Ordinal Regression for Predicting Sentiments on Twitter using LSTM

By: - Dolly

ABSTRACT

The fundamental goal of sentiment analysis is to find and categorize any views or feelings communicated in a text. Nowadays, discussing thoughts and expressing feelings through social networking sites is widespread. Consequently, a vast amount of data is generated every day, which can be mined successfully to extract valuable information. Performing sentiment analysis on such data can be useful for producing an aggregated view of particular products. Due to the prevalence of slang and misspellings, sentiment analysis on Twitter is frequently challenging.

Additionally, we are constantly exposed to new terms, which makes it more difficult to assess and compute the sentiment compared to traditional sentiment analysis. Twitter limits a tweet's length to 140 characters. Consequently, obtaining important information from brief messages is another obstacle. Knowledge-based approaches and machine learning can significantly contribute to the sentiment analysis of tweets. The amount of data produced by people, i.e., users of a certain social site, is growing exponentially due to the changing behavior of various networking sites like Snapchat, Instagram, Twitter, etc. Every day, millions upon billions of new textual, audio, and video items are posted. This is because a particular website has millions of users. These people seek to offer their thoughts and opinions on whatever topic they desire. These brief posts are solely intended to reflect the opinion of a single user regarding a specific topic. The purpose of this paper is to determine the emotions underlying these posts. We have decided to use Twitter as our platform for this. Tweets describe the updates made to this social media platform. In this study, we investigate the views expressed by Twitter users concerning certain companies. By computing a basic sentiment score and then categorizing it as positive or negative, the corporation would be provided with critical feedback about its products from individuals worldwide.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
SVM	Support Vector Machine
POM	Proportionally odds model
CML	Cumulative Link Modelling
NLP	Natural Language Processing
TF-IDF	Term Frequency-Inverse Document Frequency
HMEWA	Hybrid Mutation-Based Earth Warm Algorithm
GEA	General Entertainment Authority
MAPE	Maximum Absolute Percentage Error
BERT	Bidirectional Transformer
OSNs	Online Social Networks
NLTK	Natural Language Toolkit

CHAPTER- 1

INTRODUCTION

This chapter gives an introduction to this twitter sentiment analysis and ordinal regression. First, section 1.1 provided a brief introduction to the research topic. Second, Section 1.2 presents the details of the twitter sentiment analysis. Afterward, the third section, 1.3, discussed the sentiment analysis, in sentiment analysis also covered two topics lexicon-based approach and machine learning, then fourth section, 1.4, discussed the sentiment analysis types. Finally, the last section describes ordinal regression in detail.

1.1 Introduction

Sentiment analysis, also called opinion mining, is a technique for extracting personal information from a textual data source to better comprehend its sentiment. Brands and companies utilise it to learn more about their customers and the goods they use, which in turn helps them make smarter choices and gives them more information to keep their customers happy. To conduct such an analysis, scientists rely either on a rule-based or automated method. In a rule-based analysis, NLP methods are applied to a custom-built collection of rules. However, automatic methods employ Machine Learning strategies rather than relying on predetermined rules. The latter method will be used to investigate our goal (Ahmed *et al.*, 2021).

Twitter combines microblogging with social networking. With its growing audience and active user base, Twitter has become the ideal medium for people to share their thoughts and ideas. Because of this, analysing emotions on Twitter has become a hot topic. By gathering and categorising public opinion through examining extensive social data, research like this might be rather useful. Because of this, it can form the foundation for further research in the area.

However, some aspects of Twitter data make sentiment analysis more challenging than other data types. As with any investigation of this nature, several obstacles exist. First, tweets are limited to 140 characters in length. Second, most tweets are written in an informal linguistic manner (like English). These tweets contain slang terms. Due to these issues, research is being performed on this subject.

Using Support Vector Machine, Multinomial Logistic Regression and Random Forest machine learning algorithms, this study primarily focuses on the polarity of tweets and divides the sentiment into five groups.

Twitter is one of the largest and most varied collections of user-generated material, with around 200 million users posting 400 million daily tweets. There has been a recent explosion in the popularity of microblogging platforms like Twitter. As a result, businesses and public partnerships are looking for ways to mine Twitter for feedback about their offerings. Two essential methods used in sentiment analysis are symbolic procedure (also known as information base methodology) and AI strategies. A large dataset of predefined emotions is necessary to characterize emotions using an information database approach. An AI method for categorizing views involves using a preexisting data set to train a sentiment classifier. These approaches are used in the analysis of Twitter user feedback. Using a training set, this approach develops an emotion classifier. We present a method that involves sorting tweets into classes using a variety of artificial intelligence techniques after they have been preprocessed, having highlights extracted and having a score and balancing plan developed. (Nennuri *et al.*, 2021).

1.2 Twitter Sentiment Analysis

One application of text mining is Sentiment Analysis, which attempts to determine a text's underlying emotional tone, such as a tweet. You may share any honest emotion or viewpoint on Twitter. The problem with sentiment analysis is that it would be unnatural for any algorithm to promise 100% accuracy or prediction. Problems with the sentence and document levels may affect the sentiment analysis results on Twitter data. Searching for positive and negative phrases in a sentence, for example, is inefficient since the flavor of a text block is largely reliant on its surrounding context. There are 1.3 billion people on Twitter, with 330 million actively using the platform every month and 145 million using it daily. Since Twitter is the most understandable source of real-time public discussions, its data may be useful for gauging customer sentiment as individuals and markets respond to corporate actions.

1.3 Sentiment Analysis

Both sentiment analysis and opinion mining fall under the text mining umbrella. It's a method for anticipating how people will feel about a certain issue. It appears in a wide variety of contexts, including online reviews, social media, and news articles. Market research and governmental policymaking are two other applications for sentiment analysis findings. Many businesses rely heavily on customer feedback since it improves service quality and boosts deliveries, not to mention it helps shape policy. The primary goal of sentiment analysis is to ascertain whether or not a given text exhibits a positive or negative attitude. Conversations, forums, and weblogs are all great sources for the data collected via sentiment analysis. Sentiment analysis is becoming more common due to social media's growing significance. To gauge public reaction to its products and policies, a company may commission polls and surveys of its customers. Consumers compared prices and researched products using consumer sentiment analysis. Once upon a time, marketers would employ sentiment analysis to learn more about their target market and their experiences with their products and services. One reason is that so many websites are hosted on the web today, and that number is only expected to grow, making it impossible to collect opinion data from them all. Another is that there aren't enough resources to do so. The need for a standardized method to get the same results is evident. In addition, the internet's text corpora provide useless and useful data that might be used in the study's analysis. There is often a fine distinction between these two types of data, which creates extra work for analysts. Human readers have a hard time sifting through online content to get what they need and then summarising what they've found (Hamza and Gupta, 2022).

