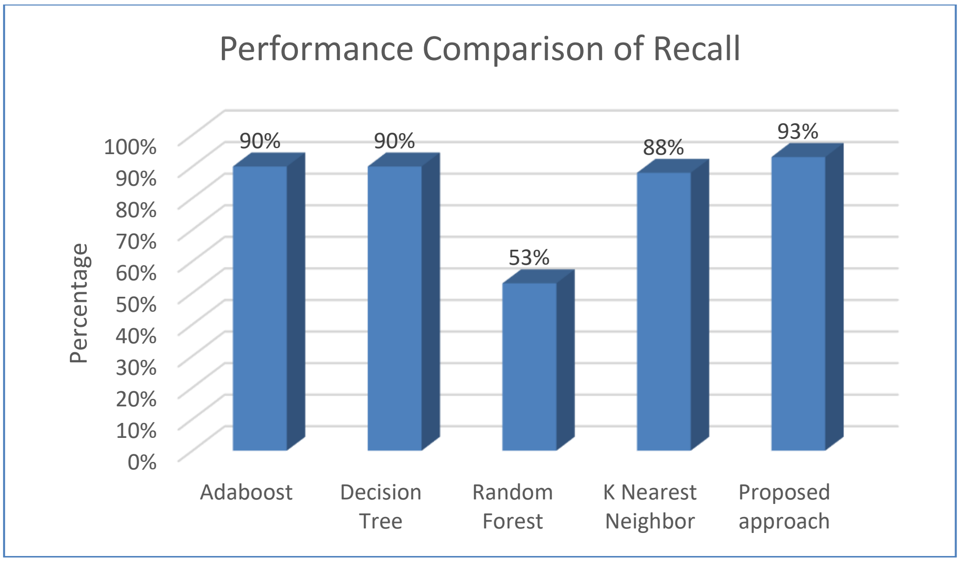


Fig 4.9: Recall Comparison

As shown in Figure 4.9, the recall of proposed approach improved by 3.33 percent, 3.33 percent, 75.47 percent and 5.68 percent, respectively, compared to Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor classifier.

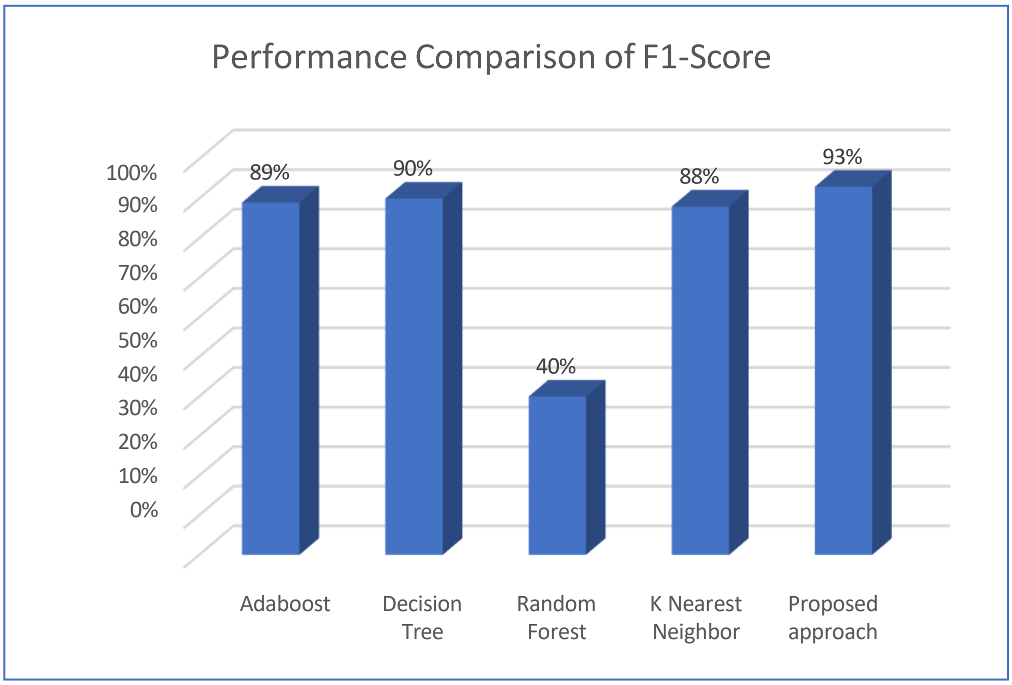


Fig 4.10: F1 -Score Comparison

As shown in Figure 4.10, the F1-score of proposed approach improved by 4.49 percent, 3.33 percent, 132 percent and 5.68 percent, respectively, compared to Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor classifier.

4.4 Performance Comparison of proposed algorithm with existing ensemble techniques

The proposed model's performance is compared to numerous state-of-the-art sarcasm detection techniques. The selected models are implemented and evaluated on the same dataset (Tweets and News Headlines) for a fair comparison [149]. The experimental findings of various models are compared in Table 4.1. The results reveal that the proposed strategy outperforms existing methods in recognizing sarcastic tweets across multiple domains as in the proposed ensemble approach, random forest classifier is applied for feature extraction, PCA has also been used for feature reduction, and later multiple classifiers are ensembled, which has improved the performance accuracy.

Table 4.1: Performance analysis of proposed ensemble classifier with existing ensemble techniques

Citation	Technique	Accuracy
[146]	GRU-based neural network	90%
[149]	CNN-LSTM	91.6%
[97]	Chi2, TFIDF, Voting classifier	89%
[113]	CNN, Glove	86.1%
[155]	Nave Bayes classifiers, Random Forest classifiers, Support Vector Machines, Logistic Regression as base learners	73.33%
	Proposed approach	**93.1%**

As seen in Table 4.1, the proposed approach has outperformed other existing ensemble techniques, achieving the highest accuracy of 93.1 %.

4.5 Summary

In this section, the results of the ensemble classifiers are compared to those of other existing ensemble classifiers and also stand-alone classifiers such as Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor, on the dataset. It was observed that the ensemble classifier generates the highest accuracy because of its robustness compared to other machine learning models. Also, the logical structure of the model enables it to achieve relatively better performance.

