Detecting tweet sentiment and sarcasm on online learning during covid-19

by

Shishir Kumar Das
22241123
Khairun Nisa
19101376
Razit Kabir
22241180
Israt Jahan Tonni
22241189

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
School of Data and Sciences
Brac University
January 2023

Declaration

It is hereby declared that

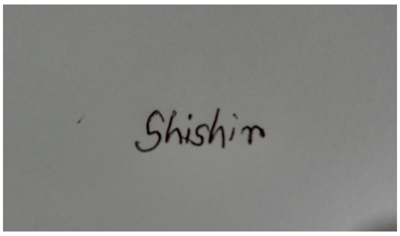
1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

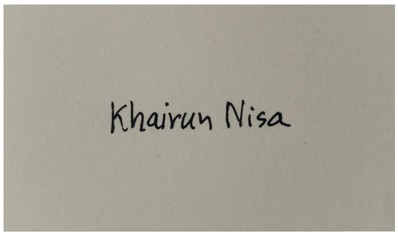
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

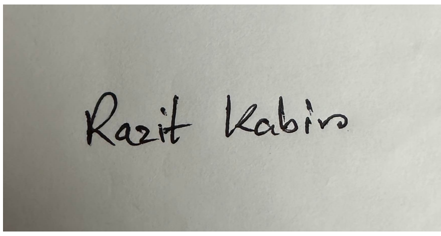
Student's Full Name & Signature:



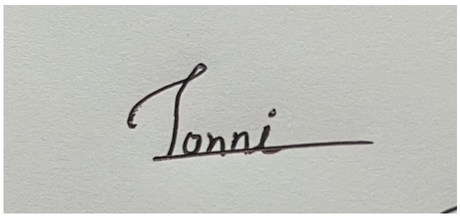
Shishir Kumar Das
22241123



Khairun Nisa
19101376



Razit Kabir
22241180



Israt Jahan Tonni
22241189

Approval

The thesis/project titled "Detecting tweet sentiment and sarcasm on online learning during covid-19" submitted by

1. Shishir Kumar Das (22241123)

2. Khairun Nisa (19101376)

3. Razit Kabir (22241180)

4. Israt Jahan Tonni (22241189)

Of Fall, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on January 17, 2023.

Examining Committee:

Supervisor:
(Member)

Dr. Amitabha Chakrabarty

Associate Professor

Department of Computer Science and Engineering

BRAC University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi

Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

COVID-19 pandemic has created a lot of challenges for student learning and education across the globe. As a result of the global increase of the state of COVID-19, numerous educational institutions in the whole world were closed in 2020 and moved to online or remote learning, which had a variety of effects on student learning. As a result teachers and students spent more time online than ever before, with both groups studying, learning, and getting acquainted themselves with information, assets, tools, and structures in order to adapt to online or virtual learning. On the basis of COVID-19, studying different opinions about online learning as the Big Data mining and analysis of tweets from the people of various countries around the world provides the opportunity to identify, quantify, and investigate the needs, challenges, and interests related to online learning in various countries around the world. Analyzing the sentiment of people they want to express through their tweets gives us a clear view of their opinion about online learning. Moreover, a huge number of tweets are sarcastic and it will not be possible to crack the sentiment of a maximum number of people without identifying the sarcastic tweets. Different types of methods were used for these analyses. Twitter is the most popular and used social media platform around the globe for many years. So, tweet data in the form of search interests related to online learning was mined for the creation of this dataset using Rapid Miner and Twitter API. As the dataset is created based on only the tweets during the covid19 the data is much less to get a more perfect result.We have analyzed the sentiment using the stemmed feature and applied a few models among which we get the best result from the logistic regression model which is 70.63% and for the sarcasm detection, we used 3 features and overall get the best accuracy 76.19% from tf-idf, 76.94% from the stemmed feature. The more information the datasets can have the more identical the changes will be. So, work on the datasets should also be continued.

Keywords: Data Mining; Machine Learning; Tweet; Sarcasm; Prediction; Decision tree; Linear Regression Analysis; online learning

Acknowledgement

Firstly, all praise to the Great Almighty for whom our thesis has been completed without any major interruption.
Secondly, to our Supervisor Amitabh sir for his unwavering support and assistance during the process of our project, which we greatly appreciate.
We are also grateful to Shahriar Hossain(RA, BRAC University) sir for his valuable feedback. Finally, we may not be able to fulfill our goals if we do not have our parents' unwavering support. We are on the approach of graduating as a result of their wonderful assistance and prayers.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

emo emoji

GNB Gaussian Naive Bayes

NLP Natural Laguage Processing

$NLTK$ Natural Language Toolkit

POS Parts Of Speech

RE Regular Expression

$repl$ Replace

$sklearn$ scikit-learn

SVD Singular Value Decomposition

$TF-IDF$ Term Frequency–Inverse Document Frequency

Chapter 1

Introduction

Social media allows users to connect and communicate with people from throughout the world. At present Twitter is the most widely used social networking site. It allows users to post short messages about their activities and share links to interesting websites or other online content. Many of these messages are public, and anyone with access to the Twitter website can view them. Furthermore, some Twitter users prefer to keep their messages private and only share them with the people who follow them. Individuals can keep their communications private from the public by substituting hashtags for @names in their messages. Words or phrases that begin with the hash sign are known as hashtags. They are used to organize tweets so that users can use the hashtag search feature to identify tweets with similar topics. For example, you could search for all tweets containing the hashtag football and then view all of the related tweets by clicking on the link associated with the hashtag. Although this feature can be useful for making private conversations more private, it can sometimes be problematic if people use their tweets to make jokes or play pranks on other users.

Twitter has significantly increased its content over the past few years, making it a prime example of what is commonly referred to as big data. On average, 450 million users use Twitter each month, and at least 500 million tweets are published every day, according to the company's official website. 97% of young people who use the internet regularly browse social media sites, demonstrating the reliability of social media as a component of online life. Images emote, or videos are regularly included in tweets on information, ideas, events, and humor on Twitter. Twitter has gained popularity as a source of big data for analytics in marketing, politics, and stock forecasts since Facebook, the most popular website on the internet, closed access to its application programming interfaces (API) in 2018. This is because Facebook had access to these APIs until 2018. Users provide businesses with information about consumers, the media, products, and politics. However, given Twitter's informal language and character limit, it may be difficult to understand user attitudes. Moreover, it's hard to conduct such an exploration. The existence of sarcasm, which occurs when someone says something other than what they mean, complicates the task significantly. Sarcasm is widely used on Twitter because it can be used to criticize and mock other people, viewpoints, ideas, and so on. Sarcasm can be expressed in a number of forms, such as direct voice, direct interaction, text, etc. Fewer stars can be given in the rating system to reflect this. Applications for sarcasm detection are numerous. It is intended to inform the analyst about the user's intention 1 and the context of the statement. Capital letters, emojis, exclamation points, and other symbols are more frequently used in sarcasm. In sentiment analysis, one of the primary duties is sarcasm identification. One can obtain a sentiment regarding the topic of the text by performing sentiment analysis, which uses a number of Natural Language Processing (NLP) techniques and algorithms to assess if a particular string of text has a good or negative sentiment linked to it [18]. Because Twitter data is unstructured, examination of it might be difficult. However, the Twitter API offers a simple way to gather a lot of information from open posts[20]. Due to Twitter's character limit, this data is typically brief and frequently only partially appropriately categorized [5] [21]. Twitter's utility in data science is further cemented by the use of text mining to create corpora [22] [23]. Despite all these advantages of utilizing Twitter data, there is still one challenge that needs to be resolved. Sarcasm is a type of irony or satire that flips the sentiment of a statement or comment, frequently by associating a surface-level positive sentiment with an intended negative sentiment [24] [25]. Even while sarcasm is popular and used in social media for both humor and criticism, the reader or listener frequently misses it[13]. Because the speaker's or author's remarks are perceived as confirming the surface sentiment rather than subverting it, this mistaken sarcasm causes perplexity. Sarcasm is frequently used in social media, which brings mistakes into sentiment analysis and opinion mining [11]. The context of speech exchanges, such as tone, inflection, or laughter [26], facial expression, and shared emotions or beliefs [27], are typically used to identify sarcasm. Even then, it often goes unnoticed because the listener needs details to support or contradict the statement's apparent aim [22]. Some Twitter users attempt to compensate for the lack of tone and context by using hashtags like not, sarcasm, sarcastic, and irony [24]. Emoticons like emojis (winking face, tongue out winking face, happy face) or emoticons (:p, ;),:-), ;P) [28] provides additional context to help the reader understand sarcastic posts. Even when employing self-labeling tools like hashtags or emotes, the author of a Tweet assumes some risk when using sarcasm that readers won't get the joke and will instead take the post seriously. Because the weight of the planned intent will be much reduced by the matched opposing polarity surface intent, the reader frequently accepts that risk in exchange for decreasing the post's emotional impact [28].

Sadly, author-tagged Tweets are not standardized even if they are useful. Apart from sarcasm, individual emotes can be used for a number of purposes, and hashtags can be used incorrectly or ironically. Therefore, developing a trustworthy system for automatic sarcasm detection still presents a design and implementation difficulty [24].

1.1 Sarcasm detection Benefits and implications

Sarcasm is frequently employed on community-based networking and web blogging, where users poke fun at or criticize in a way that makes it hard for even individuals to determine whether what is said or meant. Sarcasm is frequently cited as a barrier to sentiment analysis due to its metaphorical character (Liu, 2010). Despite having a cheerful appearance, it conveys an inferred negative connotation. Concerns have been raised about the challenges of sarcasm and the benefits of sarcasm identification to sentiment analysis in the study of computerized sarcasm detection. The term "automatic sarcasm detection" describes computer methods that determine if a text is sardonic (Joshi et al., 2017). This inspired other academics to implement sarcasm detection in a number of significant disciplines (Joshi et al. 2016). conducted research on the factors affecting the accuracy of sarcastic predictions. According to the researcher, using such a method would help in determining the value of fresh data sets. Another piece of work by Kannangara (2018) used sarcasm detection to categorize people's political viewpoints. In this regard, the researchers put up three models for categorizing the sociopolitical polarity of microblog articles. The researchers also suggested a unique sarcasm detection algorithm that classifies sarcastic beliefs using ideology and fine-grained opinion as parameters along with other language variables. Sarcasm detection has a significant industry execution in addition to the political application by utilizing social media platforms. These sites have developed into sizable ecosystems that let people express their thoughts without restriction. As a result, businesses use this ecosystem to obtain significant public opinion regarding elements of their products and services and to offer real-time customer support (Sarsam, 2019). Additionally, these businesses have a robust social media presence and a responsive workforce in charge of marketing and customer service (Rajadesingan et al., 2015). Appropriately, social media platforms that produce a vast amount of information while running these businesses can use tools like HootSuit to handle a variety of challenging tasks, which involve managing the company's content, conducting sentiment analysis, and extracting essential messages for the company's customer care professionals to respond to.

1.2 Research Problem

Like everything else, learning and teaching have undergone significant change as a result of the pandemic we have been dealing with for the past two years. Universities and colleges have switched to online learning platforms. It's a sad reminder that even if we have much more potential than we've allowed ourselves to believe, we still lack the resources and ability to effectively educate anything we do online[16]. More time than ever before is being spent online by professors and students. During the epidemic, their entire way of communicating, learning, and living changed. A lot of queries and problems about how this online-based learning is affecting have also been raised by this change in methods and lifestyle! In our project, we've tried to use social media opinions to analyze the mentality and sentiment of people on online-learning as social media engagement increased a lot during the time of the pandemic. We've also tried to analyze and identify the sarcastic tweets as it's difficult to find out the meaning of the sarcastic tweets.

When sentiment analysis is used on web blogging services like Twitter, the ability to recognize sarcasm enhances the process. Sentiment analysis and opinion mining rely on the emotional vocabulary of a text to identify its polarity. The appeared text can be also false. Sarcastic text is a common instance of that. In a comparable context, "sarcasm is not a separate logical or linguistic phenomenon", as stated by Brown et al. [4], but operates as [8] [9] proposed to discover sarcastic writing styles to identify regardless of whether a speech is sardonic. Such trends have been seen during our research and while manually annotating tweets, especially among English language learners. As a result, we concentrate on identifying and gathering these patterns from a dataset which was manually annotated, quantifying them with the intention we can compare patterns derived from a given tweet to them in order to determine whether or not it is sarcastic.

1.3 Motivation

The paper draws inspiration from various works, including those already listed and many others that were crucial to the development of the current classifier model. In the result analysis section of this work, we provide a number of adjustments that we believe would improve accuracy, such as choosing the proper collection of features.

1.4 Thesis contribution

This thesis could potentially contribute to the advancement of a machine learning model for effectively identifying the sentiment of online learning-related tweets during the period of the COVID-19 pandemic. This model could be trained on a large dataset of labeled tweets and could potentially outperform existing approaches in terms of accuracy. Another potential contribution could be the development of a method for detecting sarcasm in tweets associated with online learning during the COVID-19 pandemic. This method could take into account contextual information and irony cues, such as exaggerated language and negation, to identify sarcastic tweets. It could also potentially be combined with multimodal approaches, incorporating visual information in addition to text, to enhance the precision of identifying sarcasm. Overall, the contribution of this thesis could be the development of novel approaches for analyzing the sentiment and sarcasm of web-based learning related tweets during the COVID-19 situation, which could be useful for educators and researchers studying the effectiveness of online learning during the pandemic. These approaches could also potentially have broader applications in the field of NLP for analyzing sentiment and sarcasm on social media platforms more generally.

1.5 Research Objective

For sarcasm detection in tweets on Twitter, the study aimed to identify the most effective machine learning algorithms and features, additionally the challenges faced by researchers in this area. Sarcasm detection is important for various fields such as security, health, and marketing, as it can help companies understand customer sentiment toward their products. Sarcasm detection is also a crucial subtask of sentiment analysis, particularly for classifying tweets, as it can reveal implicit information in a message. However, detecting sarcasm in tweets can be difficult due to the 280-character limit, and the use of informal language, slang, and abbreviations. Additionally, there is no calculated formation for recognizing sarcasm on Twitter. To address these challenges, previous research has applied machine learning techniques, such as those used by Altrabsheh et al. (2015) who compared several machine learning techniques and found that Complement Naive Bayes (CNB) had the giant recall function. Other studies have also proposed different methods for sarcasm detection, such as the method proposed by Tungthamthiti et al. (2016) which combined two classification algorithms. The results of this review offer recommendations for future research on sarcasm detection in tweets on Twitter, including the most effective machine learning algorithms and features to use. Besides, this project is to create a research analysis on the sentiment of people on online learning during the covid-19 so that it becomes easier to understand and analyze the positive and negative sites of online learning. By doing so we can create necessary modifications in the learning process and also can make effective use of online learning in regular education. First, we categorize each tweet according to the sentiments of the tweets. Then we categorize each tweet in a batch of tweets according to whether it is sarcastic or not. In order to do the classification, we first extract a collection of features from each tweet, referring to a training set, then utilize machine learning methods. The features are retrieved in a method that utilizes several tweet elements and accounts for various sarcasm subtypes. The corpus of tweets that we use for our research is manually reviewed and annotated.

Chapter 2

Literature Review

Natural Language Processing (NLP) has long been focused on researching and developing methods for identifying and extracting subjective information, such as emotions and opinions, from text through the technique of sentiment analysis[16]. It involves classifying text as posi-tive, negative, or neutral in sentiment. Sarcasm detection is a more recent focus of NLP research, as sarcasm is a form of irony that can be difficult for computers to recognize[4]. The ability to detect sarcasm is significant as it can influence sentiment analysis accuracy, as sarcasm is frequently utilized to express a negative sentiment despite the use of positive language. There has been a growing interest in studying online learning and its effectiveness during the COVID-19 pandemic, as many educational institutions have moved to online platforms due to lockdowns and social distancing measures[17][13]. During the COVID-19 epidemic, there was a transition to online learning, which resulted in an increase in the use of social media platforms like Twitter, for communication and collaboration among students and teachers. Therefore, there is a requirement for developing techniques for examining the sentiment and sarcasm of online learning related tweets during this period[7][11].

Previous research has attempted to address the issue of sentiment analysis and sarcasm detection on social media platforms. Several studies have inspired further exploration of this topic, including a technique proposed by authors Sana Parveen and Sachin N. Deshmukh[4] for detecting sarcasm on Twitter through the use of Maximum Entropy and Support Vector Machine (SVM). Many of these studies have employed machine learning methods, such as Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs), to classify tweets as positive, negative, or neutral in sentiment [5]. Some studies have also utilized lexical and syntactic features, such as specific words or punctuation, to enhance the accuracy of sentiment analysis.There have also been several studies specifically focused on detecting sarcasm in tweets. These studies have employed various approaches, such as using contextual information and irony cues, such as exaggerated language and negation, to identify sarcastic tweets. Some studies have also used multimodal approaches, combining text and visual information, to enhance the precision of identifying sarcasm.The primary outlets for expressing one's ideas, thoughts, and opinions on a variety of subjects and events are microblogging sites. Because social networks and microblog-ging websites frequently Sarcasm is a complex form of irony that is commonly found on these platforms, it may lead to trolling and/or criticism of others.