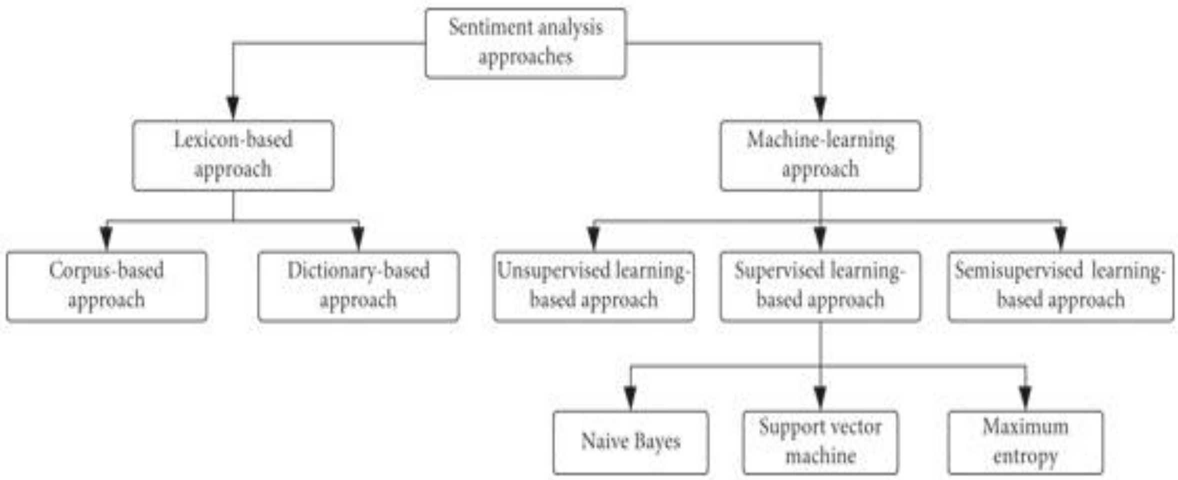

Figure 1.1: Sentiment Analysis

1.3.1 Lexicon based approach

A lexicon uses a vocabulary containing negative and positive phrases to determine the polarity of a viewpoint. In the book, the number of optimistic and pessimistic terms are explored. If the text is more favorable, it will be granted a positive score. A text receives a bad score if it contains many pessimistic terms. The score is neutral if there are about the same number of negative & positive phrases in text. A dictionary of opinion, comprising both positive and negative opinions, is being created to definitively identify whether a given the word has a good or bad connotation. There is no single best approach to dictionary building or compilation (Medhat, Hassan and Korashy, 2014).

- **Dictionary-based approach:** Gathering a few words of opinion by predetermined standards is done manually. Then, the terms are looked up in dictionaries and corpora like WordNet to find their antonyms and synonyms, and the findings are added to the set. The lexicon will shrink gradually until no more additions can take place. The dictionary scale, which determines the degree to which a feeling is intense, must be taken into account while using this approach, which is an annoyance. This strategy is flawed, particularly given the growing size of the lexicon(Jain and Dandannavar, 2017).
- **Corpus-based approach:** They rely on enormous companies as their primary source of syntactic and semantic opinion patterns. The words that are generated are dependent on the context. Thus the system requires a huge dataset with labels (Jain and Dandannavar, 2017).

1.3.2 Machine learning

Algorithms, which are collections of mathematical procedures used to determine links between variables, constitute backbone of machine learning techniques that are currently available. In this work, we will present the training process and evaluate an algorithm to determine whether or not a breast tissue sample contains cancerous cells based on the characteristics of sample itself. While the specifics of how various algorithms work may vary, there is consistent pattern in their development. For all their apparent obscurity, machine learning (ML) algorithms often bear more than a passing resemblance to more familiar forms of statistical analysis. Given their similarities, the line between statistical and machine learning methods can be difficult to draw. One method for distinguishing different techniques is examining their basic objectives. The objective of statistical techniques is to make inferences, more especially, to draw inferences about populations or to gain scientific insights from data acquired from a sample representative of that group. The requirement to make inferences about the correlations between variables motivates the use of several statistical processes, including logistics and linear regression, which can generate predictions for new data. To develop a model that captures the association between clinical characteristics and mortality following organ transplant surgery, for instance, it is necessary to understand elements that distinguish low mortality risk from high mortality risk. By doing so, we can create interventions to boost success rates and lower mortality.

Therefore, understanding the connections between the different variables should be the goal of any statistical conclusion.

In contrast, the fundamental concern in machine learning is an accurate prediction, the "what" rather than the "how." For instance, if the prediction is accurate in image recognition, the relationship between individual characteristics (pixels) and conclusion is of minimal importance. Because the link between many inputs, including pixels in a video or image and geo-location, is generally complex and non-linear, this is an essential component of machine learning methodologies. When there are a large number of predictors, each of which only adds a small amount to the model, and when the interactions between the predictors and outcomes are non-linear, it is extremely difficult to characterize the links between the predictors and outcomes in a rational fashion (Sidey-Gibbons and Sidey-Gibbons, 2019).

Supervised learning: In a nutshell, labeled training manuals are essential to the success of supervised learning. Supervised learning is a powerful classification technology that has shown highly encouraging results when used for opinion classification. Sentiment analysis typically employs the supervised classification methods of SVM (Support Vector Machine), NB (Naïve Bayes) and ME (Maximum Entropy) (Aydogan and Ali Akcayol, 2016).

- **Naïve Bayes**
- This probabilistic machine learning model is used to solve classification issues, and Bayes theorem is the model's underlying theoretical framework. (Gandhi, 2018). Below is the general Bayes equation.

$$P(A|B) = \frac{P(\mathrm{B}|\mathrm{A})P(B)}{P(B)}$$

If B has already happened, then it is possible to calculate the likelihood of A happening. Evidence (B) and a working hypothesis (x) It is assumed that the features do not interact with one another in any way. Naive Bayes classifiers come in a variety of flavours, including the multinomial, the Bernoulli, and the Gaussian. The more comprehensive Statistical n- Gram Modeling can be derived from this classifier. Compared to previous modeling methods, this one is superior for classifying texts. This method takes into account the relationships between neighboring words.