The next chapter concludes the thesis and presents future directions.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1. Conclusion

Sentiment analysis is a task which evaluates an opinion expressed as positive, negative or neutral, but the presence of sarcasm flips the polarity of the text. Sarcasm is a kind of verbal irony which emphasizes expressing ridicule. Sarcasm has a negative implied sentiment. However, it is free of negative surface sentiment. A sarcastic sentence may carry positive, negative or no surface sentiment. There are various kinds of techniques that detect sarcasm. Chapter 1 explains the basics of sentiment analysis, its process, applications and different issues. Based on the existing problems, multiple research objectives have been framed to carry out the research. Chapter 2 presents the literature review highlighting sentiment analysis and sarcasm. The Sarcasm detection techniques have different phases in which data is pre-processed, features are extracted, and reduced clustering and classification. After an exhaustive review, it becomes apparent that there is a scope for further improvement in accuracy to detect sarcasm. Thus, an ensemble approach was proposed to address the issue. Chapter 3 introduces the design and implementation of an ensemble approach to detect sarcasm. We have performed research on social networking data. The data is pre-processed using an approach of tokenization; the features are extracted using the random forest algorithm, the Principal Component Analysis algorithm is applied for the feature reduction, K-mean is used for the data clustering, and in the phase of classification, four different ensemble classifiers are designed which are the combination of multiple classifiers. The first ensemble classifier combines SVM, KNN, and decision tree. The second ensemble classifier uses SVM, logistic regression, and decision tree classifiers to identify sarcasm. MLP, logistic regression, and decision tree are combined in the MLP ensemble model, whereas MLP, logistic regression, and SVM are combined in the MLP, logistic regression, and SVM ensemble model. The performance of each ensemble model is tested on the dataset, which is a combined dataset of tweets and news headlines. Results are presented in chapter 4 where the performance of the ensemble models is measured in terms of accuracy, precision, recall, and F1 score. It is analyzed that the proposed ensemble classifier had performed well in comparison with other algorithms concerning various metrics for Sarcasm detection. The accuracy of the proposed approach improved by 4.31 percent, 3.76 percent, 76.15 percent and 6.38 percent, the precision of the proposed approach improved by 3.37 percent, 2.22 percent, 29.57 percent and 4.54 percent, the recall of the proposed approach improved by 3.33 percent, 3.33 percent, 75.47 percent and 5.68 percent, F1-score of proposed approach improved by 4.49 percent, 3.33 percent, 132 percent and 5.68 percent respectively, when compared to Adaboost, Decision Tree, Random Forest and K-Nearest Neighbor classifier. Also, in comparison with existing ensemble approaches, the proposed method has achieved the highest accuracy among others.

5.2. Future Work

Sarcasm detection is one of the main challenges of sentiment analysis. The importance of sarcasm has notably increased in the past few years. The proposed solution has been designed and tested only for textual data for detecting sarcasm. In future, this can be extended to detect sarcasm by analyzing neutral tweets would also be a potential priority since certain tweets have neither optimistic nor negative opinions.

As social media is rising day by day, there are plenty of text, graphics, audio clippings, memes and different means of communication. A consistent job is done in the text to recognize the sarcasm, but very little progress has been made in multimedia sentiment analysis. In the future, we can extend our work in this field as it is a growing area, and data is rising day by day and should be used for the detection of sarcasm.

Overall, sentiment analysis has found various promising applications like market prediction, political sentiment determination, equity value prediction, box office prediction etc. But a lot of work remains to be done, and it is a fertile area for sarcasm analysis on financial markets and government issues.

REFERENCES

[1] M. Karqmibekr and A.A. ghorbani, “Sentiment Analysis of a Social Issues,” International Conference on a Social Informatics, USA, pp. 215-221, 2012.

[2] M. Eirinaki, S. Pisal, J. Singh, “Feature-based opinion mining and ranking,” Journal of Computer and System Sciences, vol. 78, issue 4, pp. 1175- 1184, 2012.

[3] E. Haddia, X. Liua and Y. Shib, “The Role of Text Pre-processing in Sentiment Analysis,” Information Technology and Quantitative Management, vol. 17, pp. 26-32, 2013.

[4] K. Ghag and K. Shah, “Comparative Analysis of the Techniques for Sentiment Analysis,” International Conference on Advances in Technology and Engineering (ICATE), pp. 1-7, 2013

[5] Q. Vuong and A. Takasu, “Transfer Learning for Emotional Polarity Classification,” International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), vol. 2, pp. 94-101, 2014.

[6] V. Rout and A.D. Londhe, “Survey on Opinion Mining and Summarization of User Reviews on Web,” International Journal of Computer Science and Information Technologies, vol. 5, pp. 1026-1030, 2014.

[7] Shahana P.H and Bini Omman, “Evaluation of Features on Sentimental Analysis,” International Conference on Information and Communication Technologies, pp. 1585-1592, 2015.

[8] A. Tripathy, A. Agrawal and S. K. Rath, “Classification of Sentimental Reviews Using Machine Learning Techniques,” International Conference on Recent Trends in Computing, pp. 821-829, 2015.

[9] U. Farooq, T. Prasad, A. Nongaillard, Y. Ouzrout and M. Qadir, “A Word Sense Disambiguation Method for Feature Level Sentiment Analysis,” International

Conference on Software, Knowledge, Information Management and Applications (SKIMA), pp. 1-8, 2015.