Irony and sarcasm differ slightly from one another (Reyes et al., 2012). The term "sarcasm" is frequently used to refer to a verbal ironic statement (Colston, 2000). Certain forms of irony, including levity, exaggeration, rhetorical questions, and understatement, are paired with it (Gibbs, 2000). Sarcastic irony was described by Kumon-Nakamura et al. (2007) as the antithesis of non-sarcastic irony. According to Gibbs and Colston (2007), parody and satire are frequently used as comparisons for irony. So, Parmar et al. (2018) proposed the following to describe sarcasm on Twitter: Conflicting facts in a tweet include those that are (a) between a negative situation and a posi-tive sentiment, (b) between a positive situation and a negative sentiment, (c) that start with an interjection word, (d) that contradict likes and dislikes, (e) that contain conflicting universal facts, (f) that contain conflicting time-sensitive facts, and(g). Text categorization algorithms have been devised to cope with this complex emergence because of the vast volume of information being produced on social media and the requirement to properly study it. Sarcasm identification in text categoriza-tion is a crucial technique with effects on security, health, and commerce, among other domains (Jain Hsu, 2015). Businesses may examine how customers feel about their products by using sarcasm detection methods. These businesses may improve their product quality with this critical assistance (Saha et al., 2017). The ability to classify sarcasm is a significant sub-task in sentiment analysis, particularly when analyzing tweets, as it allows for the identification of implicit information conveyed within the message a person expresses or shares with others. Additionally, the way a tweet is structured may be used to forecast sarcasm (by, for example, flipping a statement's polarity from positive to negative). Sarcasm identification is challenging on Twitter due to a number of problems. Parmar et al. (2018) outlined a few of the difficulties in identifying sarcastic tweets currently present. These difficulties include (a) the nature of the collected tweets (e.g., Twitter's character limit of 280 for posting tweets, which might cause more ambiguity); (b) The challenges faced in identifying sarcasm in tweets include the presence of a high number of unusual words, slang, and abbreviations that have a more informal tone, c) the lack of a predefined structure for recognizing sarcasm on Twitter. In order to anticipate sarcasm in tweets, earlier research has used machine learning approaches (Jain Hsu, 2015). For example, Altrabsheh et al. (2015) examined several machine learning techniques, features, and preprocessing stages to detect sarcasm from student input collected through Twitter. (2015) examined a number of classifiers that Tian et al. had suggested in order to identify the sarcasm (2014). The results showed that Complement Naive Bayes (CNB) had the highest recall function, as per the findings. For the purpose of identifying sarcasm, Ren et al. (2018) created two distinct context-augmented neural models. Gradient Boost outperformed other classification algorithms in terms of prediction accuracy, according to a study by Prasad et al. (2017). A unique approach for detecting sarcasm in tweets was presented by Tungthamthiti et al. (2016) by fusing two classification algorithms (Support Vector Machine [SVM] with N-gram features and SVM). Based on this, it can be concluded that classifier performance is crucial for precise predictions when handling expressions in the textual input (Sarsam et al., 2019). Additionally, it appears that the kind of classifiers used is important in detecting sarcasm. However, few researchers have looked at how well sar-casm detection algorithms perform on Twitter in relation to the characteristics that are used. Therefore, the main sarcasm classifiers were examined, along with their classification accuracy and the parameters that affected it. This study also looked at the difficulties that earlier researchers had when attempting to identify ironic tweets. Results from this evaluation include recommendations for future researchers as well as practical consequences for the kinds of machine learning algorithms and key characteristics employed in the recognition of tweets with sarcasm.Identifying irony-containing phrases in text with real emotional aspects is recognized as sarcasm detection. For affective computing systems, doing sentiment analysis, sarcasm's figurative and creative nature character presents a significant barrier. There have been three major paradigm shifts in how researchers have handled this issue so far: 1) use of hashtag-based supervision, 2) semi-supervised pattern extraction to identify implicit sentiment, and 3) incorporation of context beyond target text. Anukarsh G Prasad, Sanjana S, Skanda M Bhat, and B S Harish (2019) proposed a method for distinguishing between sarcastic and non-sarcastic tweets based on the slang and emojis used. They looked at the slang and emoji values from the slang and emoji dictionaries. After that, to identify sarcasm in tweets from the Twitter Streaming API, these values are compared with various classification algorithms such as Gradient Boosting, Random Forest,Adaptive Boost, Logistic Regression, Gaussian Naive Bayes and Decision Tree. The best classification algorithm is chosen from among those considered and combined with various preprocessing and filtering techniques using emoji and slang dictionary mapping to produce the highest efficiency.

Overall, the literature on sentiment analysis and sarcasm detection in the context of the shift to online learning during the COVID-19 pandemic implies that machine-learning techniques can be effective in accurately identifying the sentiment and sarcasm of tweets. However,there is still potential for advancement, specifically in the identification of sarcasm, and further research is needed to develop more robust methods for analyzing sentiment and sarcasm on social media platforms.

Author	Labeling Approach	Type of Source	Algorithm	title	Category
Sana Parveen & Sachin N. Deshmukh	Sarcastic or non sarcastic	Article	ML	Opinion Mining in Twitter – Sarcasm Detection	CMLA
Gibbs & Colston	Polarity: positive and negative tweets	Journal	ML	Emotional Reactions to Verbal Irony	CMLA
Parmar et al	Polarity: positive and negative tweets	Journal	Feature-based Composite Approach (FBCA)	Feature based composite approach for sarcasm detection using MapReduce.	CMLA
Altrabsheh et al. (2015)	Sarcastic or non sarcastic	Books	Several ML	Detecting Sarcasm from Students' Feedback in Twitter	CMLA
Sarsam et al., 2019	Sarcastic or non sarcastic	Journal	SVM	Sarcasm detection using machine learning algorithms in Twitter: A systematic review	CMLA & AMLA
Prasad et al. (2017)	Sarcastic or non sarcastic	Article	Gradient Boost	Sentiment analysis for sarcasm detection on streaming short text data.	CMLA

Ren et al. (2018)	Polarity: positive and negative tweets	Journals & Books	MODEL-KEY	Context-augmented convolutional neural networks for twitter sarcasm detection.	CMLA
Tungthamthiti et al. (2016)	Sarcastic and non sarcastic	Article	SVM	Recognition of sarcasms in tweets based on concept level sentiment analysis and supervised learning approaches.	AMLA
Saha et al. (2017)	Polarity: positive, negative or neutral	Indian Journal	Naïve Bayes	Proposed approach for sarcasm detection in twitter	AMLA
Jain & Hsu (2015)	Sarcastic and non sarcastic	Article	Logistic regression	The lowest form of wit: Identifying sarcasm in social media	AMLA
Kumon Nakamura et al. (2007)	Sarcastic and non sarcastic	Books	ML	Irony in language and thoughts	CMLA
(Reyes et al., 2012)	Sarcastic and non sarcastic	Journal	ML	Sarcasm Detection on Czech and English Twitter	CMLA

Figure 2.1: Summary Table of Literature review

Chapter 3

Methodology

3.1 Data collection

"Online learning Tweet Data.csv" is the name of our first dataset. The tweets were gathered using the Advanced Search feature of the Twitter API and the Search Twitter "operator" built into Rapid-Miner Studio. It is possible to develop, implement, and test a broad variety of algorithms, processes, and applications in the domains of Big Data, Data Mining, Data Science, Artificial Intelligence, Machine Learning, and other related ones using a data science platform called RapidMiner. The built-in "operator" of RapidMiner called "Search Twitter" connects to the Twitter API and uses its default search criteria to get tweets with one or more entered keywords or phrases that were posted within two specified dates. Based on research on widely used words for web-based learning, COVID-19, and the Omicron variation, a set of words was produced. This is due to the fact that Twitter users may use a variety of languages while discussing subjects like COVID-19, the Omicron variation, and online education. These words, phrases, and antonyms were all used to compile the information. The "operator" of Search Twitter used this "method" to ensure that the search results included at least one synonym, term, or phrase for COVID-19 and at least one synonym, term, or phrase for online learning. The results of this RapidMiner "process" included the Tweet ID, Tweet Source (the program used to submit the Tweet, such as Twitter for Android or Twitter for iOS), Text of the Tweet, Retweet count, and Username of the Tweeter who sent the Tweet on Twitter, among other attributes. According to the guidelines in the Twitter API standard search regulations, all of these attributes are open data that can be mined. However, all attributes other than Tweet IDs were removed using data filters in order to adhere to Twitter's developer guidelines, privacy policies, and guidelines for content redistribution. Therefore, the collection only includes Tweet IDs. A process known as "hydration of Tweet ID" can be used to extract the entire collection of data related to a tweet, including the tweet's content, user name, user ID, date, retweet count, etc. With an emphasis on Twitter study, researchers in the fields of Big Data, Data Mining, and Natural Language Processing have created a number of tools for the hydration of Tweet IDs. Every Tweet ID that was present in this dataset at the time this paper was being published matched a tweet that had not yet been deleted. Second, the Twitter API's search function does not return a comprehensive list of tweets that were posted over a specific period of time. Since the search feature of the Twitter API did not fetch them at the time the data was gathered, a number of tweets that might have been made between these dates might not have been included in this dataset.

our initially collected dataset looks like the figure below:

Created-At	From-User	From-User-Id	To-User	To-User-Id	Language	Source	Text
2022-08-24 20	Camilo Solano	1106206240525467648		-1	en	<a href="https://	RT @ajlamesa: "Masks will no long
2022-08-24 19:2	Just Shana ??	8228914165233	An3ita5	7927238404052	en	<a href="http://t	@An3ita5 I have a friend who is a t
2022-08-24 18:4	MyJournals	52728374		-1	en	<a href="https://	IJERPH, Vol. 19, Pages 10556: Emot
2022-08-24 15:5	Red — Taylor's v	1189524852006195202		-1	en	<a href="http://t	RT @TROmhpnhs: NEWS MHPNHS PUSHES IN-PERSON CLAS After two consecutive years of dist
2022-08-24 13:1	The Republic On	1018791261304598528		-1	en	<a href="https://	NEWS MHPNHS PUSHES IN-PERSON CLAS After two consecutive years of dist (1/5) https://t.co/hv8Q9LuoQA
2022-08-24 12:2	The Freeman	262657733		-1	en	<a href="http://t	TOP STORY: Two non-government c
2022-08-24 05:1	Anthony LaMesa	277851432		-1	en	<a href="https://	"Masks will no longer be a requirer https://t.co/sK9pSZ256z
2022-08-24 04:3	Ed Barnard	25417067	ewbarnard	25417067	en	<a href="https://	46. As Covid hit, the USAFA decide
2022-08-23 23:2	A Rose Extravaga	1139763596588	RustyAbbe	1011545834	en	<a href="https://	@RustyAbbe @trad_thinker @rnas I agree that distance learning was s Teachers already get enough shit fr
2022-08-23 23:0	Michael K Barbo	12969492		-1	en	<a href="http://	Article Notice – Distance Learning S
2022-08-23 19:3	Mike Closs	8337393245417	RodKahx	9787587981763	en	<a href="https://	@RodKahx Some of us still wear m
2022-08-23 17:0	Sam Washington	1246124670312681479		-1	en	<a href="https://	RT @Brill_Education: The nine cha
2022-08-23 10:5	PHL News Inside	257277393		-1	en	<a href="http://t	RT @PhilippineStar: Millions of ma

Figure 3.1: Initially collected data from Twitter API using rapid miner

3.2 Data pre-processing

First, we cleaned up the data, eliminating any instances where the hashtag was used in one of the two first uses of the three described below as well as those that were noisy and irrelevant.

3.2.1 Pre-processed data for sentiment analysis

For pre-processing data for sentiment analysis, we've removed the repetitive words, and words from other languages except for English, also cleaned the hashtags and other signs using the python regular expression module, and then also used the PorterStemmer to clear the text for the sentiment analysis.

Our cleaned data looks like the figure below:

	Text	level_sarcasm
0	rt ajlamesa masks longer requirement students ...	1
1	an3ita5 friend teacher great one covid set sch...	0
2	ijerph vol 19 pages 10556 emotional impact cov...	0
3	rt tromhpnhs news mhpnhs pushes inperson class...	0
4	news mhpnhs pushes inperson classes two consec...	0

Figure 3.2: pre-processed data for sentiment analysis

3.2.2 Pre-processed data for sarcasm detection

1. to serve as an anchor for searches;

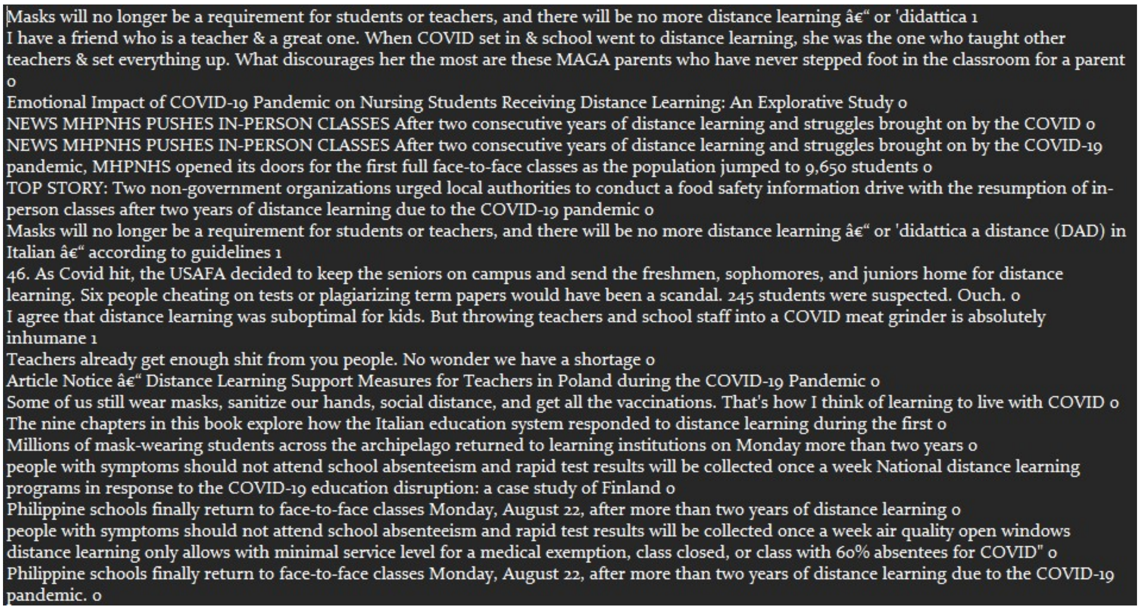
2. For example, "I failed to include sarcasm so guys like you get it!" to elucidate why sarcasm is present in a prior tweet;

3. to serve as a sarcasm marker when using very mild sarcasm, as in "Today was fun," where it can be challenging to understand the sarcasm without an explicit label. after several days, finally! Sarcasm". Regarding the non-sarcastic tweets, we gathered them and ensured that they have some sort of emotional element. For our research, we created the following 3 data sets.

Sample 1: In the first dataset the tweets in this data collection have been personally reviewed and graded 0 and 1 manually based on whether they are sarcastic or not. If they are sarcastic the value is 1 and if they are not the value is given 0. Two individuals who have no prior knowledge of the tweets or the folks who tweeted them manually annotate them. They are expected to assign the scores. It's vital to remember that manual labeling depends on the annotators' personal judgment. As a result, the classification's flaws are acknowledged. However, sarcasm is never considered to be non-sarcastic in a tweet, and the contrary is equally true.

Sample 2: We removed non-English tweets and extremely brief tweets to reduce the amount of noise(those with fewer than three words), and tweets that contained URLs. Most often, URLs are used to refer to photo links. We disregard them since we think that the snapshot has some sarcasm. The parameters we select for our features are optimized using this data set during our experimentation phase. We can call this collection the "optimization set" for the remainder of this paper.

Sample 3: Each tweet is individually examined to determine whether it contains satire or not. This set will be considered a test set to evaluate the effectiveness of our proposed method. It will be considered as the "test set" throughout this study.



Masks will no longer be a requirement for students or teachers, and there will be no more distance learning â€“ or 'didattica 1
I have a friend who is a teacher & a great one. When COVID set in & school went to distance learning, she was the one who taught other teachers & set everything up. What discourages her the most are these MAGA parents who have never stepped foot in the classroom for a parent 0
Emotional Impact of COVID-19 Pandemic on Nursing Students Receiving Distance Learning: An Explorative Study 0
NEWS MHPNHS PUSHES IN-PERSON CLASSES After two consecutive years of distance learning and struggles brought on by the COVID 0
NEWS MHPNHS PUSHES IN-PERSON CLASSES After two consecutive years of distance learning and struggles brought on by the COVID-19 pandemic, MHPNHS opened its doors for the first full face-to-face classes as the population jumped to 9,650 students 0
TOP STORY: Two non-government organizations urged local authorities to conduct a food safety information drive with the resumption of in-person classes after two years of distance learning due to the COVID-19 pandemic 0
Masks will no longer be a requirement for students or teachers, and there will be no more distance learning â€“ or 'didattica a distance (DAD) in Italian â€“ according to guidelines 1
46. As Covid hit, the USAFA decided to keep the seniors on campus and send the freshmen, sophomores, and juniors home for distance learning. Six people cheating on tests or plagiarizing term papers would have been a scandal. 245 students were suspected. Ouch. 0
I agree that distance learning was suboptimal for kids. But throwing teachers and school staff into a COVID meat grinder is absolutely inhumane 1
Teachers already get enough shit from you people. No wonder we have a shortage 0
Article Notice â€“ Distance Learning Support Measures for Teachers in Poland during the COVID-19 Pandemic 0
Some of us still wear masks, sanitize our hands, social distance, and get all the vaccinations. That's how I think of learning to live with COVID 0
The nine chapters in this book explore how the Italian education system responded to distance learning during the first 0
Millions of mask-wearing students across the archipelago returned to learning institutions on Monday more than two years 0
people with symptoms should not attend school absenteeism and rapid test results will be collected once a week National distance learning programs in response to the COVID-19 education disruption: a case study of Finland 0
Philippine schools finally return to face-to-face classes Monday, August 22, after more than two years of distance learning 0
people with symptoms should not attend school absenteeism and rapid test results will be collected once a week air quality open windows distance learning only allows with minimal service level for a medical exemption, class closed, or class with 60% absentees for COVID" 0
Philippine schools finally return to face-to-face classes Monday, August 22, after more than two years of distance learning due to the COVID-19 pandemic. 0

Figure 3.3: pre-processed text data for sarcasm detection

3.3 Analytic tools and methods

We have installed and used multiple NLP python modules and libraries on google colab for the task. We've used nltk, re, sys, os, textblob, gensim, pickle, logging and the sklearn libraries for doing different NLP tasks. Besides we used some special modules such as matplotlib for the visualization of the results and showing the metices. Due to tweets' unconnected content, use of slang, etc., and OpenNLP's PoS tagger's poor performance when used with the given model to tag them, we decided to use part-of-speech tagger instead. This PoS tagger obtains a 93 percent accuracy rate when used with Twitter data.To do the categorization using a Support Vector Machine, we used libsvm [33]. (SVM). In case of sarcasm extraction specifically for Stemming and Lemmatization we imported sentiWord net and used topic modelers.