- **Support vector machine**

Because it provides a good method for text classification, SVM (Support Vector Machine) is a discriminative classifier that is taken into consideration. It is a technique of classification that is used in statistics. Several nonlinear mappings are utilised within the SVM to transform the input (actually-valued) feature vectors into a higher-dimensional feature space. The concept of structural hazard minimization is the basis for the development of SVMs. Finding a hypothesis (h) with the lowest error probability is the goal of structural risk minimization. On the other hand, traditional pattern recognition learning algorithms are predicated on empirical risk minimization, an approach that aims to maximize the efficiency of the training data. Structural risk minimization aims to identify a working hypothesis (h) with the lowest error rate. When computing the hyper aircraft to split the statistics points, such as Education and SVM, problems with quadratic optimization arise. Because its class complexity is not dependent on the function space's dimensionality, SVM can study a greater number of pattern combinations and can scale more effectively. SVM has ability to replace the teaching styles dynamically each time there is a new sample throughout the category (Mohammad and Kiritchenko, 2015).

- **Maximum entropy**

The weights constitute the defining characteristic of maximum entropy. It is also known as the exponential classifier due to the fact that it functions by first extracting some features from the input, combining those features linearly, and then applying the result as the sum of the exponents.

Unsupervised Learning: In contrast to the supervised learning technique, this technique does not involve the creation of any predefined labels for the classes. Compared to supervised learning techniques, these methods are more difficult to implement and require more time, yet they yield fairly similar results to those methods. Unsupervised learning is a method that can be accomplished through the use of neural networks (Jasneet Kaur, 2018).

Semi-supervised:

Because it employs both labelled and unlabeled data, semi-supervised learning can be regarded as a hybridization of unsupervised and supervised approaches that have been examined up to this point. This is because semi-supervised learning uses labelled and unlabeled data, which is why this result is. This is because semi-supervised learning makes use of both labelled & unlabeled data. Consequently, it is a mode of education that falls somewhere in the middle of "learning without supervision" and "learning with supervision." In the real world, labelled data may be few in a number of scenarios, whereas unlabeled data are plentiful; as a result, semi-supervised learning is effective in these kinds of circumstances (Mohammed, Khan and Bashie, 2016). The result of any semi-supervised learning model should be a prediction that is more accurate than the one that could be made by utilising the model's labelled data alone. This is the ultimate goal of any semi-supervised learning model. Semi-supervised learning has applications in many different areas, such as identifying fraudulent activity, labelling data, categorization of text, and machine translation. (Sarker *et al.*, 2020).

1.4 Sentiment Analysis types

Sentiment analysis representations focus on polarization, emotional state, emotions, urgency, and even objectives. Dependent on how a person wants to interpret purchaser feedback and questions, you can outline and tailor your groupings to meet your sentiment study needs. In

the intervening time, here are the most widespread categories of sentiment breakdown

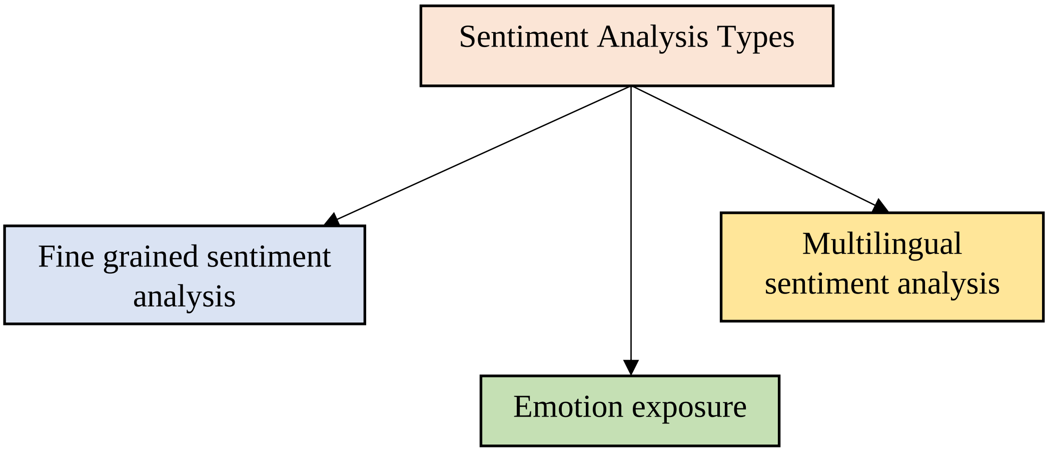

Figure 1.2: Sentiment Analysis Types

Fine-grained Sentiment Analysis

If divergence is important to corporate, we might study expanding our polarity groupings to comprise:

- Very optimistic
- Confident
- Nonaligned
- Destructive
- Very destructive

This is frequently raised to as fine-grained sentiment study and could be used to take full ratings in an appraisal, for example:

Very Constructive = 5 stars Very

Destructive = 1 star

Emotion exposure

These sentiments aim to detect happiness, hindrance, anger, grief, and so on. Many emotion recognition systems use dictionaries or complex machine learning systems. One problem with using a monolingual dictionary is that people direct emotions in diverse ways. Some arguments classically express irritation, which might also express joy (Katsurai and Satoh, 2016).

Multilingual sentiment analysis

This type of analysis of the sentiment can be hard. It contains a lot of pre-processing and properties. Most of these properties are available operationally, while others need to be produced, such as translated quantities or noise discovery processes. Still, we need to distinguish then and also how to use them. Otherwise, it could detect morphological texts automatically with language classifier and then train a tradition sentiment analysis prototype to classify manuscripts in any language (You *et al.*, 2015).