10] Jayraj M. Desai, Swapnil R. Andhariya, “Sentiment analysis approach to adapt a shallow parsing-based sentiment lexicon”, 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

[11] M. Shanmugam, A. Agawane, A. Tiwari and R. V. Deolekar, "Twitter Sentiment Analysis using Novelty Detection," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 1258-1263

[12] M. Karthiga, S. Kumar G., N. Aravindhraj and S. Priyanka, "Machine Learning-Based Sentiment Analysis of Twitter Data," 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2019, pp. 1-8

[13] T. Parlar, E. Saraç and S. A. Özel, "Comparison of feature selection methods for sentiment analysis on Turkish Twitter data," 2017 25th Signal Processing and Communications Applications Conference (SIU), 2017, pp. 1-4

[14] J. Ranganathan, A. S. Irudayaraj and A. A. Tzacheva, "Action Rules for Sentiment Analysis on Twitter Data Using Spark," 2017 IEEE International Conference on Data Mining Workshops (ICDMW), 2017, pp. 51-60

[15] M. Venugopalan and D. Gupta, "Exploring sentiment analysis on twitter data," 2015 Eighth International Conference on Contemporary Computing (IC3), 2015, pp. 241-247

[16] S. Ranjan, S. Sood and V. Verma, "Twitter Sentiment Analysis of Real-Time Customer Experience Feedback for Predicting Growth of Indian Telecom Companies," 2018 4th International Conference on Computing Sciences (ICCS), 2018, pp. 166-174

[17] P. Monish, S. Kumari and C. Narendra Babu, "Automated Topic Modeling and Sentiment Analysis of Tweets on SparkR," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1-7

[18] J. Wehrmann, W. Becker, H. E. L. Cagnini and R. C. Barros, "A character-based convolutional neural network for language-agnostic Twitter sentiment analysis," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2384-2391

[19] A. Krouska, C. Troussas and M. Virvou, "The effect of pre-processing techniques on Twitter sentiment analysis," 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA), 2016, pp. 1-5

[20] M. Bouazizi and T. Ohtsuki, "A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter," in IEEE Access, vol. 5, pp. 20617-20639, 2017

[21] Bala Durga Dharmavarapu, Jayanag Bayana, “Sarcasm Detection in Twitter using Sentiment Analysis”, 2019, International Journal of Recent Technology and Engineering (IJRTE)

[22] Athira M, Chithra C, Gayathry Anil and Smitha E S, “Sentiment Analysis - Sarcasm Detection in Twitter”, 2020, IOSR Journal of Computer Engineering (IOSR-JCE)

[23] Dr. Kalpesh H. Wandra1, Mehul Baro, “SARCASM DETECTION IN SENTIMENT ANALYSIS”, 2017, International Journal of Current Engineering and Scientific Research (IJCESR)

[24] Dwi A. P. Rahayu, Soveatin Kuntur, Nur Hayatin, “Sarcasm Detection on Indonesian Twitter Feeds”, 2018, Proceeding of EECSI 2018, Malang – Indonesia

[25] Sundararajan K., Saravana J.V., Palanisamy A., “Textual Feature Ensemble-Based Sarcasm Detection in Twitter Data”, 2021, In: Peter J., Fernandes S., Alavi A. (eds) Intelligence in Big Data Technologies—Beyond the Hype. Advances in Intelligent Systems and Computing, vol 1167. Springer, Singapore.

[26] E. Lunando and A. Purwarianti, "Indonesian social media sentiment analysis with sarcasm detection," 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2013, pp. 195-198

[27] M. Bouazizi and T. Otsuki Ohtsuki, "A Pattern-Based Approach for Sarcasm Detection on Twitter," in IEEE Access, vol. 4, pp. 5477-5488, 2016

[28] S. Hiai and K. Shimada, "A Sarcasm Extraction Method Based on Patterns of Evaluation Expressions," 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2016, pp. 31-36

[29] A. D. Dave and N. P. Desai, "A comprehensive study of classification techniques for sarcasm detection on textual data," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, pp. 1985-1991

[30] S. K. Bharti, K. S. Babu and S. K. Jena, "Parsing-based sarcasm sentiment recognition in Twitter data," 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2015, pp. 1373-1380

[31] E. Fersini, F. A. Pozzi and E. Messina, "Detecting irony and sarcasm in microblogs: The role of expressive signals and ensemble classifiers," 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2015, pp. 1-8

[32] T. Ahmad, H. Akhtar, A. Chopra and M. W. Akhtar, "Satire Detection from Web Documents Using Machine Learning Methods," 2014 International Conference on Soft Computing and Machine Intelligence, 2014, pp. 102-105

[33] M. Khokhlova, V. Patti and P. Rosso, "Distinguishing between irony and sarcasm in social media texts: Linguistic observations," 2016 International FRUCT Conference on Intelligence, social media and Web (ISMW FRUCT), 2016, pp. 1-6

[34] Md. Rakibul Hasan, Maisha Maliha, M. Arifuzzaman, “Sentiment Analysis with NLP on Twitter Data”, 2019, International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)

[35] Önder Çoban, Gülşah Tümüklü Özyer, "Word2vec and Clustering based Twitter Sentiment Analysis", 2018, International Conference on Artificial Intelligence and Data Processing (IDAP)

[36] Lei Wang, Jianwei Niu, Shui Yu, "SentiDiff: Combining Textual Information and Sentiment Diffusion Patterns for Twitter Sentiment Analysis", 2019, IEEE Transactions on Knowledge and Data Engineering

[37] Paramita Ray, Amlan Chakrabarti, "Twitter sentiment analysis for product review using lexicon method", 2017, International Conference on Data Management, Analytics and Innovation (ICDMAI)

[38] Victoria Ikoro, Maria Sharmina, Khaleel Malik, Riza Batista-Navarro, "Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers", 2018, Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS)

[39] Rashid Kamal, Munam Ali Shah, Asad Hanif, J Ahmad, "Real-time opinion mining of Twitter data using spring XD and Hadoop", 2017, 23rd International Conference on Automation and Computing (ICAC)

[40] Sheeba Naz, Aditi Sharan, Nidhi Malik, "Sentiment Classification on Twitter Data Using Support Vector Machine", 2018, IEEE/WIC/ACM International Conference on Web Intelligence (WI)

[41] Mansour Alshaikh, Mohamed Zohdy, "Sentiment Analysis for Smartphone Operating System: Privacy and Security on Twitter Data", 2020, IEEE International Conference on Electro Information Technology (EIT)

[42] Peiman Barnaghi, Parsa Ghaffari, John G. Breslin, "Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment", 2016, IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)