Chapter 4

Implementation

4.1 Techniques Used in Sentiment Analysis

Sentiment analysis techniques involve the application of Natural Language Processing (NLP) and Machine Learning Algorithms to classify and analyze text data to identify the sentiment or opinion expressed. Some prevalent techniques used in sentiment analysis include:

4.1.1 Dictionary-based approaches and opinion mining:

This technique involves the use of a predefined dictionary or lexicon of words with assigned sentiment values (e.g. positive, negative, neutral). Text is then analyzed by counting the frequency of positive and negative words and determining the overall sentiment.Opinion mining involves the techniques of NLP to identify and extract subjective information from text, including sentiments, opinions, and emotions. This information can then be used to classify the overall sentiment of the text.

4.1.2 Machine Learning Approach

The creation of analytical model construction is computerized. through the use of the data analysis technique known as Machine Learning. The foundation of the artificial intelligence field is the idea that computers are capable of learning from data, spotting patterns, and drawing conclusions with little to no human involvement. It is not so literal, though. Computers now experience emotions despite never having a brain.Simply put, machine learning creates algorithms that let computers learn on their own and provide definite answers to particular queries. Algorithm development can be done in two ways: supervised learning and unsupervised learning.This strategy, however, is the one that is most frequently utilized in sentiment analysis. This technique uses classifiers to automatically identify the labels of the new data. Due to the rapid expansion of web-based media, the majority of client-produced material is readily available. It is a very difficult process to analyze the attitudes and precise classification of this vast volume of data. Since writing is the most creative and readable way for presenting the users' thoughts and opinions, the text makes up the

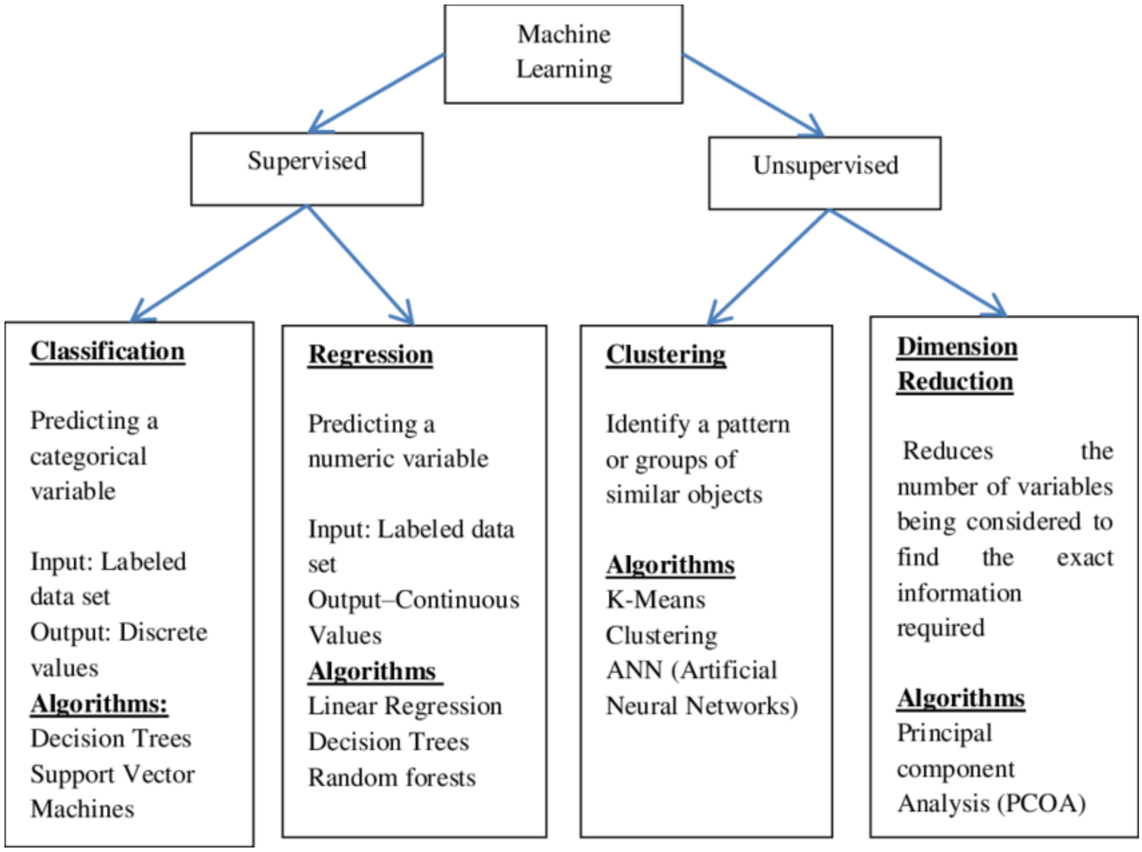


Figure 4.1: Machine Learning Classification

majority of internet information [30]. Typically, binary classification and predictions of attitudes as either positive or negative are done using machine learning techniques.

4.1.3 Lexicon Based Approach

Using lexicons to conduct sentiment analysis involves the use of a predefined dictionary or lexicon of words with assigned sentiment values (e.g. positive, negative, neutral). The text being analyzed is then divided into individual words, and each word is looked up in the lexicon to determine its sentiment value. The overall sentiment of the text is then calculated by summing the sentiment values of all the words and determining the majority sentiment (e.g. if more positive words are present, the overall sentiment is positive).

One advantage of this technique is that it is relatively simple and easy to implement. However, it can be limited in its accuracy, as it may not always accurately capture a word's meaning when used in circumstances.For example, the word "not" can negate the sentiment of a word, but a lexicon-based approach may not account for this. Additionally, some words may have multiple meanings and can have different sentiment values depending on the context in which they are used. Lexicon-based approaches.

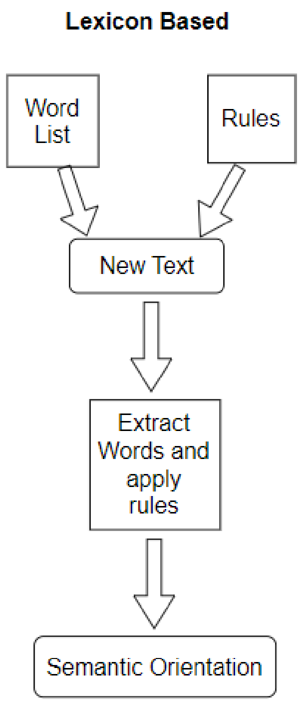


Figure 4.2: Lexicon Based approaches

4.2 Sarcasm Detection Features Extraction

There are many uses for sarcasm as a sophisticated type of communication. The data annotators came to the conclusion that these aims primarily, but not totally, fell into three categories:

1. wit
2. whimper
3. avoidance

1. Sarcasm as wit: Sarcasm is used to make others laugh; to make it obvious, the individual utilizes certain speech patterns, a tendency to exaggerate, or a tone that differs from his normal speaking style. Voice tones are translated into unique writing styles on social media platforms, including the usage of capital letters, exclamation and question marks, and certain sarcastic emoticons.

2. Sarcasm as a whimper: SA person might show his irritation or rage by using sarcasm. As a result, it might be easy to overstate the situation's seriousness or serious the situation is or to use good language to describe a bad one.

3. Sarcasm as evasion: It describes an instance where the speaker wants to avoid giving a direct response. Sarcasm is consequently employed. The speaker in this instance used convoluted sentences, odd vocabulary, and a few peculiar terms.

The classification and observations are based on the reasons why sarcasm is utilized, as opposed to the classification, which divides sarcasm into four groups depending on how sentiments emerge in the sentence. We use these observations to create our model even though they are probably biased and based on the writer's own preferences. To identify sarcastic statements while working, we largely rely on writing patterns; however, additional elements are extracted to increase classification precision and accuracy. The use of various aspects is highlighted by the separation of purposes, which we shall cover next. Extractions include elements linked to sentiment, punctuation, syntactic and semantic structure, and pattern.

4.2.1 Sentiment-related Features

The use of various aspects is highlighted by the separation of purposes, which we shall cover next. Extractions include elements linked to sentiment, punctuation, syntactic and semantic structure, and pattern. To find this distinction in unidentified tweets, they developed a lexicon-based technique that learns potential positive terms and unfavorable scenarios.

However, because negative occurrences are unexpected, learning about every potential unfavorable situation would require a large and rich source and might be impossible. In our job, we choose a simpler yet more comprehensive strategy. We thus take the emotive components of the tweet and tally them in order to discover and quantify such inconsistencies. For this reason, we maintain separate lists of words labeled as "positive words" and "negative words." Words with positive emotional connotations (such as "love," "happy," etc.) and words with negative emotional connotations (such as "hate," "sad," etc.) are included in the two lists. The two-word lists are created using the SentiStrength database. The list of emotional words in this database ranges from -1 (nearly negatively) to -5 (very negatively) and from 1 (almost positively) to 5 (extremely positively) (extremely positive). By calculating the number of positive and negative words in the tweet, we are able to extract two characteristics, PW and nw, from these two lists. More emotion is present in adjectives, verbs, and adverbs than in nouns. The number of emotionally charged positive words and negatively charged negative words, labeled PW and NW, respectively, has been calculated again using the positive and negative words with the corresponding PoS-tag. Then, three more attributes are added by counting the proportion of positive, negative, and sarcastic emoticons. Sarcastic emoticons are those that accompany sardonic or sarcastic words (for instance, ":P").

Part of Speech	Part of Speech Tag
Adjectives	"JJ", "JJR", "JJS"
Adverbs	"RB", "RBR", "RBS"
Verbs	"VB", "VBD", "VBG", "VBN", "VBP", "VBZ"

Table 4.1: First and Second Segment Coefficient (Bat Second)

When someone is attempting to be comical or to show that he is just making a joke, they may use these emoticons (i.e., when sarcasm is used as wit). Emotional content can also be found through hashtags. They are sometimes employed for making clear the Twitter user's genuine objective as stated in his post. As an illustration, the hashtag "Thank you very much for being there for me ihateyou" in the following tweet implies that the user is criticizing the addressee for not being there for him.

4.2.2 Punctuation-Related Features

To identify all sarcasm, sentiment-related criteria are insufficient. In addition, not every element of the tweet is utilized. As previously said, sarcasm is an innovative form of communication since it manipulates word meaning and also behavioral elements including low tones, facial movements, and exaggeration. These elements are turned into precise punctuation or vowel repetition when the message is written. We extract a group of properties known as punctuation-related features in order to recognize these traits. For every tweet, we detect the following values:

- Number of Exclamation Marks
- Number of Question Marks
- Number of Dots
- Number of All-Capital Words
- Number of Quotations

By examining if any of the words (such as "looooove") which consist of a repetitive vowel more than 2, we additionally include a sixth attribute. The feature can be declared as "true" if such a word present else, it is set to "false." " However, this tone is not always sarcastic. For example, the "excessive" apply of exclamation or question marks, or the repeating of a vowel, especially in an evocative term, may indicate the user's intended tone. These traits, in our opinion, are closely related to how many words are in a tweet.

- "Thank you @laur3en, it was amazing !!!"
- "Thanks for another amazing day with your amazing boyfriend!!!!"

Exclamation points are used to indicate sincere thankfulness in the first situation. In contrast, the exclamation marks in the second example indicate displeasure; The user is not genuinely thanking his mate. It is projected that this feature will add value to the categorization when combined with other features because the use of exclamation marks by themselves is useless and may not indicate if the user is sarcastic or expressing any other emotion. The final element is determined by counting the words in the tweet. All told, seven punctuation-based features were located.

4.2.3 Syntactic and Semantic Features

Some popular idioms are frequently employed in a sarcastic context in addition to punctuation-related characteristics. By contrasting these terms with the punctuation, it is possible to identify whether or not anything is being spoken in a sarcastic manner. In addition, it's normal for people to use unusual words or construct complex statements while trying to avoid giving the reader or listener a clear response. This happens frequently when sarcasm is employed as a kind of "evasion," or when a person wants to conceal their true feelings or opinions.

As a result, we extract the properties listed below, which reflect these aspects:

1. The use of unusual vocabulary.
2. The quantity of unusual vocabulary.
3. Presence of usual sarcastic expressions.
4. The interjections quantity.
5. The laughing faces quantity.

Specifically, "Existence of common sarcastic expressions", this feature is retrieved using the previous methodology as "pattern-related" features. We extracted every pattern that may exist, with lengths ranging from 3 to 6, and we only kept the patterns that showed up more than ten times. Since we were small in number, we personally reviewed the list and eliminated any unnecessary items. A total of 13 major patterns were discovered.

4.2.4 Pattern-Related Features:

The "popular sarcastic phrase" patterns chosen in the above paragraph are widely used, especially in spoken English. However, they are uncommon, few in number, and absent from the majority of tweets in both our data sets which are training set and train set. Given that, we delve deeper and pull out a different set of features. We thus suggest more effective and dependable patterns. We categorize words into two groups: one group, called "CI," contains words whose content is crucial, while the other group, "GFI" contains terms with a more significant grammatical function. A word is lemmatized if it falls under the first category; otherwise, an alternative phrase is used in its stead. The TABLE displays the phrases that were used in place of these terms.

POS tags	Definition	Class
CC	Coordinating Conjunction	CI
CD	Cardinal digit	GFI
DT	Determiner	CI
EX	Existential there	CI
FW	Foreign Words	GFI

A pattern is defined as an organized string of words. The training set is used to choose patterns, which are then chosen for acceptable durations.

$$(L_{Min} \leq Length(pattern) \leq L_{Max}) \tag{4.1}$$

LMin and LMax stand for the minimum and maximum allowed lengths of patterns in words, and length(pattern) is the expression for the pattern's length in words. There are

$$NL = (L_{Max} - L_{Min} + 1) \tag{4.2}$$

pattern lengths in total.

4.2.5 TF-IDF Features:

The TF-IDF statistic basically shows the significance relating a phrase (term) to a documentation in the corpus. The amount of times a word appears in other papers in the TF-IDF must be compared to the frequency of that word in the current text (Cong et al., 2016).To avoid word filtration in text summarize and categorization tools, TF-IDF is commonly applied. Additionally, it is applied to proportionally enhance the frequency with which a term shows in a manuscript. However, TF-IDF typically offsets the word's frequency in the corpus; as a result, this technique aids in regulating the more common usage of some terminology than others (Christian et al., 2016). Due to this benefit, TF-IDF was utilized to extract sarcasm-related characteristics in earlier studies that looked at sarcasm detection. Using TF-IDF values, Zhang et al. (2016) sorted the terms in old tweets. The writers treated the collection of historical tweets for a specific collection of data as a single document and created additional documents for TF and IDF estimation using all of the tweets in the training corpus. Additionally, Ren et al. (2018) modeled all contextual tweets as a single document and utilized TF-IDF to order all words in sarcastic tweets. They picked words with the greatest TF-IDF values that were the most significant as a feature for the Machine Learning algorithm's intake.

4.2.6 POS taggers Features:

POS taggers were created to group words into categories according to their POS types. Due to the following factors, sentiment analysis uses it frequently: (A) Pronouns and nouns typically lack emotional content. Such terms can be filtered out using a POS tagger: (b) Words that can be utilized in various POS can be found using a POS tagger. Researchers are using the POS tagger to analyze sarcastic tweets because of its advantages.For instance, Ghosh et al. (2015) employed POS to model situational data for algorithm co-training, yielding a strong corpus and precise predictions. Using the Stanford POS tagger, Kovaz et al. (2013) annotated each statement in the acquired corpora. The writers were interested in adverbs, adjectivals, and interjections, especially when they occurred immediately after an adverb or an adjective. Prasad et al. (2017) looked at the viability of classifying words in a tweet and their POS using POS tagging.A person is said to be offering clues regarding sarcastic tweets if they use a lot of adjectives. In a different study, Barbieri et al. (2015) employed POS by integrating particular variables intended to preserve both the good and bad tweets' structures.Researchers have frequently employed POS tagging in statement annotation due to its considerable contribution to the classification challenge.

4.2.7 Stemmed features

The foundation of stemming is the notion that word meanings that share the same stem are related. The world list of words will be significantly trimmed down by the stemming process (Rani et al., 2015). Stemmers can be utilized to group terms, decreasing the size of indexed files and improving the effectiveness of retrieval (Nayak et al., 2016). According to Rani et al. (2015), the method of stemming employs an affixed removal algorithm to eliminate the word's prefixes and suffixes from the text.In order to accurately perform this operation, several algorithms are used. Generally speaking, there are three categories of stemming algorithms: truncating methods (such as Porter's stemmer), statistical approaches (such as hidden Markov model stemmer), and mixed methods (e.g., Krovetz stemmer). These algorithms can be used to extract stemmed features from the task model during the data preprocessing step thanks to stemming.As a result, Signhaniya et al. (2015) applied the snowball stemmer during the preprocessing of the data step to take out the stemmed features pertinent to the sarcasm job. Using the stem Porter operator, Saha et al. (2017) produced to do polarity testing for the sarcasm detection job, stemmed words were used. Therefore, developing stemmed features during preprocessing operation may aid in the recognition of sarcasm. This is so that the performance of the Machine Learning Algorithm is not adversely impacted by stemmed features, which might minimize a word's derivationally connected forms and inflectional data forms.

Chapter 5

Experimental Results

We have used four performance analysis metrics to perform results analysis as below:

- **Accuracy:** It shows how accurately the classification was made in general.
- **Precision:** Precision shows the proportion of witty, applicable tweets that were found when searching. To put it another way, it compares the proportion of tweets that were correctly identified as being sarcastic to all tweets that were identified in this way.
- **Recall:** it shows the percentage of tweets that are sarcastic are recovered that are applicable. That is to say, it compares the overall number of sarcastic tweets to how many tweets have been successfully classified as humorous.
- **F1-score:** It is a measure that combines the precision and recall of a model by using a weighted average. It takes into account both False Positive and False Negative predictions. The formula for calculating the F1-score is $(2 \cdot Precision \cdot Recall)/(Precision + Recall)$.