1.5 Ordinal Regression

In practise, there are numerous approaches to ordinal classification. This section will investigate the many approaches that can be utilised for the same purpose. Traditional approaches, binary decomposition models, and threshold models are the primary categories that have been distinguished within this methodologies. Ordinal problems are typically broken down into more manageable standard pattern recognition problems using conventional approaches. Methods such as error-weighted categorization and regression are included here. The process of binary decomposition involves dividing the target classes into a number of different binary variables, the values of which are determined by using different classification models. This class includes both the strategy that uses neural networks and methods that use a large number of models. The umbrella term "threshold models" encompasses a wide variety of approaches to the ordinal classification process. Figure 3 is a pictorial representation of a hierarchical classification of the approaches that were discussed in the previous section. Following are some sections that go into further detail regarding each of these approaches.

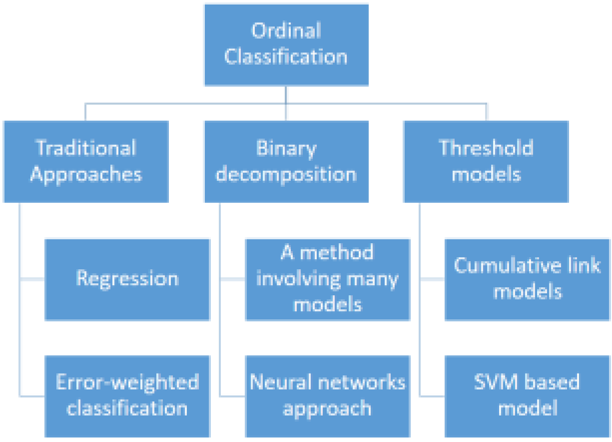

Figure 1.3: Ordinal regression

Traditional approach

In order to accomplish this, an ordinal problem must be converted into a standardised problem that has a predetermined answer. There are primarily two approaches that are conventional in their character. They are:

- **Regression**

Classification of the data is then accomplished using conventional regression-based methods and algorithms, given that data labels are now actual values. However, there are several drawbacks associated with using this method. To begin, the distance that exists between classes in ordinal data is unknown, which makes it challenging to transform ordinal numbers into real-valued entities. This is one of the reasons why. Second, the utilisation of these genuine values could potentially hinder the execution of the regression process. Finally, the position of a label relative to others is less of a concern to regression algorithms than is the label's absolute weight (Sánchez-Monedero *et al.*, 2013).

- **Error weighted classification**

The ordinal scale is used in this method to quantify the mistaken classification error. (Tu and Lin, 2010) contrasts with Kotsiantis's absolute approach to the cost variable by making use of concept of the relative distance between actual and expected classes. Each step of the procedure is represented by a single entry in a cost values matrix, M[i][j], where the entry indicates the expense associated with the jth step. Multiple options exist for determining the type of cost variable to use. (Lin and Li, 2012) gives the price variable an asymmetrical Gaussian distribution. The fact that there is no way to determine in advance which form of cost variable would work best for a particular classification problem is most significant drawback of error-weighted classification method.

Binary decomposition

The final ordinal target categories can also be divided into several binary variables, which is another approach to completing ordinal regression. Classification models will be used to assign values to these variables.

- **A method involving many models**

Using ordinal categorization, one can determine if a group is higher or lower than another. There is no doubt that we are dealing with a case of a two-way classification dilemma. If you have a pattern y and you want to know if it exceeds a certain level l, you can do so. If we assume that there are n ordinal categories and that each input must fall into one of those n categories, then we may use the above binary classification on each of those n categories to make an educated guess as to which target category the input may fall into. The input's categorical label can be established with high accuracy when the results from all n classes are aggregated. Each of the n target labels is put through a binary classifier in this approach. The probability output values of binary classifiers are utilized to make ultimate conclusion about the target.

- **A neural network approach**

A typical neural network architecture consists of a group of input neurons, a layer or two of hidden neurons, and a set of output neurons. One node with a value of zero or one represents the output for a broad category of binary classification tasks. Nevertheless, in order to solve a problem with n output classes, a neural network with n nodes in output layer is utilised. If the intended output value is strictly categorical, the standard procedure for training a neural network can be employed. As this is an ordinal problem, we need to think about how the n output nodes are connected to one another.

Threshold modelling

Threshold-based techniques are those that base their models on this premise (Verwaeren, Waegeman and De Baets, 2012). When dealing with ordinal data, threshold modelling is the most typical method used. Traditional modelling approaches are improved upon by threshold methods. Threshold models are those in which the separation between classes is not fixed at the outset but rather determined over a series of iterations, which is the fundamental difference between the two modelling systems. Modeling at the threshold is analogous to modelling at the 0-1 decomposition level. Threshold modelling diverges primarily in that it employs a single mapping scheme within a specified interval for each class.

- **Cumulative link models**

Probabilistic study of thresholds in relation to the ordinal scale is at the heart of the POM (Proportionally odds model) that is part of the CML (Cumulative Link Modelling) method. The formal model for this procedure is the generalized ordered partial proportional odds models. In the existing set of variables, Peterson only applies probabilistic ideas to a select few. The term "Partial Proportion Modeling" describes this method.

- **SVM based model**

When it comes to ordinal-based classification, SVM is most used classifier. For two reasons, this is the case. First, SVM has a high value for accuracy and good general performance. Furthermore, threshold-based structures can be readily included in the SVM structure. In this studies, HerbRich developed a SVM-based technique which takes feature pairs into consideration and creates a brand-new dataset. This new dataset contains all of values of x d ij = xi xj and yij = O(yi) O(yj), where yi, yj Set of classes.

1.6 Dissertation outline

Chapter 1 explains how ordinal regression and twitter sentiment analysis may be performed with the help of machine learning. Sentiment analysis, Twitter sentiment, and ordinal regression are just some of the topics we cover in this thoroughly annotated compendium.

Chapter 2 describes previous related work, grouped into 3 categories: corpus-based approaches, lexicon-based approaches and combined approaches in literature review.

We have a data set for sentiment analysis, and Chapter 3 shows how it's organised for the Twitter network.

Filtering and minimising noise in the data set are covered in detail in Chapter 4.

The concepts of Natural Language Processing (NLP) and data preparation for a machine-learning algorithm are introduced in Chapter 5.

CHAPTER- 2

LITERATURE REVIEW

This chapter discusses the background of twitter sentiment analysis and ordinal regression using machine learning. This chapter provides the related work for Sentiment Analysis for Tweets using Machine Learning. After this, the research gap is discussed along with the final summary of the chapter.