[43] Vallikannu Ramanathan, T. Meyyappan, “Twitter Text Mining for Sentiment Analysis on People’s Feedback about Oman Tourism”, 2019, 4th MEC International Conference on Big Data and Smart City (ICBDSC)

[44] Hussain AlSalman, “An Improved Approach for Sentiment Analysis of Arabic Tweets in Twitter Social Media”, 2020, 3rd International Conference on Computer Applications & Information Security (ICCAIS)

[45] Cem Baydogan, Bilal Alatas, “Sentiment analysis using Konstanz Information Miner in social networks”, 2018, 6th International Symposium on Digital Forensic and Security (ISDFS)

[46] Nann Hwan Khun, Hninn Aye Thant, “Visualization of Twitter Sentiment during the Period of US Banned Huawei”, 2019, International Conference on Advanced Information Technologies (ICAIT)

[47] Ankita Rane, Anand Kumar, “Sentiment Classification System of Twitter Data for US Airline Service Analysis”, 2018, IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)

[48] Abdullah Ammar Karcioğlu, Tolga Aydin, “Sentiment Analysis of Turkish and English Twitter Feeds Using Word2Vec Model”, 2019, 27th Signal Processing and Communications Applications Conference (SIU)

[49] Mudassir Khan, Aadarsh Malviya, “Big data approach for sentiment analysis of twitter data using Hadoop framework and deep learning”, 2020, International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)

[50] Xinyi Guo, Jinfeng Li, “A Novel Twitter Sentiment Analysis Model with Baseline Correlation for Financial Market Prediction with Improved Efficiency”, 2019, Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)

[51] Abdelmounaam Rezgui, Daniel Fahey, Ian Smith, “AffinityFinder: A System for Deriving Hidden Affinity Relationships on Twitter Utilizing Sentiment Analysis”, 2016, IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)

[52] Vishu Tyagi, Ashwini Kumar, Sanjoy Das, “Sentiment Analysis on Twitter Data Using Deep Learning approach”, 2020, 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)

[53] Hossam Elzayady, Khaled M. Badran, Gouda I. Salama, “Sentiment Analysis on Twitter Data using Apache Spark Framework”, 2018, 13th International Conference on Computer Engineering and Systems (ICCES)

[54] Mamta Patil, H.K. Chavan, “Event Based Sentiment Analysis of Twitter Data”, 2018, Second International Conference on Computing Methodologies and Communication (ICCMC)

[55] Deni Kurnianto Nugroho, “US presidential election 2020 prediction based on Twitter data using lexicon-based sentiment analysis”, 2021, 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)

[56] Shihab Elbagir Saad, Jing Yang, “Twitter Sentiment Analysis Based on Ordinal Regression”, 2019, IEEE Access

[57] Mariam Khader, Arafat Awajan, Ghazi Al-Naymat, “The Effects of Natural Language Processing on Big Data Analysis: Sentiment Analysis Case Study”, 2018, International Arab Conference on Information Technology (ACIT)

[58] A Jenifer Jothi Mary, L Arockiam, “Jen-Ton: A framework to enhance the accuracy of aspect level sentiment analysis in big data”, 2017, International Conference on Inventive Computing and Informatics (ICICI)

[59] Sumandeep Kaur, Geeta Sikka, Lalit Kumar Awasthi, “Sentiment Analysis Approach Based on N-gram and KNN Classifier”, 2018, First International Conference on Secure Cyber Computing and Communication (ICSCCC)

[60] Rincy Jose, Varghese S Chooralil, “Prediction of election result by enhanced sentiment analysis on twitter data using classifier ensemble Approach”, 2016, International Conference on Data Mining and Advanced Computing (SAPIENCE)

[61] MSR Hitesh, Vedhosi Vaibhav, Y.J Abhishek Kalki, Suraj Harsha Kamtam, Santoshi Kumari, “Real-Time Sentiment Analysis of 2019 Election Tweets using Word2vec and Random Forest Model”, 2019, 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)

[62] Brandon Joyce, Jing Deng, “Sentiment analysis of tweets for the 2016 US presidential election”, 2017, IEEE MIT Undergraduate Research Technology Conference (URTC)

[63] Rania Othman, Youcef Abdelsadek, Kamel Chelghoum, Imed Kacem, Rim Faiz, “Improving Sentiment Analysis in Twitter Using Sentiment Specific Word Embeddings”, 2019, 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)

[64] Gede Rizky Gustisa Wisnu, Ahmadi, Ahmad Rizaqu Muttaqi, Aris Budi Santoso, Prabu Kresna Putra, Indra Budi, “Sentiment Analysis and Topic Modelling of 2018 Central Java Gubernatorial Election using Twitter Data”, 2020, International Workshop on Big Data and Information Security (IWBIS)

[65] Zhao Jianqiang, Gui Xiaolin, Zhang Xuejun, “Deep Convolution Neural Networks for Twitter Sentiment Analysis”, 2018, IEEE Access

[66] Jajam Haerul Jaman, Rasdi Abdulrohman, “Sentiment Analysis of Customers on Utilizing Online Motorcycle Taxi Service at Twitter with the Support Vector Machine”,

2019, International Conference on Electrical Engineering and Computer Science (ICECOS)

[67] Malika Acharya, Shilpi Sharma, “Semantic Analysis of Twitter Posts”, 2018, International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)

[68] Andry Alamsyah, Wirawan Rizkika, Ditya Dwi Adhi Nugroho, Farhan Renaldi, Siti Saadah, “Dynamic Large Scale Data on Twitter Using Sentiment Analysis and Topic Modeling”, 2018, 6th International Conference on Information and Communication Technology (ICoICT)

[69] Usman Naseem, Katarzyna Musial, “DICE: Deep Intelligent Contextual Embedding for Twitter Sentiment Analysis”, 2019, International Conference on Document Analysis and Recognition (ICDAR)

[70] Prema Monish, Santoshi Kumari, C Narendra Babu, “Automated Topic Modeling and Sentiment Analysis of Tweets on SparkR”, 2018, 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)