5.1 Positive, Negative Neutral Sentiments

The results will be shown in a pie chart that will display the distribution of positive, negative, and neutral sentiment hashtags as a percentage. It illustrates the distribution of sentiments. Neutral hashtags are those that have been assigned a value of zero.

The pie chart in Fig. 4 shows the percentage of neutral, negative and positive, sentiment hashtags with different colors. Here we can see that, Positive sentiment is 52.0%, Negative sentiment is 28.1% and Neutral sentiment is 19.9%.

Then we used wordcloud to find the most frequent words in positive tweets. Wordcloud is a visualization tool used in python to display the most frequent words in a text or a group of texts in the form of a cloud.

The size of each word in the cloud represents its frequency, with the terms that are used the most appearing larger and less frequently used words appearing smaller.

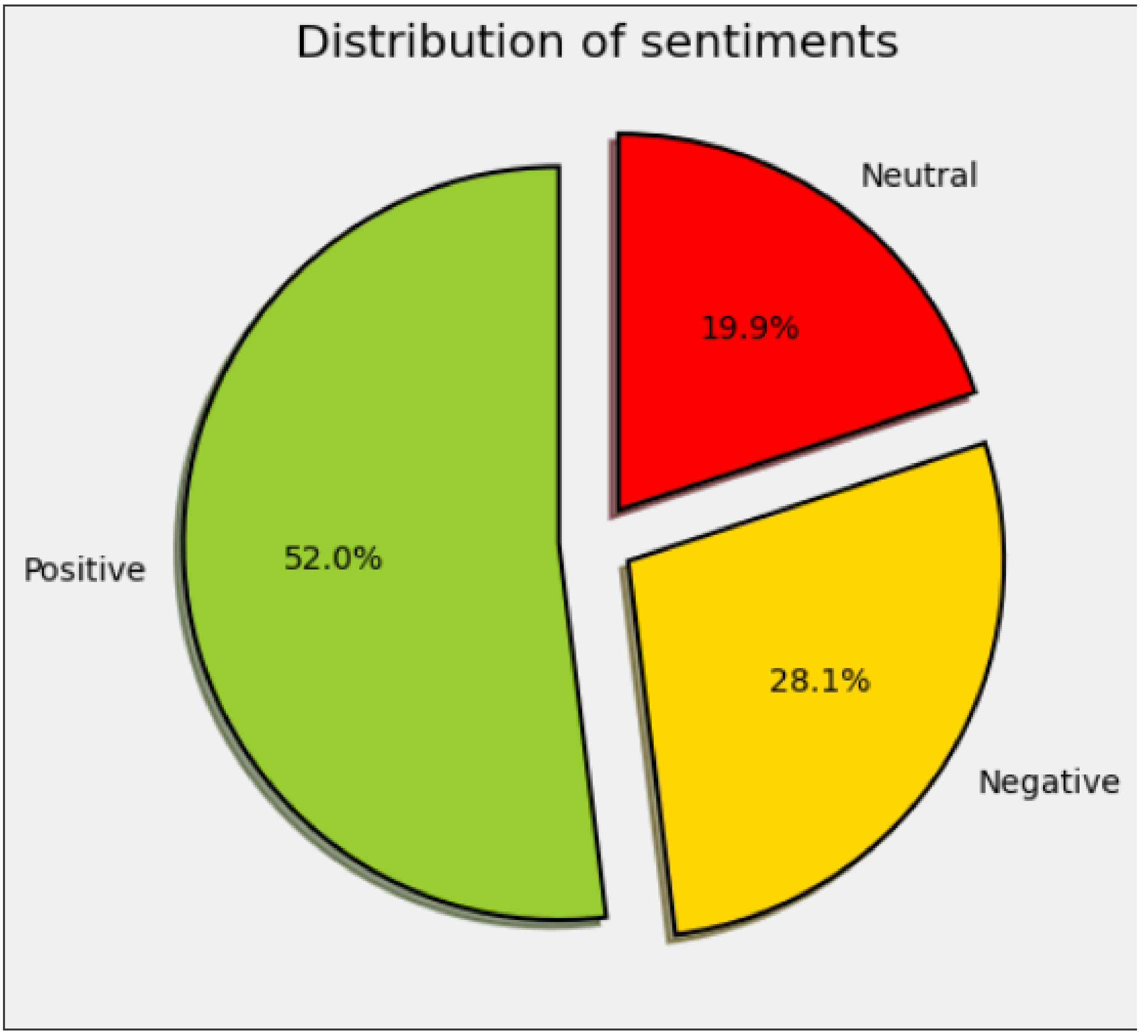


Figure 5.1: Pie Chart Of Distribution of Sentiments

In the same way, the most common phrases used in unfavorable tweets were found.

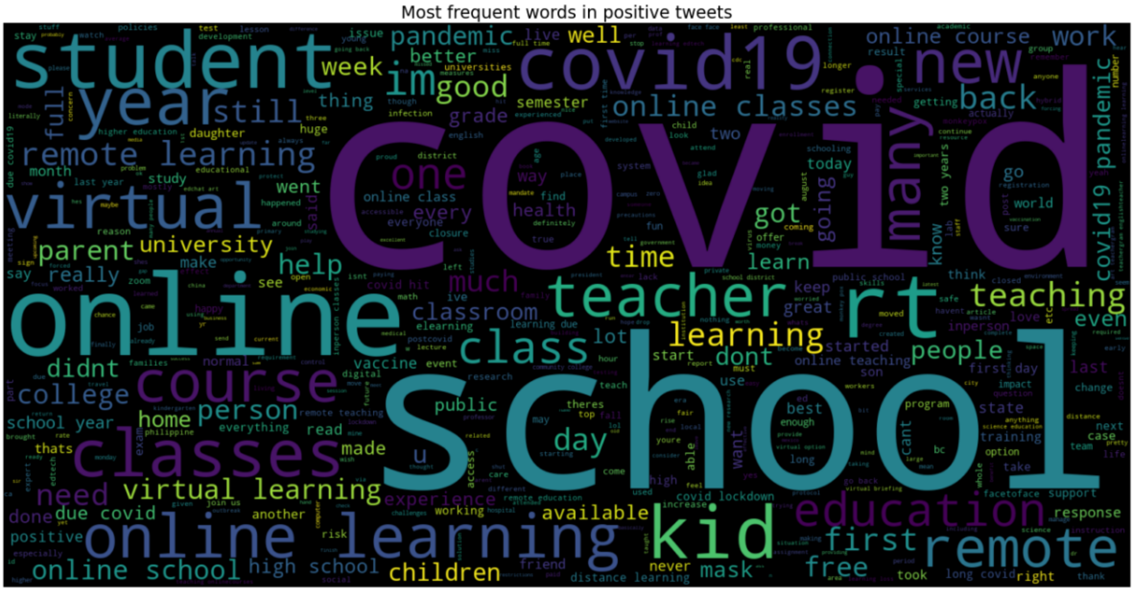


Figure 5.2: Most Frequently used Words in Positive Tweets

Again, the most frequently used words in Neutral tweets were found.

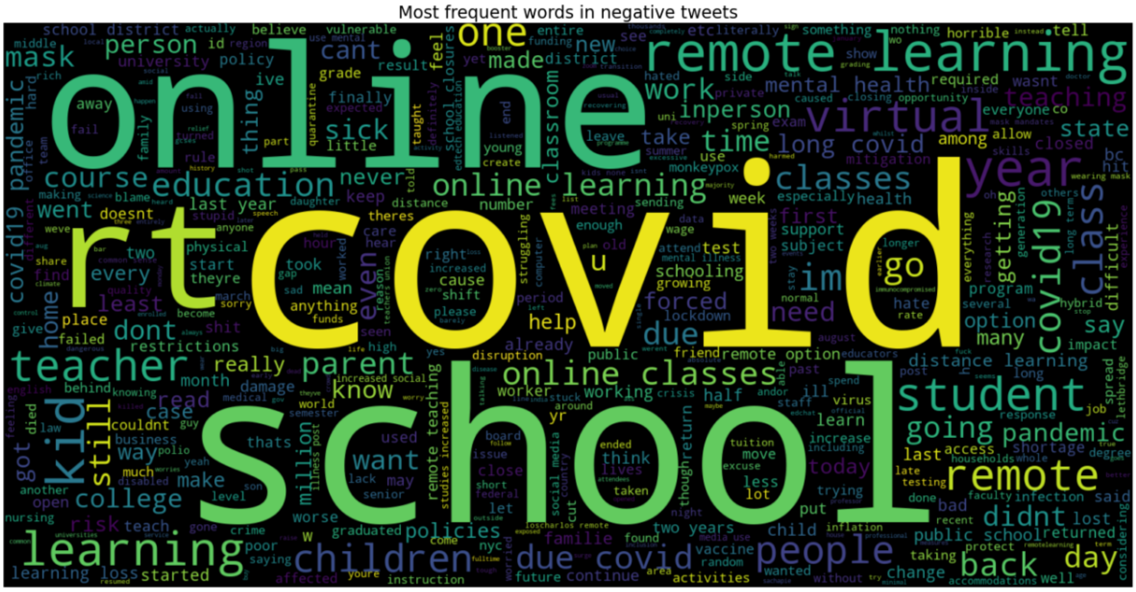


Figure 5.3: Most Frequently used Words in Negative Tweets

5.2 Applying the Logistic Regression

With the help of the Classification report tool we evaluated the performance of the Logistic Regression Algorithm. It is generated by applying the model to a set of test data and comparing the predicted labels to the true labels. The classification report displays several evaluation metrics, including precision, recall, f1-score, and support.

Tweets	Precisionn	Recall	f1-score	Support
Positive	69%	88%	77%	154
Negative	57%	53%	62%	90
Neutral	74%	53%	61%	59

Table 5.1: Logistic Regression Model Classification Report

According to the above table, the Decision Tree classifier with only positive feature set produced a recall value of 88% which is greater than the other two features. The precision value of the negative feature set is 57% and for neutral it is 74% which is higher than the positive and negative features precision value. After training the data on the Logistic Regression model the accuracy is 70.63%.

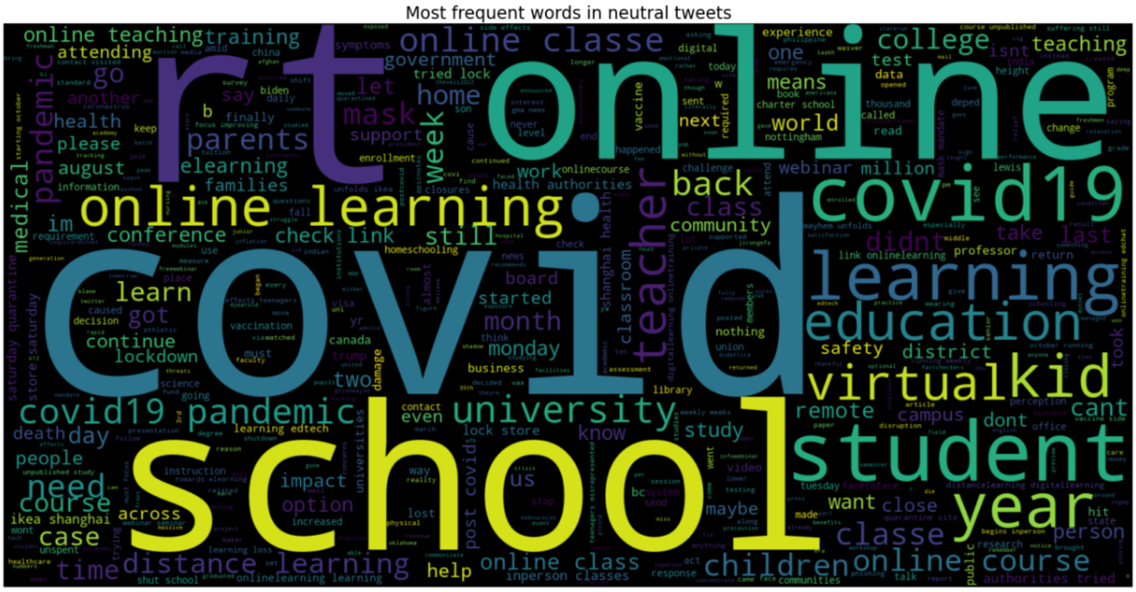


Figure 5.4: Most Frequently used Words in Neutral Tweets

5.3 Tuned Logistic Regression (Hyperparameter Tuning)

When we applied hyperparameter tuning the accuracy level changed to 69.64%. Because of hyperparameter tuning on the Logistic Regression model the accuracy level became slightly lower. The hyperparameter tuning was done by using GridSearchCV. GridSearchCV is a method for hyperparameter tuning that allows you to specify a grid of hyperparameters to search over. It will train a model using each combination of hyperparameters and evaluate the performance using cross-validation. We are using a Support Vector Machine (SVM) classifier and searching over two hyperparameters: 'C' and 'gamma'. The grid of hyperparameters to search over is defined in the 'param grid' dictionary, which specifies that the 'C' hyperparameter should be searched over the values 0.1, 1, and 10, and the gamma hyperparameter should be searched over the values 1, 0.1, and 0.01. The 'cv' parameter specifies the number of folds to use for cross-validation. Finally, the 'scoring' parameter specifies the performance metric to use for evaluation.After fitting the GridSearchCV object to the training data, the 'best params ' attribute gives us the best combination of hyperparameters found, and the 'best estimator ' attribute gives us the best model.

Tweets	Precisionn	Recall	f1-score	Support
Positive	69%	86%	77%	154
Negative	73%	53%	62%	90
Neutral	68%	51%	58%	59

Table 5.2: Updated Classification after Tuned Logistic Regression

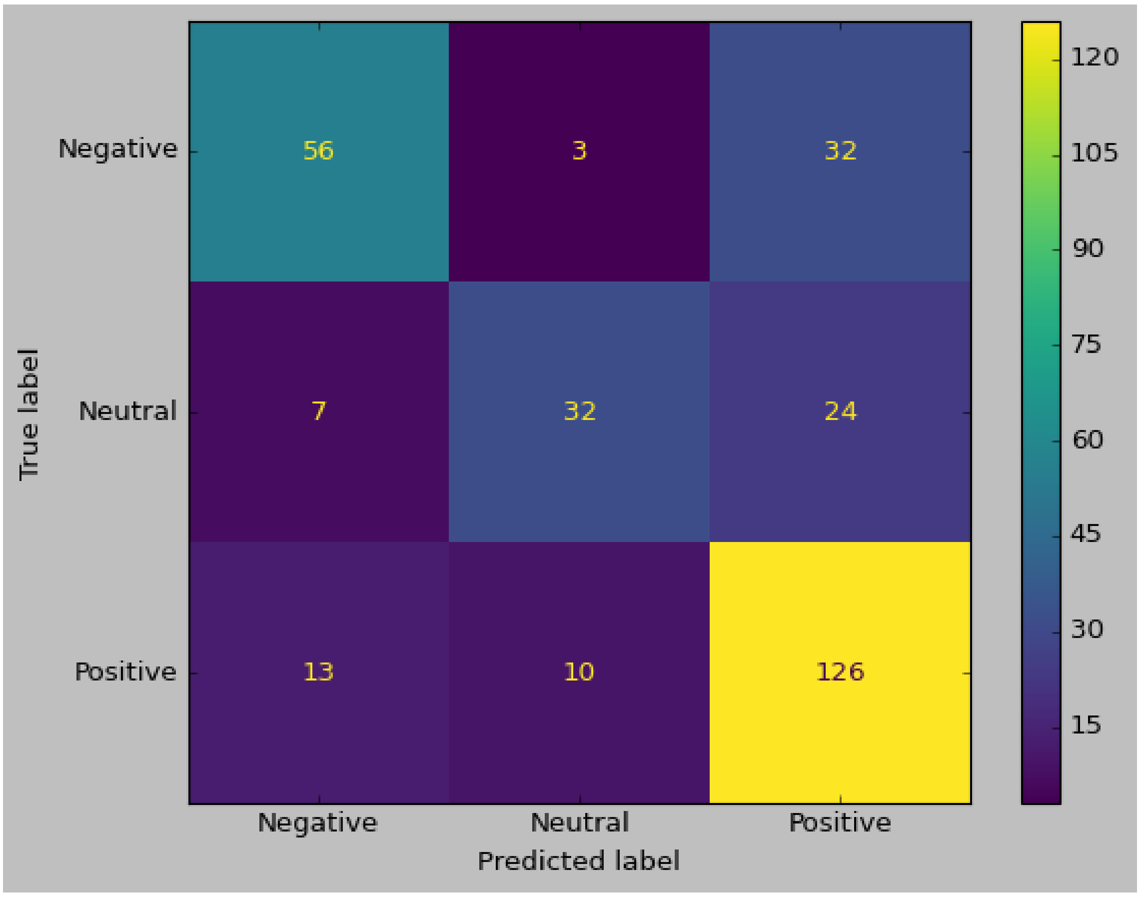


Figure 5.5: Confusion Matrix from Logistic Regression Model

5.4 Applying Support Vector Machine(SVM)

Then we ran the model on another algorithm called Support Vector Machine(SVM). We got 70.30% accuracy from the SVM model. In this algorithm the 'fit' method is used to use the training data to train the model, and the 'predict' method is applied to make predictions about the testing set.