2.1 Background

The popularity of microblogging platforms has increased in recent years. People spend a lot of time on these sites, discussing and sharing their thoughts on current events, their future goals, and products and services they have used, both extensively and briefly. Businesses and organizations are investing time and resources into analyzing the content of social media postings and tweets from all around the world. This data could be utilized to gain insights into current topics that people are interested in, as well as to determine that services and products people are discussing and whether they are interested in them or not. Twitter has been incredibly popular due to its unique features, such as its emphasis on free expression, its 140-character message restriction, and the widespread use of the hashtag. On Twitter, the standard tweet length is 140 characters, or around 13 to 15 words (Jain and Dandannavar, 2017). As a consequence, conducting sentiment analysis (SA) on tweets is a fascinating endeavor because it can readily yield insightful information. But it has its own difficulties for analysis, such as informal language, frequent misspellings, a wide diversity of vocabulary, and usage of symbolic terms, all of which present substantial obstacles for processing and interpreting data in tweets (Liu *et al.*, 2012).

Sentiment Analysis refers to the practise of extracting characteristics from people's opinions in order to gather and categorise data based on their feelings, ideas, and expressions. It is the process of classifying vast quantities of information into distinct sentiment classifications, which might be positive, neutral or negative. In order to get the best outcomes, many people nowadays are turning to ML (machine learning) approaches and algorithms. These methods are effective because they can accurately determine the sentiment of any given phrase, as well as because they can learn from their analyses in past. Given the abundance of information available online, the authors proposed a method that involves gathering datasets, preprocessing

Twitter data, and utilising feature extraction techniques, then employing multiple ML algorithms to classify tweets into distinct categories.

Sentiment analysis is a subfield of NLP that attempts to determine the underlying tone, opinion, or emotional state of a human-written text. Social media, customer reviews, and weblogs are the primary sources of text. Despite the broadness of the phrase, the study of public opinion accomplishes a variety of goals. Some examples include extracting the opinion on a product from a review or recognising sentiment polarity in communication, as well as determining the position on a target or issue, such as "Climate Change." Multiple language levels are used for sentiment analysis. Standard ones include document, sentence and aspect level. In addition to its obvious use in the realms of commerce, marketing, and politics, sentiment analysis has many other practical uses. For many industries, preparing for the future relies heavily on understanding public opinion (Appel *et al.*, 2016). Success of a new product in the market can be gauged with the help of SA to better inform strategic decisions. Measuring customer satisfaction is a reliable predictor of future sales trends. Public sentiment, which can be collected from statements in social media or product reviews, is used by many marketing agencies to suggest the best approach for selling a product to businesses. Additionally, political parties use public opinion gleaned from text on social networks, blogs, and forums as a basis for their campaign strategies.

2.2 Related work

Analyzing social media data, especially that which is timely to current events, has been the topic of a number of recent publications. The widespread usage of social media has piqued the interest of academics, who have launched numerous studies to learn more about these phenomena. The discipline of Twitter sentiment analysis has seen significant growth in recent years.

2.2.1 Machine learning-based Twitter sentiment analysis

(Pavitha *et al.*, 2022) suggested an approach that would use Cosine Similarity to recommend films to users based on their preferences. Existing recommendation algorithms perform their job, but they do not justify whether or not a movie is worth one's time. This system evaluates sentiment of selected movie reviews utilizing ML to enhance overall user experience. The supervised ML techniques of SVM & naive bayes (NB) classifier are utilized to raise efficiency and precision. Additionally, this study compares NB and SVM using precision, accuracy, recall, F1 Score metrics. In terms of accuracy, SVM scored 98.63 percent, while NB scored 97.33 percent. Consequently, SVM outperforms NB and is a better fit for Sentiment Analysis.

(Singh, Kumar and Kumar, 2022) This study suggests using machine-learning to evaluate social-media information for text sentiment analysis. The described procedure consists of 3 independent steps. During the first stage, called preprocessing, raw text data is cleaned up and prepared for analysis. In the next step, they use term frequency-inverse document frequency (TF-IDF) technique is utilized to extract relevant features from documents. In addition, the third step entails using the retrieved features to provide predictions for classifier. The study applied a free, downloadable Twitter dataset from US Airlines. For classification and analysis, multiple ML approaches are employed. There are a number of metrics for evaluating performance displayed in the outcomes, including recall, precision, F1 score, and accuracy. Results were most reliable when processed with a SVM.

In this research, (Zhao *et al.*, 2021) developed a novel ML method for analysing online reviews' emotional tone by superimposing an adapted version of bat algorithm on an Elman neural network (LSIBA-ENN). The following are four primary steps that are recommended by Sentiment Analysis: i) Preprocessing, ii) Data collection (DC), iii) Term weighting (TW) or Features extraction (FE), Polarity or Sentiment Classification (SC), and Feature Selection (FS). The first step in data collection process is to scrape E-commerce sites for consumer feedback on the products of interest. The retrieved data from a web scrape is then preprocessed. These preprocessed data are passed through FS and TW for further processing utilizing LTF-MICF (log term frequency-based modified inverse class frequency) and HMEWA (hybrid mutation-based earth warm algorithm). Last but not least, the LSIBA-ENN converts the HM-EWA data into a sentiment classification of positive, negative, and neutral for customer reviews. The proposed and existing classifiers are compared using two benchmark datasets. The outcome demonstrate that LSIBA-ENN performs exceptionally well in SC when compared to other algorithms. A reviewer's comments are spot on. Recall for state-of-the-art ENN is improved from 83.55 to 87.79 with the help of the suggested LTF-MICF scheme, but it drops to 83.48,85.48, and 86.04 when ENN is trained using the TF-IDF, TF-DFS, TFB and W2V methods.

(Jayakody and Kumara, 2021) Twitter posts based on product reviews were analysed using k-nearest neighbour, SVMs, logical regression, and ML algorithms, as well as count vectorizer and tfidf mechanisms for converting texts into vectors, as shown. Finally, a machine-learning model was trained using this information. The results of this analysis were then presented in the form of a graph. To determine the optimal "model vectorising" combinations, this research employs a systematic assessment of 6 different alternatives. "Logistic Regression with Count Vectorizer" had the greatest accuracy score (88.26 percent) of all the methods tested. At last, they'll compute true positive and false negative rates.