[71] Joonatas Wehrmann, Willian Becker, Henry E. L. Cagnini, Rodrigo C. Barros, “A character-based convolutional neural network for language-agnostic Twitter sentiment analysis”, 2017, International Joint Conference on Neural Networks (IJCNN)

[72] Mamoon Rashid, Aamir Hamid, Nazir Ahmad, Mohammed Burhanur Rehman, Mir Mohammad Yousuf, “Novel Machine Learning Approach for Sentiment Analysis of Real Time Twitter Data with Apache Flume”, 2020, Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)

[73] Andia Enggar Mayasari, Anggit Dwi Hartanto, “User Satisfaction Levels Sentiment Analysis Toward Goods Delivery Service On Twitter Using Support Vector Machine Algorithm (SVM)”, 2019, 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)

[74] Shiv Dhar, Suyog Pednekar, Kishan Borad, Ashwini Save, “Sentiment Analysis Using Neural Networks: A New Approach”, 2018, Second International Conference on Inventive Communication and Computational Technologies (ICICCT)

[75] Li-Chen Cheng, Song-Lin Tsai, “Deep Learning for Automated Sentiment Analysis of Social Media”, 2019, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

[76] Upma Kumari, Arvind K Sharma, Dinesh Soni, “Sentiment analysis of smart phone product review using SVM classification technique”, 2017, International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)

[77] W.P. Ramadhan, S.T.M.T. Astri Novianty, S.T.M.T. Casi Setianingsih, “Sentiment analysis using multinomial logistic regression”, 2017, International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)

[78] K. Padmaja, Nagaratna P. Hegde, “Twitter sentiment analysis using adaptive neuro-fuzzy inference system with genetic algorithm”, 2019, 3rd International Conference on Computing Methodologies and Communication (ICCMC)

[79] P Kalaivani, D Dinesh, “Machine Learning Approach to Analyze Classification Result for Twitter Sentiment”, 2020, International Conference on Smart Electronics and Communication (ICOSEC)

[80] Della Fitrayani Budiono, Anto Satriyo Nugroho, Afrizal Doewes, “Twitter sentiment analysis of DKI Jakarta's gubernatorial election 2017 with predictive and descriptive approaches”, 2017, International Conference on Computer, Control, Informatics and its Applications (IC3INA)

[81] James Ho, Dominik Ondusko, Brandon Roy, D. Frank Hsu, “Sentiment Analysis on Tweets Using Machine Learning and Combinatorial Fusion”, 2019, IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence

and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)

[82] Mehmet Umut Salur, İlhan Aydin, Shahed Alali Alghrsi, "SmartSenti: A Twitter-Based Sentiment Analysis System for the Smart Tourism in Turkey", 2019, International Artificial Intelligence and Data Processing Symposium (IDAP)

[83] R. Monika, S. Deivalakshmi, B. Janet, "Sentiment Analysis of US Airlines Tweets Using LSTM/RNN", 2019, IEEE 9th International Conference on Advanced Computing (IACC)

[84] Raheela Bibi, Usman Qamar, Munazza Ansar, Asma Shaheen, "Sentiment Analysis for Urdu News Tweets Using Decision Tree", 2019, IEEE 17th International Conference on Software Engineering Research, Management and Applications (SERA)

[85] Rai B Shamantha, Sweekriti M Shetty, Prakhyath Rai, "Sentiment Analysis Using Machine Learning Classifiers: Evaluation of Performance", 2019, IEEE 4th International Conference on Computer and Communication Systems (ICCCS)

[86] Sahar A. El Rahman, Feddah Alhumaidi AlOtaibi, Wejdan Abdullah AlShehri, "Sentiment Analysis of Twitter Data", 2019, International Conference on Computer and Information Sciences (ICCIS)

[87] Huma Parveen, Shikha Pandey, "Sentiment analysis on Twitter Data-set using Naive Bayes algorithm", 2016, 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)

[88] Li Yang, Ying Li, Jin Wang, R. Simon Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning", 2020, IEEE Access

[89] Fanglong Yao, Xian Sun, Hongfeng Yu, Wenkai Zhang, Wei Liang, Kun Fu, "Mimicking the Brain's Cognition of Sarcasm From Multidisciplines for Twitter

Sarcasm Detection", 2021, IEEE Transactions on Neural Networks and Learning Systems

[90] Dwi A. P. Rahayu, Soveatin Kuntur, Nur Hayatin, "Sarcasm Detection on Indonesian Twitter Feeds", 2018, 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)

[91] Saurabh Porwal, Gaurav Ostwal, Anagha Phadtare, Mohini Pandey, Manisha V. Marathe, "Sarcasm Detection Using Recurrent Neural Network", 2018, Second International Conference on Intelligent Computing and Control Systems (ICICCS)

[92] Shubham Rendalkar, Chaitali Chandankhede, "Sarcasm Detection of Online Comments Using Emotion Detection", 2018, International Conference on Inventive Research in Computing Applications (ICIRCA)

[93] Dana Al-Ghadhban, Eman Alnkhilan, Lamma Tatwany, Muna Alrazgan, "Arabic sarcasm detection in Twitter", 2017, International Conference on Engineering & MIS (ICEMIS)

[94] Muhammad Abulaish, Ashraf Kamal, "Self-Deprecating Sarcasm Detection: An Amalgamation of Rule-Based and Machine Learning Approach", 2018, IEEE/WIC/ACM International Conference on Web Intelligence (WI)

[95] Suyash Sangwan, Md Shad Akhtar, Pranati Behera, Asif Ekbal, "I didn't mean what I wrote! Exploring Multimodality for Sarcasm Detection", 2020, International Joint Conference on Neural Networks (IJCNN)

[96] K Sreelakshmi, P C Rafeeque, "An Effective Approach for Detection of Sarcasm in Tweets", 2018, International CET Conference on Control, Communication, and Computing (IC4)

[97] Rahul Gupta, Jitendra Kumar, Harsh Agrawal, Kunal, "A Statistical Approach for Sarcasm Detection Using Twitter Data", 2020, 4th International Conference on Intelligent Computing and Control Systems (ICICCS)

[98] Kumar, H. M., Harish, B. S., & Darshan, H. K. (2019). Sentiment Analysis on IMDb Movie Reviews Using Hybrid Feature Extraction Method. *International Journal of Interactive Multimedia & Artificial Intelligence*, *5*(5).