Tweets	Precision	Recall	f1-score	Support
Positive	72%	84%	78%	154
Negative	73%	57%	64%	90
Neutral	60%	54%	57%	59

Table 5.3: Classification Report after SVM model

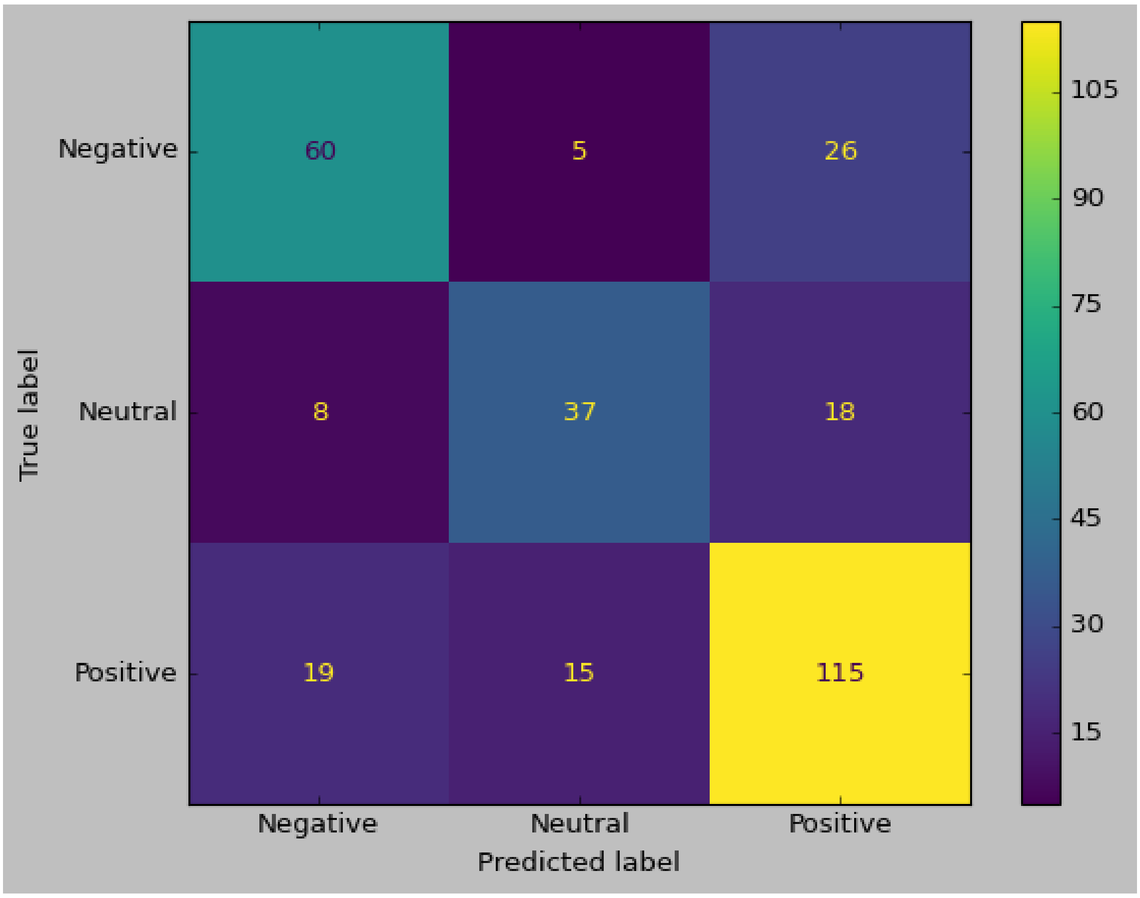


Figure 5.6: Confusion Matrix from SVM model

5.5 After SVM Tuning

Tuning the hyperparameters of an SVM can help improve its performance on a given dataset. The kernel is a function that determines the similarity between two data points. Common kernels include Linear, Polynomial, and Radial Basis Function (RBF).The C hyperparameter controls the regularization strength in the SVM. A larger value of C corresponds to higher model complexity and a lower value corresponds to lower model complexity. The gamma hyperparameter controls the influence of each training example. we are searching over three hyperparameters: C, gamma, and kernel. The grid of hyperparameters to search over is defined in the param_grid dictionary, which specifies that the C hyperparameter should be searched over the values 0.1, 1, and 10, the gamma hyperparameter should be searched over the values 1, 0.1, and 0.01, and the kernel hyperparameter should be searched over the values 'linear', 'poly', and 'rbf'. The cross-validation option controls how many folds are used. Finally, the scoring parameter specifies the performance metric to use for evaluation.

After fitting the GridSearchCV object to the training data, the best_params_ attribute gives us the best combination of hyperparameters found, and the best_estimator_ attribute gives us the best model.

After tuning the hyperparameters of SVM the accuracy was the same as the SVM model which is 70.30%.

Model	Accuracy
Logistic Regression	70.63%
Tuned Logistic Regression	69.64%
SVM	70.30%
Tuned SVM	70.30%

Table 5.4: Performance of applied models

5.6 Performance of each set of Features for sarcasm extraction

First, we evaluated how well each set of feature combinations performed in terms of categorization.

5.6.1 Applied Models

Random Forest: This classifier is renowned for minimizing bias based on overfitting and imbalanced tweet classes. Sarcastic tweets were most frequently identified using the polarity method. Bouazizi and Ohtsuki (2016) classified their data as good or negative using this approach. Bouazizi and Ohtsuki (2018), used the method of polarity later in their work to categorize tweets as positive, negative, or neutral. Previous studies reported performance ranging from 45.9% to 83.1%.

Decision Tree: This is a model made up of a number of "questions" arranged in a tree-like hierarchy. Common names for the questions include condition, split, and test. In this class, the word "condition" will be used. Each leaf node and each non-leaf node each carry a condition and a prediction. Unlike botanical trees, which typically have the root (or first node) at the bottom, decision trees typically have the root (or first node) at the top.

Gradient boosting: One of the variants of the ensemble technique that combines a number of weak models for better overall performance is known as gradient boosting.It is a very powerful model and it showed the best result on the feature data.

Logistic Regression : We can get a clear idea of this model from the sentiment analysis.On the TF-IDF feature data it also showed very good results.

Adaboost : In an ensemble setting, a machine learning method known as AdaBoost, also known as adaptive boosting, is used. Decision trees with one level or one split are the most popular methods used with AdaBoost. Decision Stumps are another name for these trees. The SVM algorithm has a low F1-score of 33.8 % with a precision of 98.1% for identifying sarcastic tweets. This means that while it can accurately identify sarcastic tweets, it only captures a small fraction of them, approximately 20%. In comparison, other classifiers like k-NN and Maximum Entropy have high F1-scoresand also high accuracy, but Random Forest performs the best overall. Therefore, the results from the Random Forest classifier will be used for further experimentation and analysis. It should be noted that this effectiveness may be limited because there are fewer sarcastic tweets in a typical stream of tweets compared to the dataset used for this study.

5.6.2 Adding the features to TF-IDF and train the model

The relevance of a word in a corpus document can be ascertained through the application of TF-IDF features in detection of sarcasm. By helping Machine Learning Algorithms focus on only the key terms when detecting sarcastic tweets, this could hasten the process of finding sarcastic tweets. Adding the features to TF-IDF and training the model we get a result:

Applied model scores

label	Precision	Recall	f1-score	Support
0	68%	65%	67%	197
1	67%	70%	69%	202
accuracy			68%	399
macro avg	68%	68%	68%	399
weighted avg	68%	68%	68%	399

Table 5.5: Performance of DecisionTree model

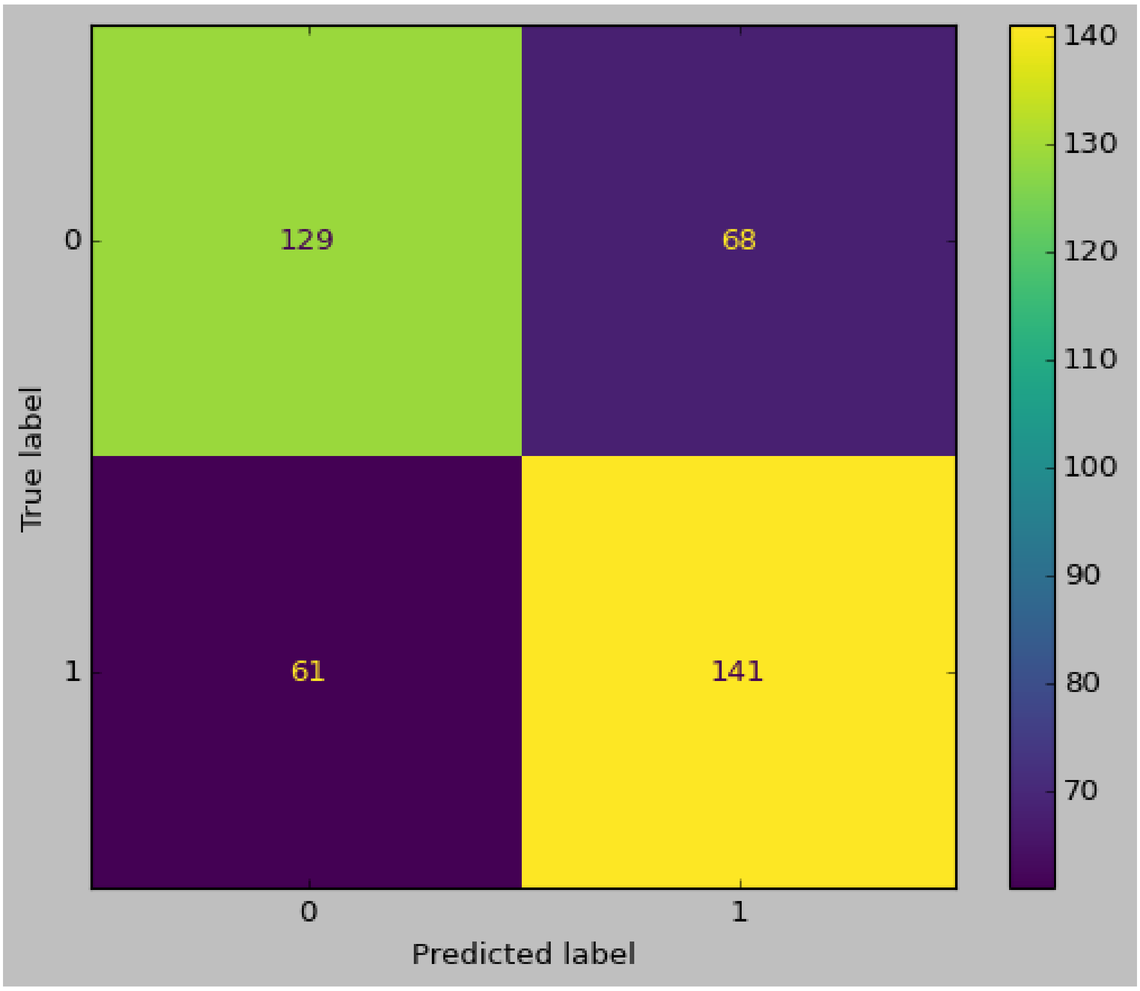


Figure 5.7: Confusion Matrix from DecisionTree model

label	Precision	Recall	f1-score	Support
0	75%	73%	74%	197
1	74%	77%	75%	202
accuracy			75%	399
macro avg	75%	75%	75%	399
weighted avg	75%	75%	75%	399

Table 5.6: Performance of RandomForest model

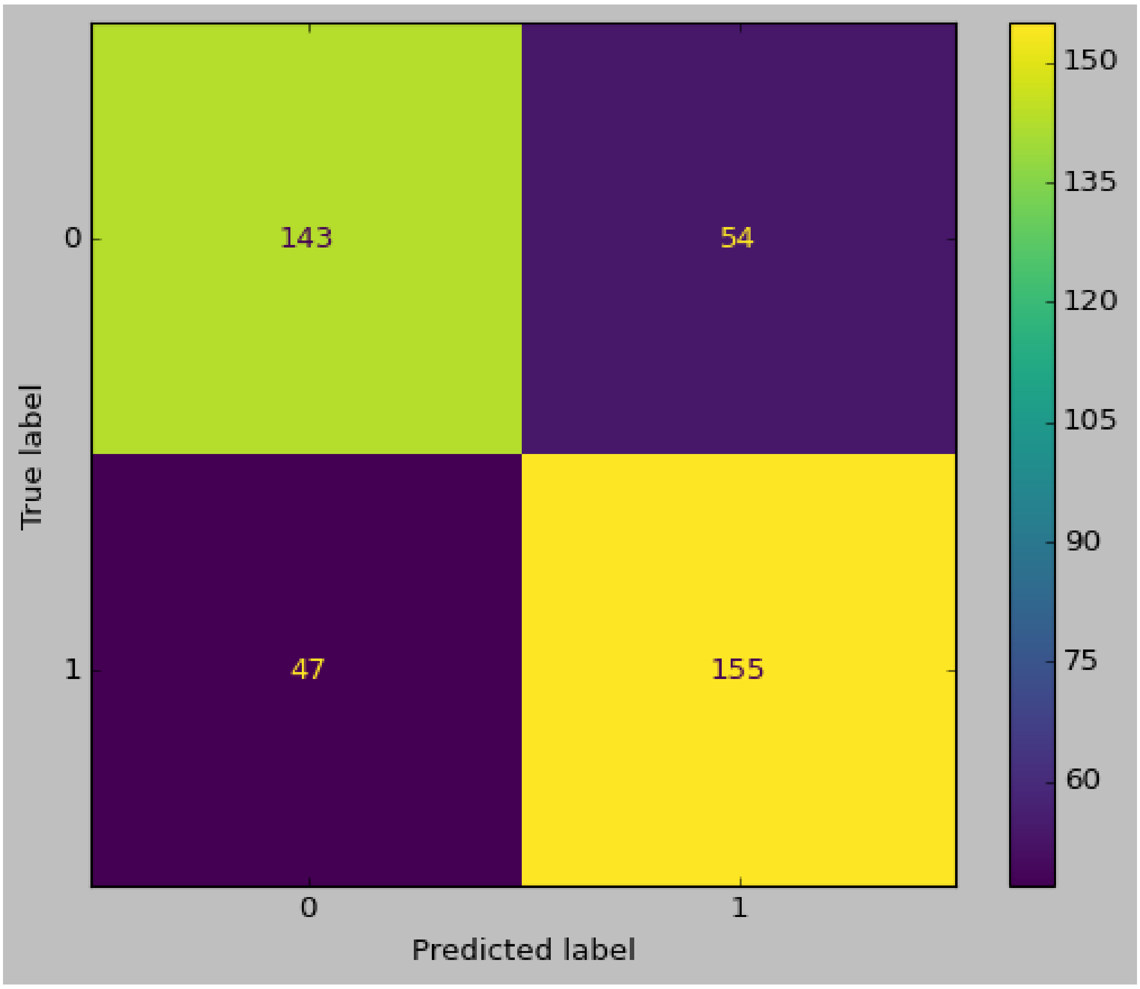


Figure 5.8: Confusion Matrix from RandomForest model

label	Precision	Recall	f1-score	Support
0	78%	68%	72%	197
1	72%	82%	77%	202
accuracy			75%	399
macro avg	75%	75%	75%	399
weighted avg	75%	75%	75%	399

Table 5.7: Performance of Adaboost model

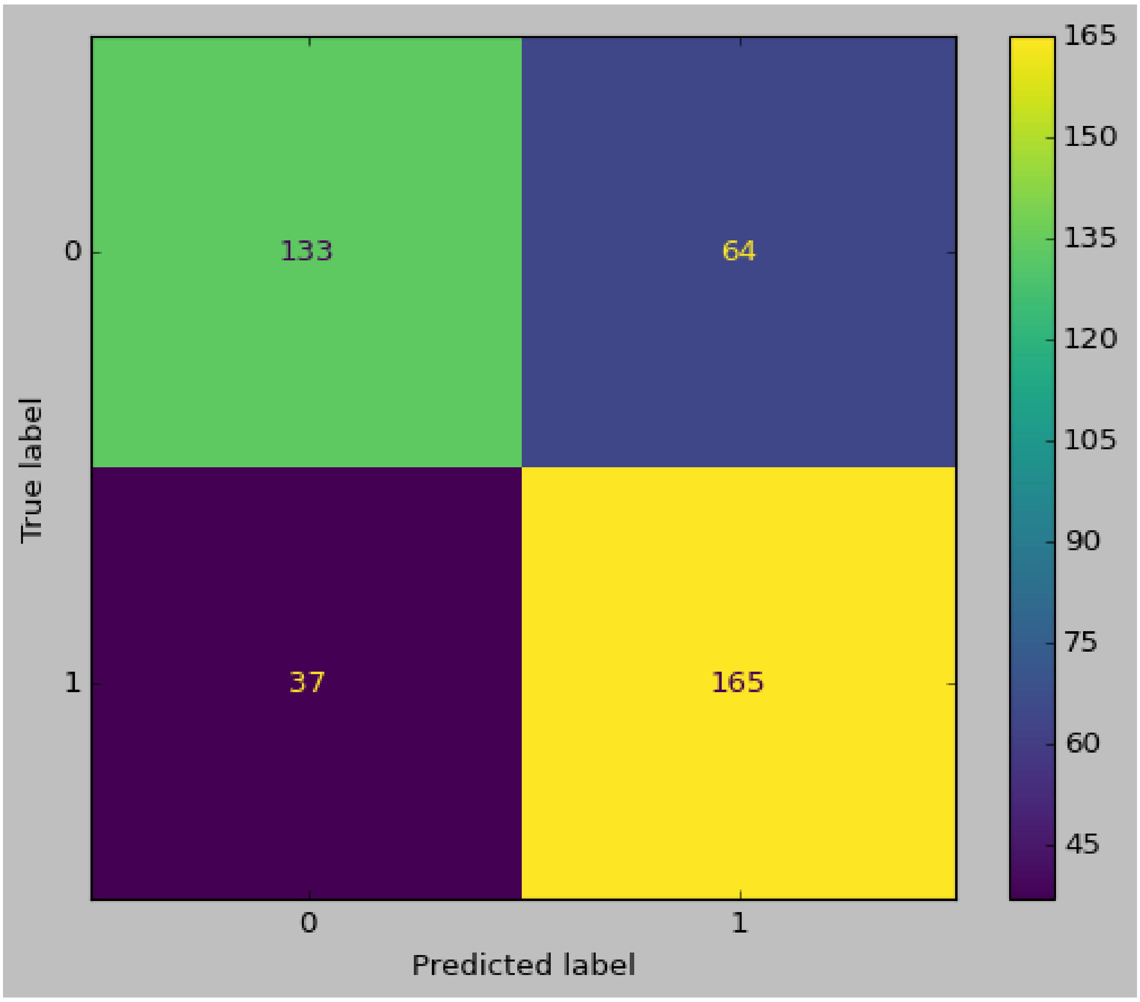


Figure 5.9: Confusion Matrix from Adaboost model

label	Precision	Recall	f1-score	Support
0	79%	68%	73%	197
1	72%	82%	77%	202
accuracy			75%	399
macro avg	76%	75%	75%	399
weighted avg	76%	75%	75%	399