The goal of (Biradar, Gorabal and Gupta, 2021) was to develop big data technologies that could be used to collect and manage massive volumes of unstructured information from actual social media with the aim of performing sentiment analysis to ultimately identify brands and services. The algorithm for sentimental analysis of customer reviews was developed through a series of steps, including data clustering according to specific domains, pre-possessing the datasets, constructing a feature vector using tf-idf vectors and n-gram models for the extraction of synonyms, and finally, performing a classification sentiment analysis. The results show that utilising sentiment analysis in conjunction with unsupervised data clustering into domains of interest and supervised machine learning methods is an efficient way to handle massive amounts of Twitter data. This aids in interpreting, analyzing, and computing interactions and linkages among individuals, subjects, and ideas. The developed solution transfers data from databases to a Hadoop cluster at a rate 1.5 times faster, with an accuracy of roughly 80%.

In this work (Alsalman, 2020), a corpus-based method is suggested for analyzing Arabic tweets that have been marked as positive or negative on Twitter. This method combines a NB classifier with a stemmer, a (TF-IDF) algorithm, and an N-gram tokenizer. The proposed method for sentiment analysis is put to the test using a publicly available Twitter dataset utilising a wide range of metrics for measuring effectiveness. The results validated the viability of the approach shown. Additionally, the outcomes demonstrate the strategy's superiority over the comparable task and increased accuracy by 0.3%.

(Alkhaldi *et al.*, 2020) The purpose of this research was to make use of ML in order to conduct an analysis of tweets relating to Saudi General Entertainment Authority (GEA). RapidMiner was used to collect and analyse 3,817 tweets. Three ML methods, MLP (multi-layer perceptron), SVM, RF (random forest), and one DL (deep learning) approach, namely RNN (recurrent neural network), were utilized to categorise tweets as pro- or anti-campaign. The classification model was evaluated using both k-fold validation tests and percentage split. The findings suggest that the populace is pleased and approves the actions of GEA. In terms of gender, the percentage of females who supported the motion was higher than that of males. The

RF approach also shines beyond the capabilities of rival algorithms in terms of classification accuracy and error rate. This is because RF method uses a supervised learning approach.

In this study, (Rashid *et al.*, 2020) utilizing Flume to store the Twitter Streaming Data within Hadoop's HDFS and following it up with Apache Hive to extract the data. The next step is to use Apache Mahout and some ML categorization techniques to figure out what kind of feelings are being expressed. For better results when analysing real-time tweets for sentiment, a novel approach is used that combines DT (Decision Trees) and NB Algorithms. In compared to Nave Bayes Classifier, the executed research strategy attained an accuracy of 86.44 percent.

(El Rahman, AlOtaibi and AlShehri, 2019) The primary purpose of this study is to create a model that can use actual data mined from Twitter to do sentiment analysis. Due to its lack of structure, Twitter's data is difficult to analyse. The combination of unsupervised and supervised ML methods distinguishes their proposed model from previous work in this sector. The following steps are involved in performing sentiment analysis: The extraction of Tweets from the Twitter API, followed by data cleansing and discovery. The data was used to train several models. Sentiment was assigned to each retrieved tweet to determine if it was good, neutral, or negative. Data were taken from two subjects, KFC and McDonald, for establishing which is the most well-liked. It used a number of ML methods. The results of these models were examined using a variety of testing criteria, including cross-validation and f-score. In addition, our model excels at mining texts that have been scraped straight from Twitter.

(Wongkar and Angdresey, 2019) Employed using python programming language to conduct a twitter sentiment analysis of 2019 presidential contenders inRepublic of Indonesia. This research of sentiment comprises a variety of stages, namely data collection using Python tools, text processing, evaluating training data, and text categorization using NB method. The NB technique is helpful for categorising groups or measuring the extent of public opinion. According to the findings of survey, pair consisting of Jokowi and Ma'ruf Amin had a favorable sentiment score of 45.45percent and a negative score of 54.55percent, whilst pair consisting of Prabowo and Sandiaga had a favourable sentiment score of 44.32percent and a negative score of 55.68percent. After training on data from all of the presidential candidates, they ran a test on entire dataset and got overall accuracy of 80.90% ~ 81%. In this study, they compared the accuracy of SVM, NB, & K-nearest neighbour (KNN) approaches using RapidMiner, and they found that SVM method had an accuracy of 63.99percent, NB method had an accuracy of 75.58percent, and K-NN method had an accuracy of 73.34percent.

(Saad and Yang, 2019) The goal of this study is to apply ML techniques in order to conduct a comprehensive ordinal regression-based analysis of the sentiment included inside tweets from social media platform Twitter. The suggested technique involves first preprocessing tweets and then employing an effective feature extraction algorithm. Then, scoring and balancing will be performed on these attributes under various classifications. DT, RF, SVM and multinomial LR (logistic regression) are the techniques used to classify sentiment in the proposed framework. This technique is really implemented using a publicly accessible twitter dataset provided via NLTK corpus resources. The suggested technique can accurately identify ordinal regression utilizing ML techniques, according to experimental results. Furthermore, studies show that DTs beat all other algorithms and get the greatest outcomes.

In this work, (Shamantha Rai, Shetty and Rai, 2019) present feelings about tweets or reviews that have been published on Twitter can be found by looking for a specific term within the tweets and evaluating the polarity of tweets as either negative or positive. Feature selection for every score word is used to assess the feelings of tweets. Choosing the most useful characteristics. The characteristics of words are trained and tested using NB Classifier (NBC), and polarity of every tweet's sentiment is assessed using NBC as well. 3 ML classifiers—SVM, RF and NB —are evaluated in terms of performance evaluation metrics like accuracy, precision, and time SVM.

(Rathi *et al.*, 2018) The primary focus of this study is classification of emotional data extracted from Twitter messages. Previous studies that relied on existing machine-learning methodologies for Sentiment Analysis found that these techniques did not yield optimal results for sentiment classifications. The proposed solution is being applied to ensemble ML algorithms to increase its efficacy and reliability, hence enhancing classification outputs in domain of sentiment analysis. The experiments show that when DT plus SVM are combined, resulting classification results are superior to those obtained using individual classifiers with regards to accuracy and f-measure.