[99] Yufeng Diao, Hongfei Lin, Liang Yang, Xiaochao Fan, Yonghe Chu, Kan Xu, Di Wu, "A Multi-Dimension Question Answering Network for Sarcasm Detection", 2020, IEEE Access

[100]Christopher Ifeanyi Eke, Azah Norman, Liyana Shuib, Faith B. Fatokun, Isaiah Omame, "The Significance of Global Vectors Representation in Sarcasm Analysis", 2020, International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS)

[101] Satoshi Hiai, Kazutaka Shimada, "Sarcasm Detection Using Role Pairs: Bootstrapping Role Pairs and Weighting", 2020, International Conference on Asian Language Processing (IALP)

[102] Mohd Suhairi Md Suhaimin, Mohd Hanafi Ahmad Hijazi, Rayner Alfred, Frans Coenen, "Natural language processing based features for sarcasm detection: An investigation using bilingual social media texts", 2017, 8th International Conference on Information Technology (ICIT)

[103] Yang Wu, Yanyan Zhao, Xin Lu, Bing Qin, Yin Wu, Jian Sheng, Jinlong Li, "Modeling Incongruity between Modalities for Multimodal Sarcasm Detection", 2021, IEEE MultiMedia

[104] Jiangnan Li, Hongliang Pan, Zheng Lin, Peng Fu, Weiping Wang, "Sarcasm Detection with Commonsense Knowledge", 2021, IEEE/ACM Transactions on Audio, Speech, and Language Processing

[105] Yazhou Zhang, Yaochen Liu, Qiuchi Li, Prayag Tiwari, Benyou Wang, Yuhua Li, Hari Mohan Pandey, Peng Zhang, Dawei Song, "CFN: A Complex-valued Fuzzy Network for Sarcasm Detection in Conversations", 2021, IEEE Transactions on Fuzzy Systems

[106] Manoj Y. Manohar, Pallavi Kulkarni, “Improvement sarcasm analysis using NLP and corpus based approach”, 2017, International Conference on Intelligent Computing and Control Systems (ICICCS)

[107] Amina Ben Meriem, Lobna Hlaoua, Lotfi Ben Romdhane, “A fuzzy approach for sarcasm detection in social networks”, 2021, Procedia Computer Science

[108] Lu Ren, Bo Xu, Liang Yang, “Sarcasm Detection with Sentiment Semantics Enhanced Multi-level Memory Network”, 2020, Neurocomputing

[109] Mayank Shrivastava, Shishir Kumar, “A pragmatic and intelligent model for sarcasm detection in social media text”, 2020, Technology in Society

[110] Satoshi Hiai, Kazutaka Shimada, “A Sarcasm Extraction Method Based on Patterns of Evaluation Expressions”, 2016, 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)

[111] Pawar, N., & Bhingarkar, S. (2020, June). Machine Learning based Sarcasm Detection on Twitter Data. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)* (pp. 957-961). IEEE.

[112] Pradeesh Parameswaran, Andrew Trotman, David Eyers, “Detecting the target of sarcasm is hard: Really??”, 2021, Information Processing & Management

[113] Parnavi Shrikhande, Vikram Setty, Dr. Ashish Sahani, “Sarcasm Detection in Newspaper Headlines”, 2020, IEEE 15th International Conference on Industrial and Information Systems (ICIIS)

[114] Bouazizi, M., & Ohtsuki, T. (2018). Multi-class sentiment analysis in Twitter: What if classification is not the answer. IEEE Access, 6, 64486-64502.

[115] Taj, S., Shaikh, B. B., & Meghji, A. F. (2019, January). Sentiment Analysis of News Articles: A Lexicon based Approach. In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (pp. 1-5). IEEE.

[116] Ren, R., Wu, D. D., & Liu, T. (2019). Forecasting stock market movement direction using sentiment analysis and support vector machine. IEEE Systems Journal, 13(1), 760-770.

[117] Xu, G., Yu, Z., Yao, H., Li, F., Meng, Y., & Wu, X. (2019). Chinese Text Sentiment Analysis Based on Extended Sentiment Dictionary. IEEE Access, 7, 43749-43762.

[118] Iqbal, F., Hashmi, J. M., Fung, B. C., Batool, R., Khattak, A. M., Aleem, S., & Hung, P. C. (2019). A Hybrid Framework for Sentiment Analysis Using Genetic Algorithm Based Feature Reduction. IEEE Access, 7, 14637-14652.

[119] Jianqiang, Z., Xiaolin, G., & Xuejun, Z. (2018). Deep convolution neural networks for twitter sentiment analysis. IEEE Access, 6, 23253-23260.

[120] Fu, X., Yang, J., Li, J., Fang, M., & Wang, H. (2018). Lexicon-enhanced LSTM with attention for general sentiment analysis. IEEE Access, 6, 71884-71891.

[121] Fang, Y., Tan, H., & Zhang, J. (2018). Multi-strategy sentiment analysis of consumer reviews based on semantic fuzziness. IEEE Access, 6, 20625-20631.

[122] Agarwal, A., & Toshniwal, D. (2018, June). Application of lexicon based approach in sentiment analysis for short Tweets. In 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE) (pp. 189-193). IEEE.

[123] Dutta, S., Mehta, A., & Bandyopadhyay, S. K. A (May 2020). Novel Integrated Framework for Sarcasm Detection in Social Platform. International Journal of Engineering and Advanced Technology(IJEAT)

[124] Lemmens, J., Burtenshaw, B., Lotfi, E., Markov, I., & Daelemans, W. (2020, July). Sarcasm detection using an ensemble approach. In *Proceedings of the Second Workshop on Figurative Language Processing* (pp. 264-269).

[125] Ilavarasan, E. (2020, March). A Survey on Sarcasm detection and challenges. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 1234-1240). IEEE.