Table 5.8: Performance of GradientBoosting model

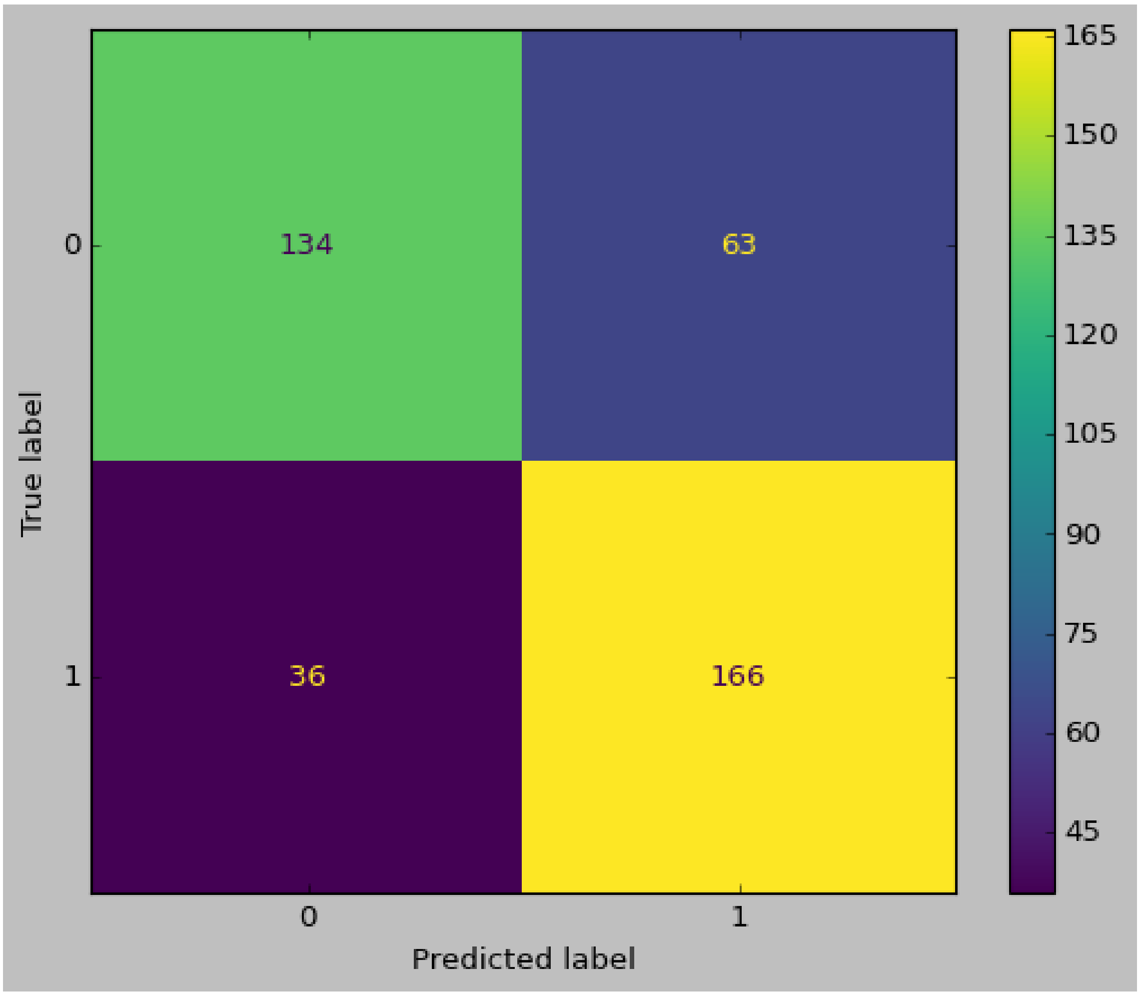


Figure 5.10: Confusion Matrix from GradientBoosting model

label	Precision	Recall	f1-score	Support
0	73%	65%	69%	197
1	69%	76%	72%	202
accuracy			71%	399
macro avg	71%	71%	71%	399
weighted avg	71%	71%	71%	399

Table 5.9: Performance of GNB model

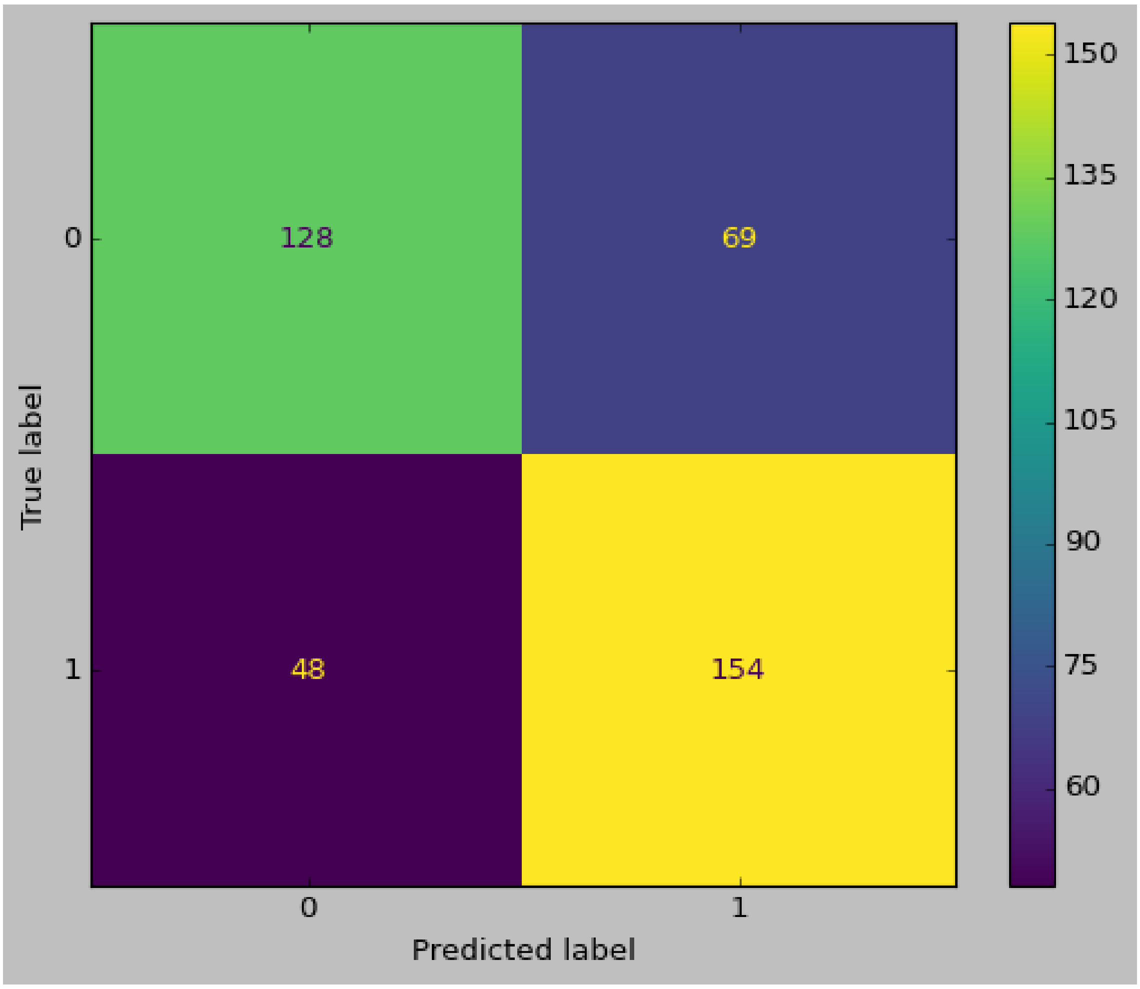


Figure 5.11: Confusion Matrix from GNB model

label	Precision	Recall	f1-score	Support
0	75%	78%	76%	197
1	77%	75%	76%	202
accuracy			76%	399
macro avg	76%	76%	76%	399
weighted avg	76%	76%	76%	399

Table 5.10: Performance of LogisticRegression model

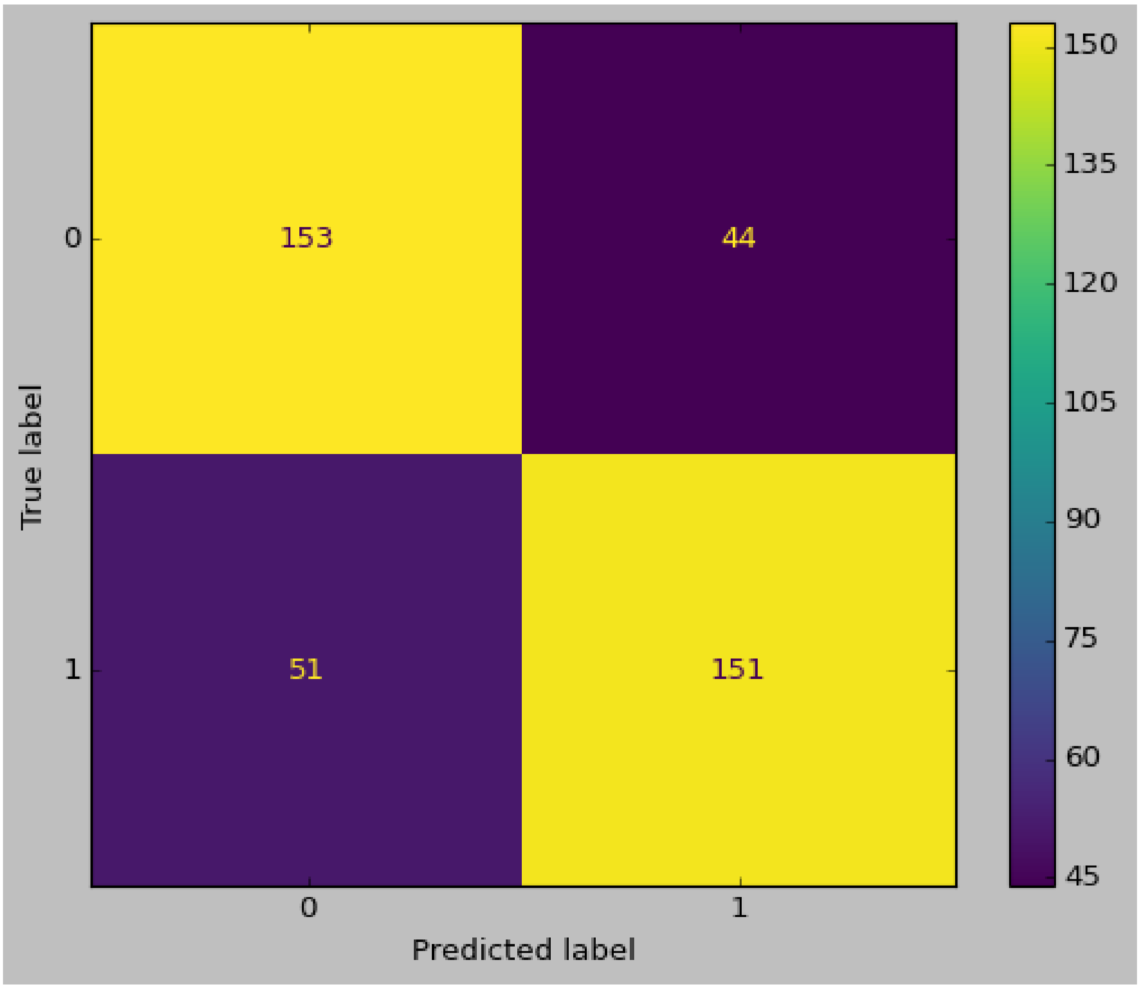


Figure 5.12: Confusion Matrix from logistic regression model

Model	Accuracy
DecisionTree	67.67%
RandomForest	74.69%
Adaboost	74.69%
GradientBoosting	75.19%
GNB	70.68%
LogisticRegression	76.19%

Table 5.11: Performance of the models on tf-idf feature dataset

5.6.3 Using Part-Of-Speech (POS) taggers and training the model

Applying the POS tagging process in complex tasks like sarcasm detection is required before performing the classification procedure to determine if the statement is labeled as sarcastic or not. Model Results on Sarcasm Labeled Dataset after applying POS Taggers are given below.

label	Precision	Recall	f1-score	Support
0	92%	95%	94%	307
1	95%	91%	93%	293
accuracy			93%	600
macro avg	93%	93%	93%	600
weighted avg	93%	93%	93%	600

Table 5.12: Performance of logistic regression model

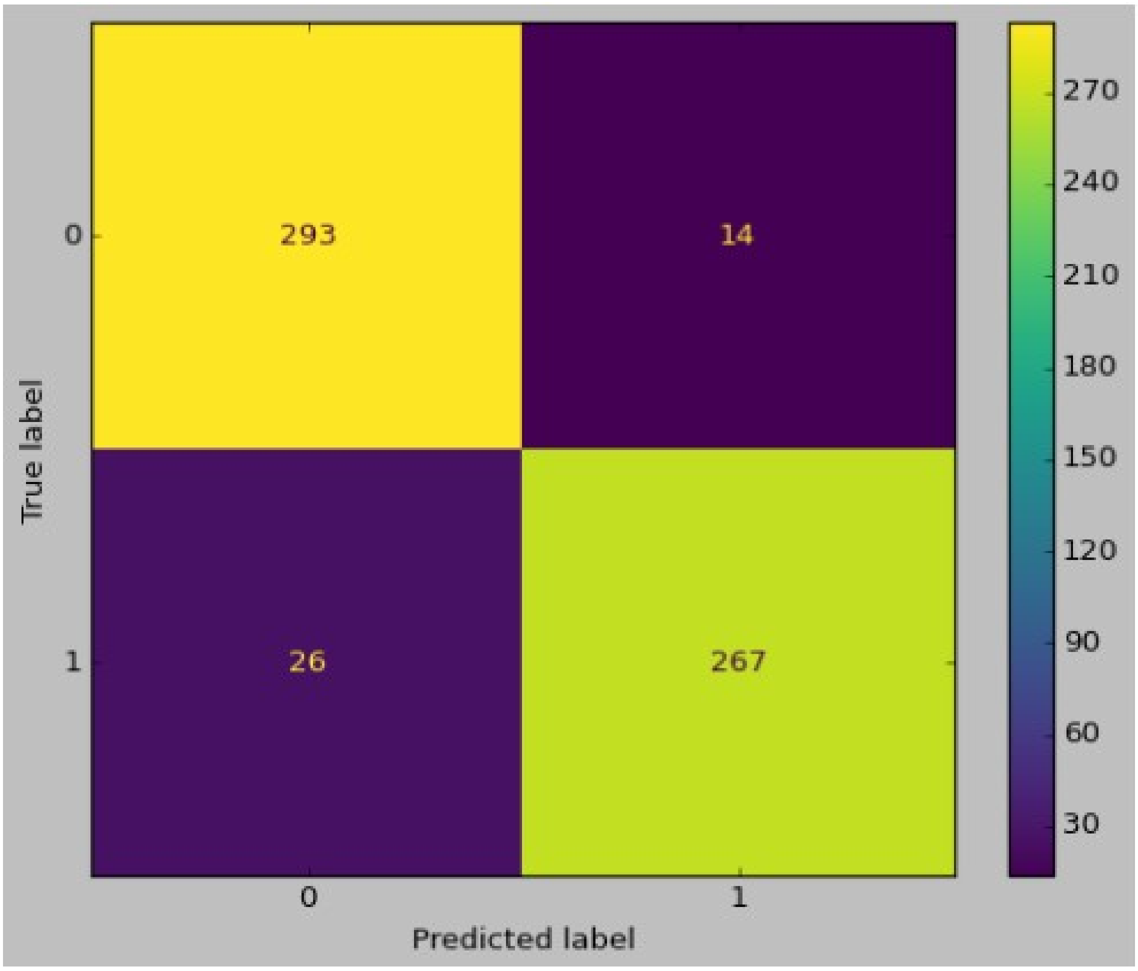


Figure 5.13: Confusion Matrix from logistic regression

5.6.4 Applying stemmed and Gram feature and training the model

These algorithms can be used to extract stemmed features from the task model during the data preprocessing step thanks to stemming. Stemming features can eliminate illogical data forms and forms of a word that are derived from them, which can affect how well a machine learning system performs.N-gram model predicts the likelihood of a specific N-gram in any language's word order. If we have a strong N-gram model, we may forecast p(w — h), which is the likelihood that the word w will appear given a his-tory of words h that contains n-1 words.After Applying stemmed and Gram feature the cross validation results will be

Applied model scores

label	Precision	Recall	f1-score	Support
0	72%	69%	70%	202
1	69%	72%	71%	197
accuracy			70%	399
macro avg	70%	70%	70%	399
weighted avg	70%	70%	70%	399