To address the challenge in understanding the sentiments, (Abid *et al.*, 2019) devised a combined architecture which combines RNN and CNNs with a global average pooling layer for capturing long-term dependencies and a word embedding approach utilising GloVe acquired through unsupervised learning in context of huge Twitter corpora. On Twitter corpora, the experimental models perform better than the baseline model, with an accuracy of 90.59 percent on the Stanford Twitter sentiment corpus, 89.46 percent on the sentiment strength

Twitter data, and 88.72 percent on the Healthcare reform dataset. These results show that the experimental models outperform the baseline model.

(Juneja and Ojha, 2017) The purpose of this study is to classify Twitter data into negative and positive attitudes utilizing multiple supervised ML classifiers to forecast outcomes of the Delhi Corporation Elections and to select the most reliable ML classifier. This research analyses how well different classifiers perform on a dataset obtained from Twitter that contains information about various political parties. Experiments have shown that the Multinomial Nave Bayes classifier is the best accurate predictor of sentiment, with a rate of 78 percent.

This work (Mishra, Rajnish and Kumar, 2017) addresses Twitter sentiment analysis, existing sentiment analysis tools, relevant work, the framework utilized, a case example to illustrate the effort, and finally, the outcomes. 50 percent of collected opinions are positive, 20 percent are negative, and 30 percent are neutral, as shown by results.

2.2.2 Polarity study of Twitter sentiment using machine learning

This work (Zervoudakis *et al.*, 2021) demonstrates Opinion Mine with Twitter Data, a Bayesian framework for opinion mining. Their system begins with a large import of Tweets via the Twitter API. Afterward, the imported Tweets undergo additional automated processing for purposes of building an untrained set of rules and random variables. The Bayesian Network is then constructed by employing an assortment of untrained rules, a collection of random variables, and a body of information. After that, trained model may be utilised for the analysis of newly published Tweets to identify quality of those Tweets. Lastly, the built-in model can be gradually retrained to increase its robustness. They chose tourism as the application domain for development of their methodology because it is one of the most well-liked subjects on social media. Their approach can forecast users' travel intentions. Their framework takes advantage of a learning-by-doing strategyIn other words, the final model may be gradually retrained with current data to improve its precision. The rules that make up final model are automatically and effectively constructed, and their design is adaptable enough to be used for opinion mining on social media across a wide variety of subjects.

The purpose of (Arif and Binte Hossain, 2021) was to develop and train a model that can accurately and automatically classify customer twitter reviews of leading cell phone brands. For obtaining the polarity values of each tweet review in dataset, they employed the Python TextBlob module. In addition, they utilized Logical Regression, Nave Bayes, RF SVM, and DT, as well as TF-IDF vectorizers and Bag of Words, to train and construct a model which will classify customer tweet reviews into 5 opinion groups: Weakly Negative, Strongly Negative Strongly Positive and Neutral, Weakly Positive. They found that employing the Bag of Words vectorizer, LR, and SVM methods outperformed other algorithms with an accuracy of 88 percent while utilizing TF-IDF vectorizer, SVM outperformed other algorithms with an accuracy of 87 percent.

This study (Saad, 2020) suggested a machine-learning model for classifying Twitter postings as either negative, positive or neutral. They applied their approach to a database consisting tweets from six different US-based airlines. They began their model with preprocessing procedures in which they cleaned tweets and extracted characteristics to represent them as feature vectors, and then they constructed their Bag of Words (BoW) model. Tweets were classified using six ML approaches throughout the classification stage: DT, LR, SVM, RF, and XgBoost. For the goal of testing and validating the data, they utilized the k-fold cross-validation method after dividing the data into 70 percent training and 30 percent testing during the validation stage. After evaluating each classifier, they calculated its Precision, Accuracy, Recall, and F1 Score. After looking over results from every classifier, they settled on SVM as having the highest accuracy (83.33%).

In this study,(Srivastava, Singh and Mangal, 2020) Using Twitter's textual data and a sentiment analysis method, the suggested model analyzes the views of Indian populace on plastic ban, and results show that a machine classifier based on a learning method can achieve an accuracy of 77.94percent in its classifications across a variety of datasets. The outcome will assist determine how efficient and successful India's ban on polybags would be when fully implemented.

(Pavel *et al.*, 2017) In this study, the authors look into the practises of Twitter users in regards to the use of shortened URLs. This study delves at the phenomenon of shortened URLs on Twitter. To be more particular, the objective is to investigate the content that is linked to short URLs as well as potential effect on effectiveness with which tasks involving sentiment analysis (opinion mining) are carried out. Twitter-based opinion mining has been applied to a variety of fields, namely healthcare, identifying public opinion on political topics, financial modeling, or advertising. However, previous study has completely ignored tweets, including URLs. Given that Twitter users frequently share URLs to articles backing a certain political figure, pieces from prominent financial publications, and product reviews, it is not difficult to imagine how opinion mining could be enhanced. This study is based on an investigation of 3 independent

Twitter datasets with different numbers of tweets containing short URLs. Popular ML algorithms utilized in opinion mining were implemented in a variety of experimental situations to see which solutions were the most profitable.

(Mridula and Kavitha, 2018) entailed analyzing tweets with hashtag #Make in India, which is associated with an effort launched by Prime Minister of India, Mr. Narendra Modi. Here, tweets serve as datasets. Collecting, cleaning, and filtering the tweets is referred to as preprocessing. The retrieved features are then chosen for analysis. Utilizing methods from ML, classification, assessment, and visualization is performed. In this study, "R Studio" is utilized for implementation and analysis.

(de Godoi Brandão and Calixto, 2019) The purpose of this research was to evaluate efficacy of SVM classifier for Twitter opinion mining using five different publically available datasets. N-Gram and TF-IDF were used in the classifier modelling phase, with k-folds cross-validation procedures used for feature extraction. Accuracy ranged from 63.93% to 81.06% across several iterations of N-Gram using L-gram, 2-gram, and 3-gram with k-folds cross-validation at 10, 15, and 20 levels. It is possible to enhance acquired satisfactory results by employing an optimization strategy to alter the classifier parameters.

(Jumadi *et al.*, 2016) Suggested opinion mining using SVM technique for classifying massive amounts of twitter opinion data. The State Islamic University of Sunan Gunung Djati Bandung, which is one of Indonesia's major universities, will be the subject of this opinion mining in order to gain insight into public's perception of the institution. Positive and negative opinions are two categories for classifying opinions. Before classification, pre-processing phase involves case folding, emotion tokenization, stop words removal, data cleaning, and stemming. An accuracy of 0.838 was observed, with a recall for the positive class of 0.76. Then, for the negative class, the accuracy was 0.78, and the recall was 0.853%. This research's SVM classification of opinions is 78.75 percent accurate.