[126] Pandey, A. C., Seth, S. R., & Varshney, M. (2019). Sarcasm detection of amazon alexa sample set. In Advances in Signal Processing and Communication (pp. 559-564). Springer, Singapore.

[127] John, A., John, A., & Sheik, R. (2019, April). Context Deployed Sentiment Analysis Using Hybrid Lexicon. In 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT) (pp. 1-5). IEEE.

[128] Khan, J., Alam, A., Hussain, J., & Lee, Y. K. (2019). EnSWF: effective features extraction and selection in conjunction with ensemble learning methods for document sentiment classification. Applied Intelligence, 1-23.

[129] Basak, R., Sural, S., Ganguly, N., & Ghosh, S. K. (2019). Online Public Shaming on Twitter: Detection, Analysis, and Mitigation. IEEE Transactions on Computational Social Systems, 6(2), 208-220.

[130] Luintel, S., Sah, R. K., & Lamichhane, B. R. (2019). A Hybrid Approach for Sarcasm Detection. *Technical Journal*, *1*(1), 1-9.

[131] Al-Twairesh, N., & Al-Negheimish, H. (2019). Surface and Deep Features Ensemble for Sentiment Analysis of Arabic Tweets. IEEE Access, 7, 84122-84131.

[132] Hiai, S., & Shimada, K. (2018, February). Sarcasm detection using features based on indicator and roles. In International Conference on Soft Computing and Data Mining (pp. 418-428). Springer, Cham.

[133] Das, D., & Clark, A. J. (2018, October). Sarcasm detection on facebook: A supervised learning approach. In Proceedings of the 20th International Conference on Multimodal Interaction: Adjunct (pp. 1-5).

[134] Alshari, E. M., Azman, A., Doraisamy, S., Mustapha, N., & Alkeshr, M. (2018, March). Effective method for sentiment lexical dictionary enrichment based on

Word2Vec for sentiment analysis. In 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP) (pp. 1-5). IEEE.

[135] Abulaish, M., & Kamal, A. (2018, December). Self-Deprecating Sarcasm Detection: An Amalgamation of Rule-Based and Machine Learning Approach. In 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI) (pp. 574-579). IEEE.

[136] Rendalkar, S., & Chandankhede, C. (2018, July). Sarcasm Detection of Online Comments Using Emotion Detection. In 2018 International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 1244-1249). IEEE.

[137] Ahuja, R., Bansal, S., Prakash, S., Venkataraman, K., & Banga, A. (2018). Comparative Study of Different Sarcasm Detection Algorithms Based On Behavioral Approach. Procedia computer science, 143, 411-418.

[138] Ren, Y., Ji, D., & Ren, H. (2018). Context-augmented convolutional neural networks for twitter sarcasm detection. Neurocomputing, 308, 1-7.

[139] Zhao, W., Guan, Z., Chen, L., He, X., Cai, D., Wang, B., & Wang, Q. (2017). Weakly-supervised deep embedding for product review sentiment analysis. IEEE Transactions on Knowledge and Data Engineering, 30(1), 185-197.

[140] Gidhe, P., & Ragha, L. (2017, December). Sarcasm detection of non# tagged statements using MLP-BP. In 2017 International Conference on Advances in Computing, Communication and Control (ICAC3) (pp. 1-4). IEEE.

[141] Razali, M. S., Halin, A. A., Norowi, N. M., & Doraisamy, S. C. (2017, December). The importance of multimodality in sarcasm detection for sentiment analysis. In 2017 IEEE 15th Student Conference on Research and Development (SCOReD) (pp. 56-60). IEEE.

[142] Mukherjee, S., & Bala, P. K. (2017). Sarcasm detection in microblogs using Naïve Bayes and fuzzy clustering. Technology in Society, 48, 19-27.

[143] Deliens, G., Antoniou, K., Clin, E., & Kissine, M. (2017). Perspective-taking and frugal strategies: Evidence from sarcasm detection. Journal of pragmatics, 119, 33-45.

[144] Ludlow, A. K., Chadwick, E., Morey, A., Edwards, R., & Gutierrez, R. (2017). An exploration of sarcasm detection in children with attention deficit hyperactivity disorder. Journal of communication disorders, 70, 25-34.

[145] Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, *57*, 117-126.

[146] Majumder, N., Poria, S., Peng, H., Chhaya, N., Cambria, E., & Gelbukh, A. (2019). Sentiment and sarcasm classification with multitask learning. *IEEE Intelligent Systems*, *34*(3), 38-43.

[147] Da Silva, N.F., Hruschka, E.R., & Hruschka Jr, E.R. (2014). Tweet sentiment analysis with classifier ensembles. Decision Support Systems, 66, 170-179.

[148] Das, D., & Clark, A.J. (2018). Sarcasm detection on facebook: A supervised learning approach. In Proceedings of the 20th International Conference on Multimodal Interaction: Adjunct, (pp. 1-5).

[149] Jamil, R., Ashraf, I., Rustam, F., Saad, E., Mehmood, A., & Choi, G. S. (2021). Detecting sarcasm in multi-domain datasets using convolutional neural networks and long short term memory network model. *PeerJ Computer Science*, *7*, e645.

[150] Vinodhini, G., & Chandrasekaran, R.M. (2013). Effect of feature reduction in sentiment analysis of online reviews. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 2(6), 2165-2172.

[151] Dash, B., Mishra, D., Rath, A., & Acharya, M. (2010). A hybridized K-means clustering approach for high dimensional dataset. International Journal of Engineering, Science and Technology, 2(2), 59-66.

[152] Malli, S., Nagesh, H.R., & Rao, B.D. Approximation to the K-Means Clustering Algorithm using PCA. International Journal of Computer Applications, 975, 8887.

[153] Jotheeswaran, J., & Koteeswaran, S. (2016). Feature selection using random forest method for sentiment analysis. Indian Journal of Science and Technology, 9(3), 1-7.

[154] Bharti, S.K., Vachha, B., Pradhan, R.K., Babu, K.S., & Jena, S.K. (2016). Sarcastic sentiment detection in tweets streamed in real time: a big data approach. Digital Communications and Networks, 2(3), 108-121.