Table 5.13: Performance of DecisionTree model

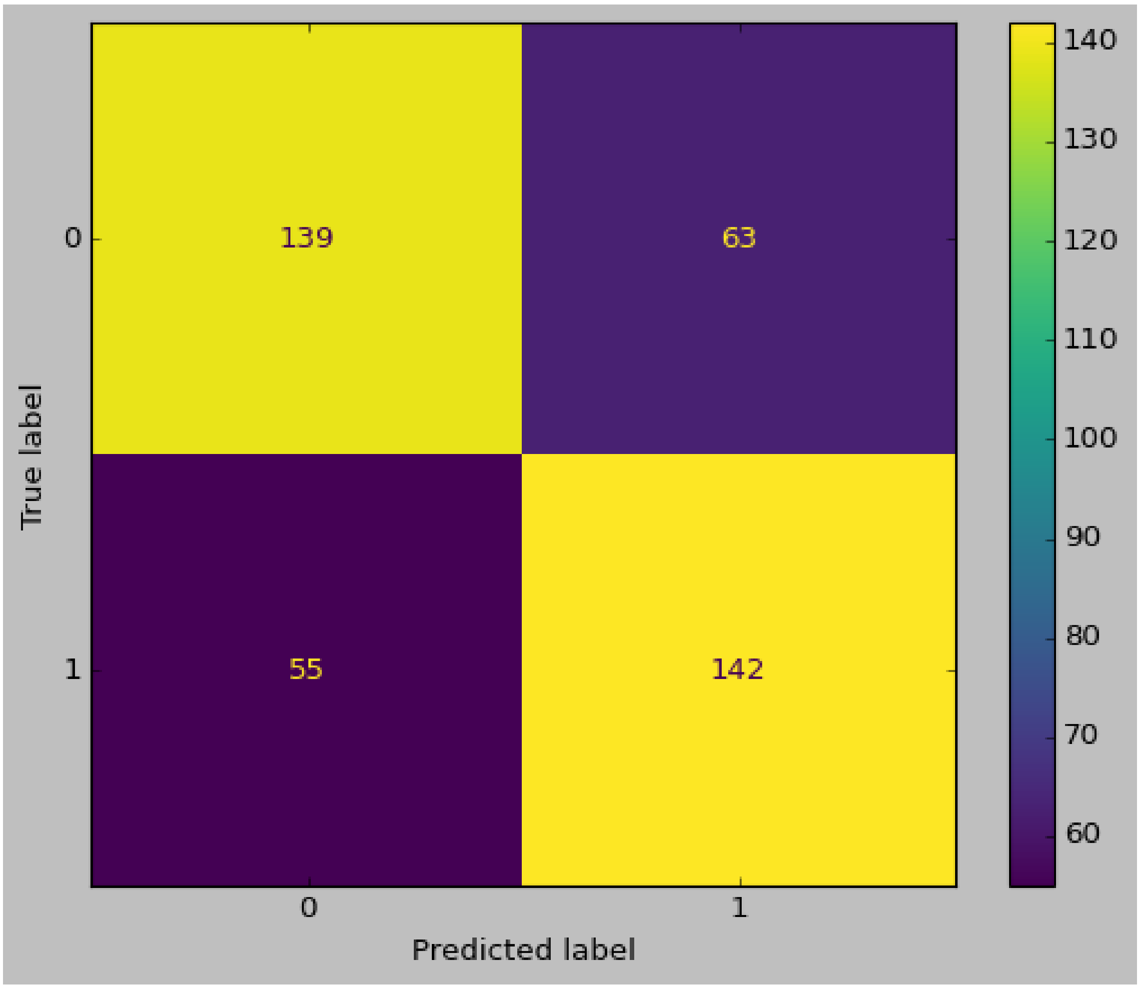


Figure 5.14: Confusion Matrix from DecisionTree model

label	Precision	Recall	f1-score	Support
0	75%	82%	78%	202
1	79%	72%	76%	197
accuracy			77%	399
macro avg	77%	77%	77%	399
weighted avg	77%	77%	77%	399

Table 5.14: Performance of RandomForest model

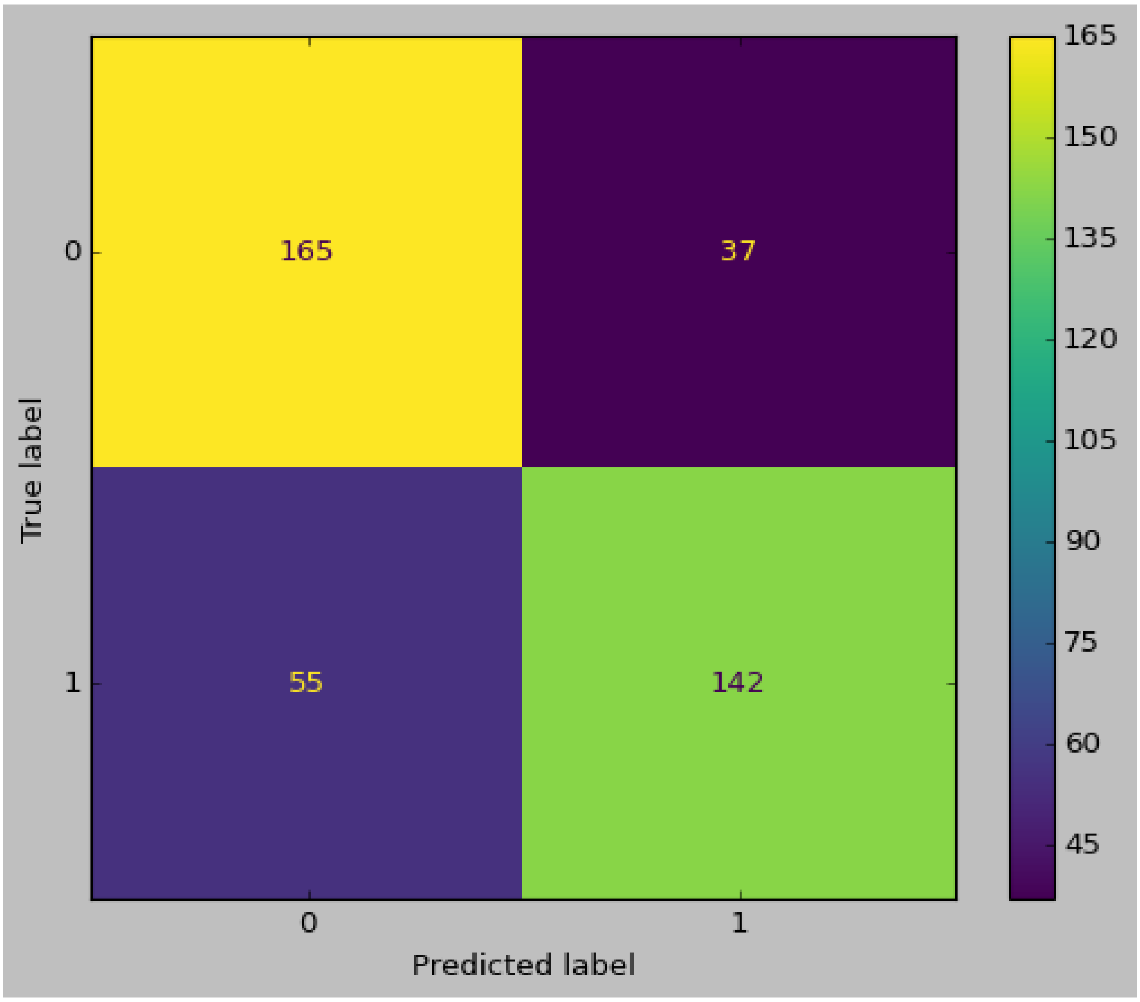


Figure 5.15: Confusion Matrix from RandomForest model

label	Precision	Recall	f1-score	Support
0	76%	62%	68%	202
1	67%	80%	73%	197
accuracy			71%	399
macro avg	72%	71%	71%	399
weighted avg	72%	71%	71%	399

Table 5.15: Performance of Adaboost model

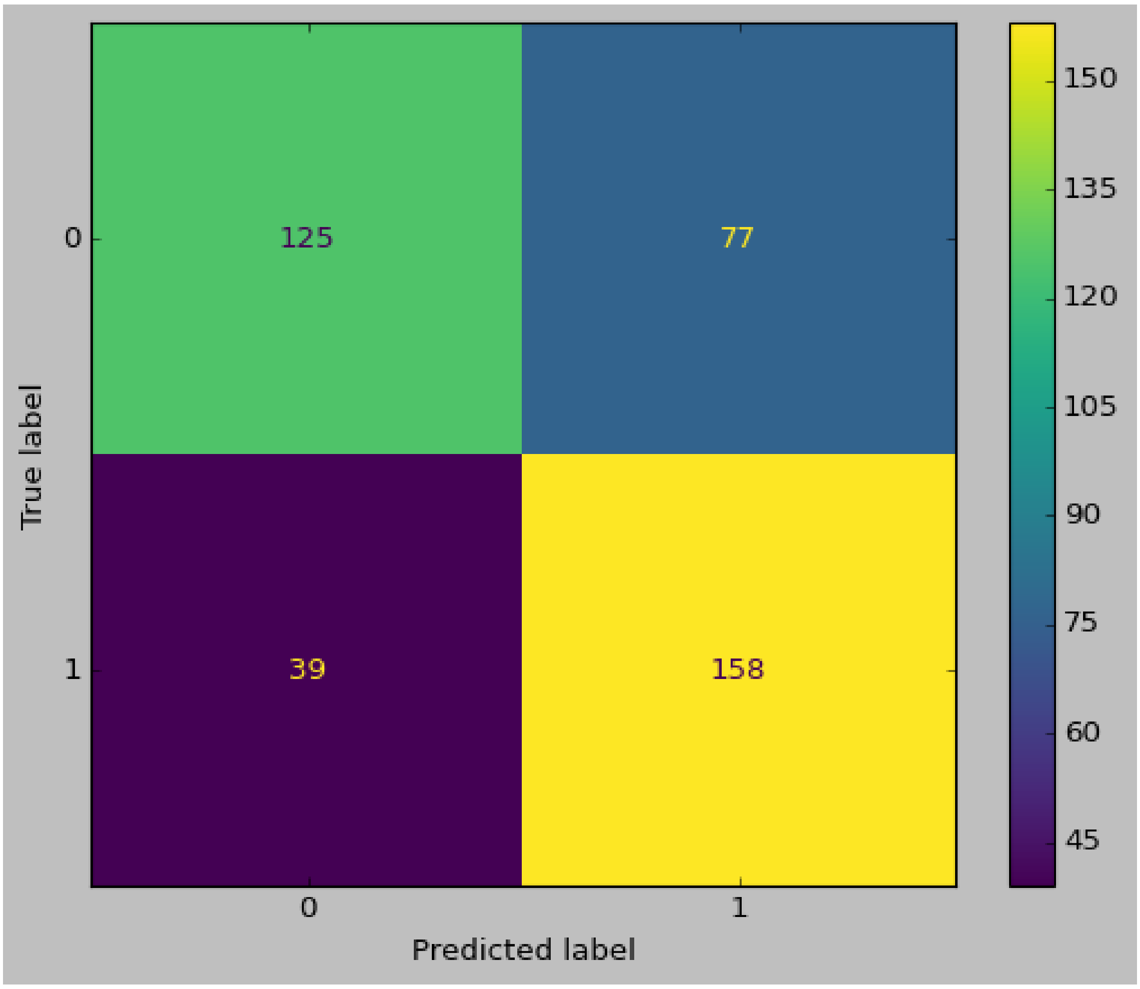


Figure 5.16: Confusion Matrix from Adaboost model

label	Precision	Recall	f1-score	Support
0	76%	75%	76%	202
1	75%	76%	76%	197
accuracy			76%	399
macro avg	76%	76%	76%	399
weighted avg	76%	76%	76%	399

Table 5.16: Performance of GradientBoosting model

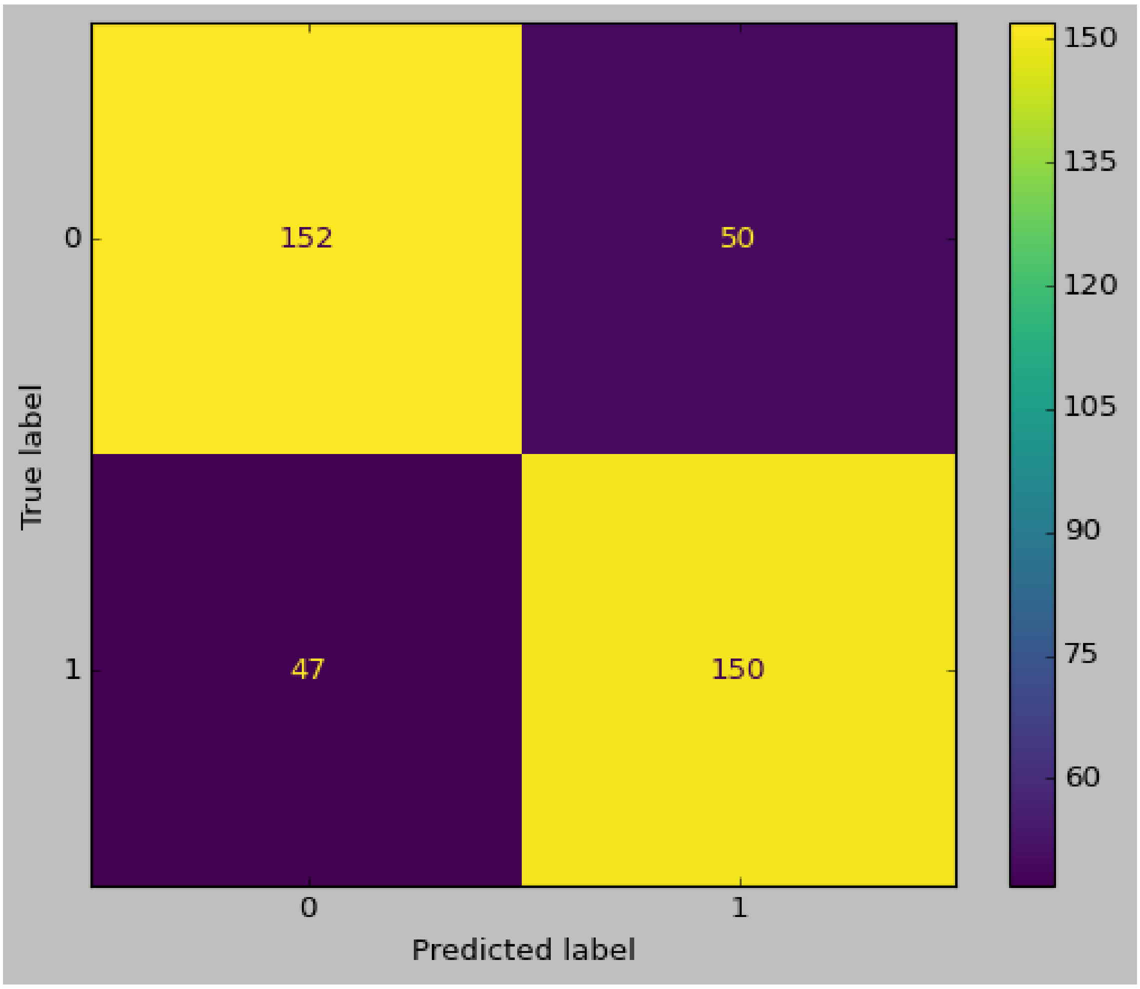


Figure 5.17: Confusion Matrix from GradientBoosting model

label	Precision	Recall	f1-score	Support
0	75%	64%	69%	202
1	68%	78%	73%	197
accuracy			71%	399
macro avg	71%	71%	71%	399
weighted avg	71%	71%	71%	399

Table 5.17: Performance of GNB model

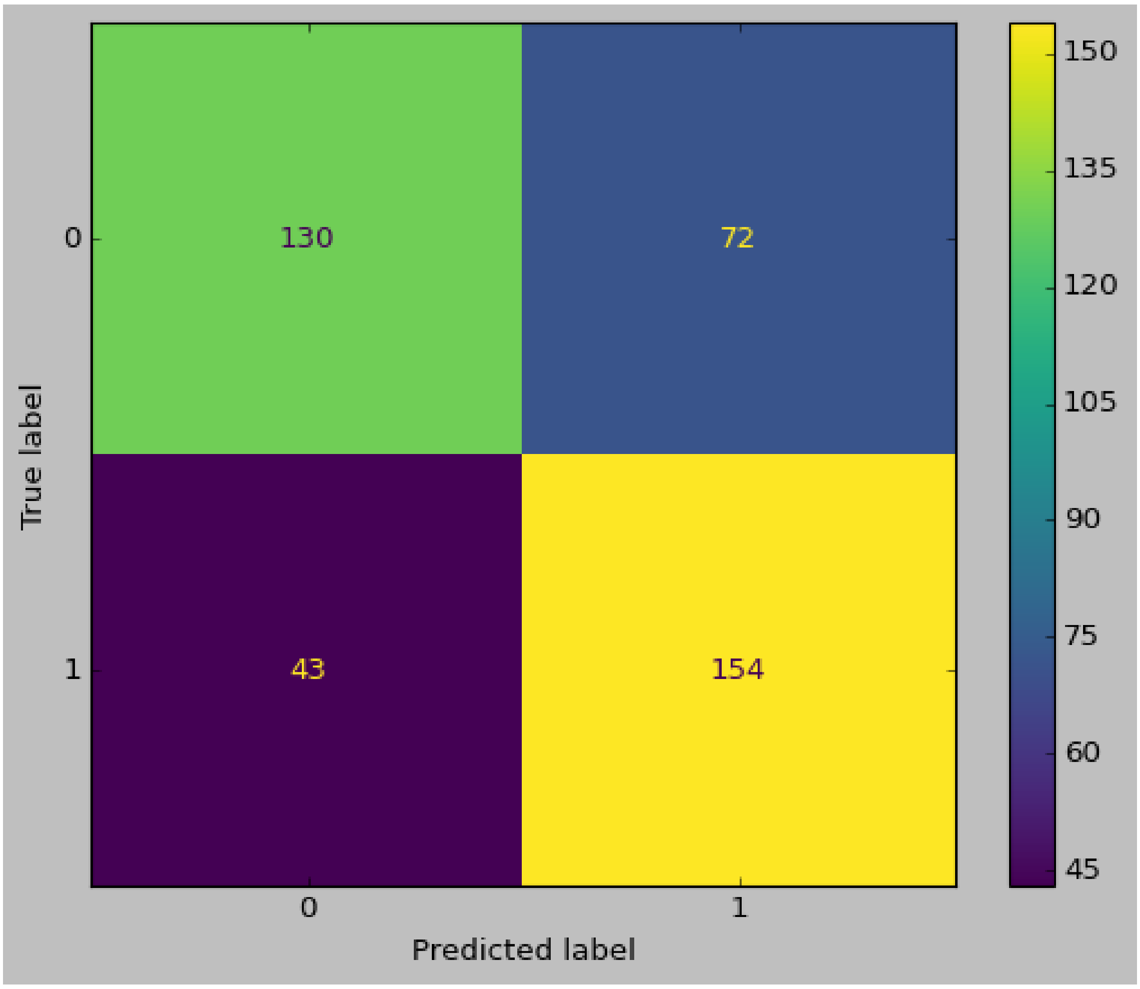


Figure 5.18: Confusion Matrix from GNB model

label	Precision	Recall	f1-score	Support
0	75%	78%	77%	202
1	76%	74%	75%	197
accuracy			76%	399
macro avg	76%	76%	76%	399
weighted avg	76%	76%	76%	399

Table 5.18: Performance of LogisticRegression model

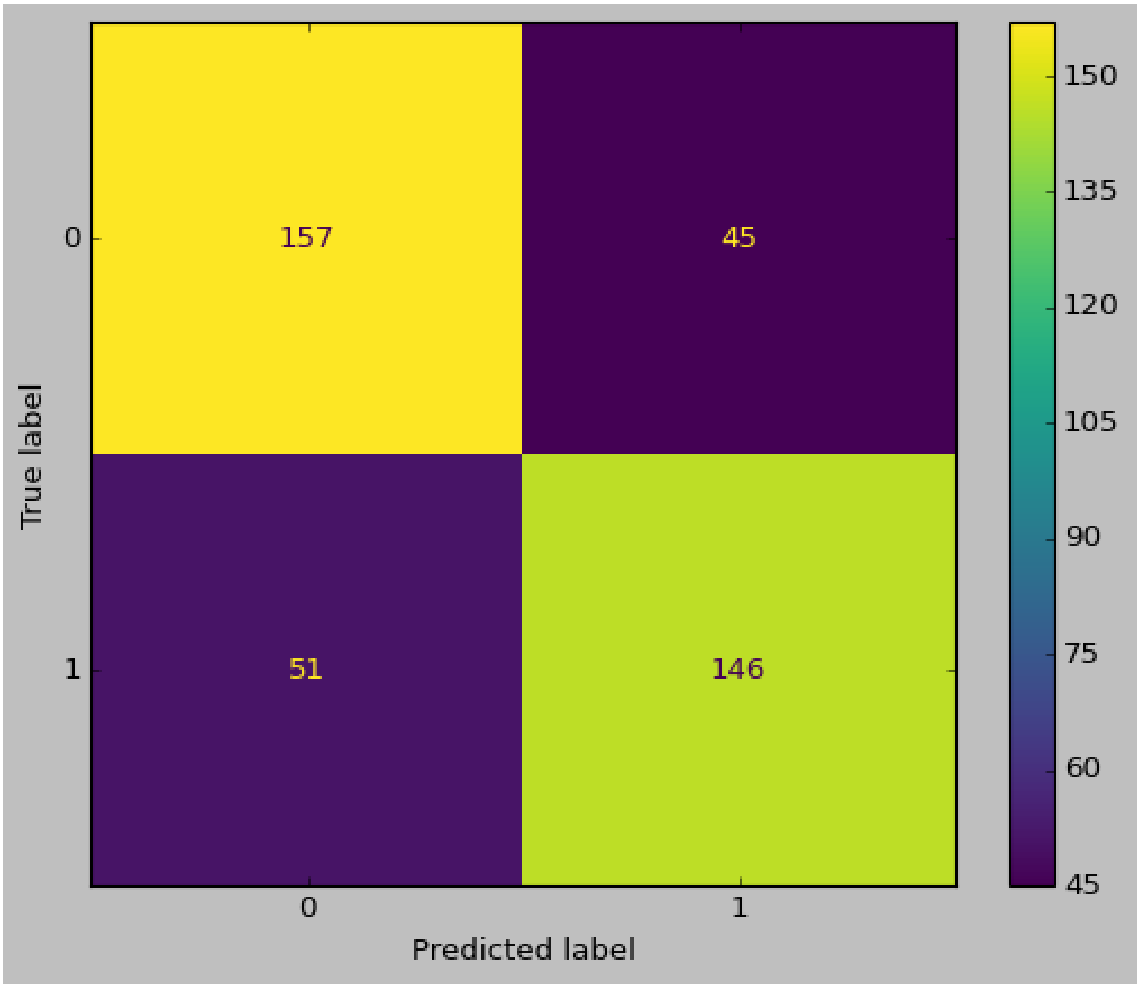


Figure 5.19: Confusion Matrix from logistic regression model

Model	Accuracy
DecisionTree	70.43%
RandomForest	76.94%
Adaboost	70.93%
GradientBoosting	75.69%
GNB	71.18%
LogisticRegression	75.94%

Table 5.19: Performance of the models after stemming and lemmetizing

5.6.5 Cross-validation

During cross-validation, we can observe that on the tf-idf features Logistic Regression shows the highest accuracy which is 76.19%, and Gradient Boosting provides the second-best score of 75.19%. The pattern-related features perform quite well. Each tweet in the training set becomes the value of as chosen, as was stated in the preceding section most comparable to itself. On the other hand, the accuracy and recall of the syntax-related features are comparatively poor. The features don't seem to be particularly good at distinguishing between sarcastic and non-sarcastic tweets when used alone. One of the factors is the small amount of these attributes in the data set. The likelihood that each feature in the training set exists is displayed in TABLE 7.1 to 7.15. Additionally, the PoS-tagger performs worse in Twitter's informal language and volume than it would in a formal text. The PoS-tagger typically classifies interjections as nouns, making them extremely challenging to distinguish. However, the near level of precision that this combination of variables provided demonstrates the requirement for such qualities in the recognition of sarcastic components. It is the percentage of all tweets that have been classified as sarcastic. Even when certain characteristics don't work well, they could nonetheless be more valuable or common than other ones. SThe prediction rates of sentiment- and punctuation-related features are higher. They are less effective than pattern-related qualities despite being more effective. Additionally, compared to precision, sentiment-based traits' accuracy is substantially lower. In other words, tweets that have been classified as sarcastic have a high prediction rate. This is comprehensible by the fact that when there is a mismatch in emotional content, sarcastic tweets are more likely to be written. They would therefore be seen as sarcastic if they were observed. However, after applying the pos-taggers on the sarcasm-labeled data, when we train the models we find a very high accuracy rate of 93% which is much higher. And after applying the stemming and Lemmatization when we train the models we find that the RandomForest model scores the ideal outcome with accuracy of 76.94%.

5.6.6 On a test set

Performance for cross-validation is unquestionably better than for classifying unknown data. However, we can observe that the feature sets with the highest merit in cross-validation are likewise the highest-merit individuals.when classifying tweets from the test set. The low existence of features relevant to syntax in the test set is another factor contributing to their low accuracy. Regarding patterns, they perform better and offer more features. Precision and accuracy have values that are extremely similar. This is due to the fact that unlike sentiment-based features, which, for example, evaluate whether particular traits linked with sarcasm are present in the tweets, patterns are created from both sardonic and non-sardonic tweets, and the proximity to these patterns is confirmed.

5.7 Overall Performances of the Approach

The features work better as a group than they do individually. When all the features are employed, Both the precision and accuracy during cross-validation are greater than 90%. The recall is less than 89% of the time. More surprisingly, the accuracy for the test set before pattern enrichment exceeded 71% with precision above 72%. This demonstrates that the various feature sets perform better when integrated. Despite the fact that the majority of the sarcastic tweets in our data set are difficult to characterize even by humans (we used the hashtag "sarcasm" to do so), the accuracy was high. The approach's potential was expanded through the enrichment process, which also improved classification accuracy. Comparing the precision to that without enrichment, it also improved. It illustrates how true sarcasm is in the majority of tweets that have been labeled as such. On the contrary, a recall has a lesser value, however, it is still an improvement over before enrichment. It demonstrates that a significant portion of the sarcastic tweets had poor classification. We assume that many of the sarcastic tweets that were not labeled as sarcastic come into this group. However, this can be improved if we include more tweets in the training set or for enrichment.

We use the technique put out by Riloff et al. [33] as well as the n-gram-based approaches as our baseline to assess the viability of our method. We define a fourth KPI in addition to the three previously mentioned ones, which is the F1 score, which is defined as follows:

$$F1 \leq (2 \times (precision \times recall) \div (precision + recall)) \tag{5.1}$$

It combines precision and recall, making it a much more dependable KPI when comparing various strategies. For the given data set, our suggested strategy notably surpasses the baseline approaches: not only does it have superior accuracy and precision, but its F1 score is also noticeably greater than the baseline approaches. The Riloff et al. [33] technique, which excels at recognizing a particular sort of sarcasm, falls short as the majority of satirical tweets don't fit in our data collection under the category of sarcasm where a positive mood contrasts with a negative circumstance. In comparison to more complex algorithms like those put out by Davidov et al. [9] or Rajadesingan et al. [12], our methodology offers competitive results despite not requiring a substantial training data set or user knowledge. Due to the lack of prior user information (as in [12]) and the lack of 5.9 million tweets to divide terms into context words and very common words[9], The two methods weren't tried out again on our data set. However, the accuracy of our suggested method is near to 74%, and our F1 score is close to 76% (on the Twitter data set).

5.8 Future Works

For the meantime, we are all aware of the variety of data when it comes to sarcasm; nevertheless, because it is not restricted by any type of specification, it is challenging for a computer to comprehend a person's feelings, let alone sarcasm. A machine must execute algorithms repeatedly in order to become used to these recurrent problems. Examining the sentiment of tweets offers an interesting window into the general public's perceptions of a particular event. Therefore, the major objective of this study is to identify sarcasm in the dynamic Twitter data. When reading reviews or complaints from consumers, understanding their genuine motivations and thinking can help after-sales services or consumer assistance become more successful.This research seeks to identify sarcasm as well as categorize feelings into good, unpleasant, and neutral attitudes. Before any inferences can be made, the data must be preprocessed using the Twitter API. Our goal's preparation involves a number of classifiers. It is evident that the different supervised algorithms are reliable and adequate for sarcasm detection.To identify sarcastic tweets as good or bad, further work with classical Machine Learning (ML) and Natural Language Processing (NLP) may be applied. This field of research may offer greater clarity for a computer to recognize sarcasm.The same method may also be used to categorize more complex kinds of sarcasm including irony, pun, banter, comedy, and so on, helping the computer comprehend better and arrive at the intended result.

Conclusion

Irony with a sophisticated twist, sarcasm, was widely used on Twitter. Identifying sarcastic tweets is crucial for text classification and has numerous ramifications. To categorize tweets that are sarcastic on Twitter, this study examined a number of machine learning techniques. In this research, we offer a method for analyzing sentiment on web-based learning as well as a method for sarcasm detection in tweets based on online learning. The various elements of the tweet are used in the suggested method for sarcasm detection. We use Part-of-Speech-Tags to identify patterns that describe the sarcasm level on Twitter. However, if we used a larger training sample, the approach might yield even better results because the patterns retrieved from the existing collection may not be entirely captured. The other features Tf-IDF and Stemming also showed a great performance in sarcasm detection. we got an accuracy 76.19% from tf-idf, 76.94% from stemmed feature. We suggested a more effective method to enhance our collection utilizing a set of trained tweets with more patterns that are sarcastic on online learning as a starting point. In the future, we will look at ways to leverage the results of the current study to enhance data on online learning as well as sentiment analysis and opinion mining performance.

Bibliography

[1] J. M. Soler, F. Cuartero, and M. Roblizo, "Twitter as a tool for predicting elections results," in Proc. IEEE/ACM ASONAM, pp.1194–1200, Aug. 2012.

[2] U. R. Hodeghatta, "Sentiment analysis of Hollywood movies on Twitter," in Proc. IEEE/ACM ASONAM, pp. 1401–1404, Aug. 2013.

[3] S. Homoceanu, M. Loster, C. Lofi, and W-T. Balke, "Will I like it? Providing product overviews based on opinion excerpts," in Proc.IEEE CEC, pp. 26–33, Sept. 2011.

[4] S. Attardo, "Irony as relevant inappropriateness," Irony in language and thought, pp. 135–174, June 2007.

[5] R. L. Brown, "The pragmatics of verbal irony," Language use and the uses of language, pp. 111–127, 1980.

[6] H. Grice, "Further notes on logic and conversation," Pragmatics: syntax and semantics, pp. 113–127, Academic Press, 1978.

[7] R. W. Gibbs and J. O? Brien., "Psychological aspects of irony understanding," Journal of Pragmatics, pp. 523–530, Dec. 1991.

[8] D. Davidov, O. Tsur, and A. Rappoport, "Semi-supervised recognition of sarcastic sentences in Twitter and Amazon," In Proc.14th Conf. on Computational Natural Language Learning, pp. 107–116, July 2010.

[9] O. Tsur, D. Davidov, and A. Rappoport. "ICWSM? A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews," in Proc. AAAI Conf. Weblogs and Social Media, pp162-?169, May 2010.

[10] D. Maynard, and M. Greenwood, "Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis," in Proc. 9th Int. Conf. Language Resources Evaluation, pp. 4238–4243, May 2014.

[11] C. Liebrecht, F. Kunneman, and A. Van Den Bosh, "The perfect solution for detecting sarcasm in tweets not," in Proc. WASSA 2013, pp. 29–37, June 2013.

[12] A. Rajadesingan, R. Zafarani, and H. Liu, "Sarcasm detection on Twitter: A behavioral modeling approach," in Proc. 18th ACM Int. Conf. Web Search Data Mining, pp. 79–106, Feb. 2015.

[13] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short text classification in Twitter to improve information filtering," in Proc. 33rd Int. ACM SIGIR Conf. Research and development in information retrieval, pp. 841–842, July 2010.

[14] B. Pang, L. Lillian, and V. Shivakumar, "Thumbs up?: Sentiment classification using machine learning techniques," in Proc. ACL-02 Conf. Empirical Methods in Natural Language Process., vol. 10, pp.79–86, July 2002.

[15] M. W. Berry, "Survey of text mining: Clustering, classification, and retrieval", 2004.

[16] M. Boia, B. Faltings, C.-C. Musat, and P. Pu, "A :) is worth a thousand words: How people attach sentiment to emoticons and words in tweets," in Proc. Int. Conf. Social Computing, pp. 345–350, Sept. 2013.

[17] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive recursive neural network for target-dependent Twitter sentiment classification," in Proc. 52nd Ann. Meeting on Assoc. for Computational Linguistics, vol. 2, pp. 49–54, June 2014.

[18] K. Manuel, K. V. Indukuri and P. R. Krishna, "Analyzing internet slang for sentiment mining," in Proc. 2nd Vaagdevi Int. Conf. Inform. Technology for Real World Problems, pp. 9–11 Dec. 2010.

[19] S. G. Shamay-Tsoory, R. Tomer, and J. Aharon-Peretz, "The neuroanatomical basis of understanding sarcasm and its relationship to social cognition," Neuropsychology, vol. 19, no. 3, pp. 288–300, May 2005.

[20] F. Jr. Sting, The meaning of irony. New York, State University of NY, 1994.

[21] C. Burfoot and T. Baldwin., "Automatic satire detection: Are you having a laugh?" in Proc. ACL-IJCNLP 2009, pp. 161–164, Aug. 2009.

[22] D. Wilson, "The pragmatics of verbal irony: Echo or pretense?," Lingua, Vol. 116, no. 10, pp. 1722–1743, Oct. 2006.

[23] S. L. Ivanko and P. M. Pexman, "Context incongruity and irony processing," Discourse Processes, vol. 35, no. 3, pp. 241–279, 2003.

[24] J. Tepperman, D. Traum, and S. S. Narayanan, "Yeah right: Sar-casm recognition for spoken dialogue systems," in Proc. InterSpeech, pp. 1838–184, Sept. 2006.

[25] F. Barbieri, H. Saggion, and F. Ronzano, "Modeling sarcasm in Twitter, a novel approach," in Proc. WASSA, pp. 50-58, June 2014.

[26] T. Veale and Y. Hao, "Detecting ironic intent in creative comparisons," in Proc. ECAI, pp. 765–770, Aug. 2010.

[27] D. Ghosh, W. Guo, and S. Muresan, "Sarcastic or not: Word embeddings to predict the literal or sarcastic meaning of words," in Proc EMNLP, pp.1003-1012, Sep. 2015.

[28] Z. Wang, Z. Wu, R. Wang, and Y. Ren, "Twitter sarcasm detection exploiting a context-based model," in Proc. Web Inform. Syst. Eng. (WISE), pp. 77–91, Nov. 2015.

[29] S. Muresan, R. Gonzalez-Ibanez, D. Ghosh, and N. Wacholder,"Identification of nonliteral language in social media: A case study on sarcasm," Journal Assoc. Inform. Sci. and Technology, Jan. 2016.

[30] E. Riloff, A. Qadir, P. Surve, L. De Silva, N. Gilbert, and R. Huang, "Sarcasm as contrast between a positive sentiment and negative situation," in Proc. Conf. Empirical Methods Natural Language Processing, pp. 704–714, Oct. 2013.

[31] S. K. Bharti, K. S. Babu, and S. K. Jena, "Parsing-based sarcasm sentiment recognition in Twitter data," in Proc. IEEE/ACM ASONAM 2015, pp. 1373-?1380, Aug. 2015.

[32] E. Fersini, F. A. Pozzi, and E. Messina, "Detecting irony and sarcasm in microblogs: The role of expressive signals and ensemble classifiers," in Proc. IEEE Data Sci. and Advanced Analytics (DSAA), pp. 1–8, Oct. 2015.

[33] A. Reyes, P. Rosso, and T. Veale, "A multidimensional approach for detecting irony in Twitter," Language Resources and Evaluation, vol. 47, no. 1, pp. 239–268, Mar. 2013.

[34] A. Reyes, P. Rosso, and D. Buscaldi, "From humor recognition to irony detection: The figurative language of social media," Data Knowledge Engineering, vol. 74, pp. 1–12, Apr. 2012.

[35] D. Bamman and N. A. Smith, "Contextualized sarcasm detection on Twitter,"in Proc. AAAI Int. Conf. on Web and Social Media (ICWSM), pp. 574–77, May 2015.

[36] A. Joshi, V. Sharma, and P. Bhattacharyya, "Harnessing context incongruity for sarcasm detection," in Proc. Annu. Meeting Assoc. Computational Linguistics, Int. Joint Conf. Natural Language Processing (ACL-IJCNLP), vol. 2, pp. 757?-762, July 2015.

[37] L. Derczynski, A. Ritter, S. Clark, and K. Bontcheva, "Twitter part-of-speech tagging for all: Overcoming sparse and noisy data," in Proc. Int. Conf. RANLP, pp. 198–206, Sept. 2013

[38] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp 20:1–27:27, Apr. 2011.

[39] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten "The WEKA data mining software: An update',' SIGKDD Explor. Newsk., vol. 11, no. 1, pp. 10–18, June 2009.