[155] Saleena, N. (2018). An ensemble classification system for twitter sentiment analysis. Procedia computer science, 132, 937-946.

[156] Vinodhini, G., & Chandrasekaran, R.M. (2014). Sentiment classification using principal component analysis based neural network model. In International Conference on Information Communication and Embedded Systems (ICICES2014) (pp. 1-6). IEEE.

[157] Garg, A., & Duhan, N. (2020). Sarcasm Detection on Twitter Data Using Support Vector Machine. ICTACT Journal of soft computing, 10(4), 2165-2170.

[158] Ashwitha, A., Shruthi, G., Shruthi, H.R., Upadhyaya, M., Ray, A.P., & Manjunath, T.C. (2021). Sarcasm detection in natural language processing. Materials Today: Proceedings, 37, 3324-3331.

[159] Suhaimin, M.S.M., Hijazi, M.H.A., Alfred, R., & Coenen, F. (2017). Natural language processing based features for sarcasm detection: An investigation using bilingual social media texts. In 2017 8th International conference on information technology (ICIT) (pp. 703-709). IEEE.

[160] Ilavarasan, E. (2020, March). A Survey on Sarcasm detection and challenges. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 1234-1240). IEEE.

[161] Kumar A, Sangwan SR, Arora A, Nayyar A, Abdel-Basset M, others . Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network. IEEE Access 2019; 7: 23319–23328.

[162] John, N. (2020). Tweets with sarcasm and irony. *dirección: https://www. kaggle. com/nikhiljohnk/tweets-with-sarcasm-and-irony*.

[163] Jurek, A., Bi, Y., Wu, S., & Nugent, C. (2014). A survey of commonly used ensemble-based classification techniques. The Knowledge Engineering Review, 29(5), 551.

[164] Misra R. 2019. News headlines dataset for sarcasm detection. Available at https://www.kaggle.com/ rmisra/news-headlines-dataset-for-sarcasm-detection.

[165] Wang, G., Sun, J., Ma, J., Xu, K., & Gu, J. (2014). Sentiment classification: The contribution of ensemble learning. Decision support systems, 57, 77-93.

[166] Perikos, I., & Hatzilygeroudis, I. (2016). Recognizing emotions in text using ensemble of classifiers. Engineering Applications of Artificial Intelligence, 51, 191-201.

[167] Lemmens, J., Burtenshaw, B., Lotfi, E., Markov, I., & Daelemans, W. (2020, July). Sarcasm detection using an ensemble approach. In Proceedings of the Second Workshop on Figurative Language Processing (pp. 264-269).

[168] Sundararajan, K., & Palanisamy, A. (2020). Multi-Rule Based Ensemble Feature Selection Model for Sarcasm Type Detection in Twitter. Computational Intelligence and Neuroscience, 2020.

[169] Luintel, S., Sah, R. K., & Lamichhane, B. R. (2019). A Hybrid Approach for Sarcasm Detection. Technical Journal, 1(1), 1-9.

[170] Farha, I. A., & Magdy, W. (2021). A comparative study of effective approaches for arabic sentiment analysis. *Information Processing & Management*, *58*(2), 102438.

[171] Wang G, Sun J, Ma J, Xu K, Gu J. Sentiment classification: The contribution of ensemble learning. Decision support systems 2014; 57: 77– 93.

[172] Ravi K, Ravi V. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems 2015; 89: 14–46.

[173] Kumar A, Jaiswal A. Systematic literature review of sentiment analysis on Twitter using soft computing techniques. Concurrency and Computation: Practice and Experience 2020; 32(1): e5107.

[174] Kharde V, Sonawane P, others . Sentiment analysis of twitter data: a survey of techniques. arXiv preprint arXiv:1601.06971 2016.

[175] Li YM, Li TY. Deriving market intelligence from microblogs. Decision Support Systems 2013;

[176] Khan J, Alam A, Hussain J, Lee YK. EnSWF: effective features extraction and selection in conjunction with ensemble learning methods for document sentiment classification. Applied Intelligence 2019; 49(8): 3123–3145.

[177] Zhao W, Guan Z, Chen L, et al. Weakly- supervised deep embedding for product review sentiment analysis. IEEE Transactions on Knowledge and Data Engineering 2017; 30(1): 185–197.

[178] Shayaa, S., Jaafar, N. I., Bahri, S., Sulaiman, A., Wai, P. S., Chung, Y. W., ... & Al-Garadi, M. A. (2018). Sentiment analysis of big data: Methods, applications, and open challenges. *IEEE Access*, *6*, 37807-37827.

[179] Nandal, N., Tanwar, R., & Pruthi, J. (2020). Machine learning based aspect level sentiment analysis for Amazon products. *Spatial Information Research*, 1-7.

[180] Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*, 101003.

[181] Prasad AG, Sanjana S, Bhat SM, Harish B. Sentiment analysis for sarcasm detection on streaming short text data. In: IEEE.; 2017: 1–5.

Publication Details

1. Godara, J., & Aron, R. (2021). An Approach to Detect Sarcasm in Tweets. In *Data Analytics and Management* (pp. 41-50). Springer, Singapore.
   [Scopus Indexed]

2. Godara, J., & Aron, R. (2021, March). Support Vector Machine Classifier with Principal Component Analysis and K Mean for Sarcasm Detection. In *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Vol. 1, pp. 571-576). IEEE.
   [Scopus Indexed]

3. Godara, J., Aron, R., & Shabaz, M. (2021). Sentiment analysis and sarcasm detection from social network to train health-care professionals. *World Journal of Engineering*. **[Scopus and ESCI Indexed]**

4. Godara, J., & Aron, R. (2021). Sarcasm Detection on Social Network: A Review. *Annals of the Romanian Society for Cell Biology*, *25*(6), 3761-3771. **[Scopus Indexed]**

5. Godara, J., Batra, I., Aron, R., & Shabaz, M. (2021). Ensemble Classification Approach for Sarcasm Detection. *Behavioural Neurology*, *2021*. **[SCIE Indexed]**