In this work, (Barnaghi, Ghaffari and Breslin, 2016) provide a negative or positive sentiment for tweets employing well-known ML technique for text classification. Additionally, they create a trained approach to complete a task using manually labeled negative/positive tweets. The objective of this study was establishing a relationship among Twitter sentiment and actual events. The trained model employs BLR (Bayesian Logistic Regression) classification technique. They employed external lexicons to determine whether tweets were objective or subjective. adding bigram and unigram features, filtered out features by using (Term

Frequency-Inverse Document Frequency) TF-IDF. They selected FIFA World Cup 2014 as their case study, and they mined, filtered, and processed tweets utilizing Twitter Streaming API and some of the official world cup hashtags. Their goal was to investigate the public emotion reflections towards unexpected events. The similar methodology can be utilized to anticipate future events.

2.2.3 Twitter Sentiment Analysis-Based Regression Algorithms

(Sufi, 2022) The consequences of bad news and breaking news are analysed in this study using (Artificial Intelligence) AI-based techniques as sentiment analysis, automated regression, and entity detection. The methodology of this study was implemented using a cutting-edge algorithm, which allowed for the discovery of all factors and themes that influence people's pessimistic views on international news. From June 2, 2021, to September 1, 2021, the solution was hosted in the cloud. From 2,397 news outlets in 192 countries, it automatically gathered and analysed 22,425 news pieces from around the world. After this time period, 34,975 entities will have been automatically organized into thirteen distinct entity types. Recall, Precision, and F1-score measurements of entity detection categorization accuracy were respectively 0.995, 0.992, and 0.994. In addition, accuracies of LR and linear regression were determined to be, on average, 0.895 for AUC or Area Under the Curve and 0.255 for maximum absolute percentage error (MAPE). Finally, the proposed approach was effectively implemented in a variety of contexts, such as tablets, smartphones, and desktop computers.

(Kaur, 2022) used well-known supervised machine-learning approaches like NB and LR to analyze real-time news data and determine how they affect data. Additionally, ML techniques are used to research and study use of sentiment analysis. This simplistic method provides some much-needed context for the posts and blog entries seen on popular news websites by categorising the views presented therein. The purpose of the study is merging supervised learning approaches with online scraping approaches in order to optimize precision, accuracy, and recall of news article findings.

In this study, (Leow, Nguyen and Chua, 2021) suggest two innovative models, (SMPT) Sentimental MPT and (SAW) Sentimental All-Weather , which capture the most recent market circumstances using feelings from Twitter via Google's Bidirectional Transformer (BERT) model. To optimize the models for various goals, such as maximizing cumulative returns and minimizing volatility, genetic algorithms were applied. Utilizing common portfolio performance indicators such as cumulative returns, value-at-risk, and sharpe ratio, proposed models were shown to be superior to buy-and-hold SPY indices, MPT method, and CRB framework for an All-Weather Portfolio after being tested on an out-of-sample period spanning January 2020–April 2020. This was determined by testing models on an out-of-sample period.

(Khan *et al.*, 2021) This study's focus is on Twitter, and it provides a method for spotting trends in real time as they unfold across tweet streams. The proposed method ranks the most popular phrases and hashtags in real time and identifies hot topics on Twitter. Further, the paper explores the research behind social media trend forecasting; Finding the themes and words inside those topics is the focus of this investigation, which employs exploratory data analysis in addition to the biterm topic model (BTM), Tf-IDF (term frequency-inverse document frequency), CCA (combined component approach) methods. Investors, advertisers, industries, and everyone else with a stake in today's business climate will benefit from the information provided by this study. An in-depth study of the data that could direct their attention, resources, and efforts toward the most promising areas of development, markets, and products.

(Alicioglu, Sun and Ho, 2020) In this research, an original framework for evaluating the potential for accidents is proposed, one that is powered by ordinal regression. The lack of non-accident data in the research of transportation safety is a difficulty for risk assessment. In light of this challenge, they also provide a method for producing negative data that employs the feature weights generated by multinomial LR. The proposed ordinal regression framework was shown experimentally on two real-world datasets from Transport for Greater Manchester in United Kingdom and National Highway Traffic Safety Administration in United States, showing its practicability and robustness. Four ordinal regression approaches are compared: the logistic immediate-threshold method, the logistic all-threshold method, the least absolute deviation method, and the ordinal ridge method. These algorithms are used to analyse ordinal data. They also look into how the proposed risk assessment methodology would change if they randomly oversampled or under-sampled the US dataset. Specifically, they provided empirical evidence that the bagging with random oversampling approach to ordinal regression using logistic all-threshold ordinal regression method provides best prediction.

(Ullah *et al.*, 2020) One study used a variety of features, such as TF-IDF, Bag of Words, emoticon lexicons, and N-gram, in conjunction with DL and ML algorithms to deduce sentiment from airline tweets. This study shows that the emotional tone of a message is most strongly influenced by the emoticons used in it. The superiority of DL algorithms over ML algorithms has also been established.

The purpose of (Saad and Yang, 2019) is to use machine-learning approaches to do an in-depth ordinal regression-based sentiment analysis of tweets. In the suggested method, tweets are first subjected to pre-processing, after which a feature extraction approach is used to generate an effective feature set. The scoring and balancing of these aspects can then be found under various different classes. With the findings of sentiment analysis as input, the proposed system uses classification techniques RF, DTs, Support Vector Regression, and Multinomial Logistic Regression to assign ratings to statements. The developers of system rely on a Twitter dataset that is made available by NLTK corpus resources in order to put together the version of the system that is actually functional. Experiments validate the effectiveness of the proposed strategy for detecting ordinal regression with ML methods. In addition, the data shows that among all the algorithms tested, DTs performs the best.

(Windha Mega and Haryoko, 2019) The findings of this opinion poll are expected to be useful for all parties involved, but especially for Go-Jek. The SVM algorithm is dependable for classification and regression since it can handle high-dimensional data sets, classification problems, kernel linear and non-linear regression. However, there is a problem with SVM when trying to optimise it by picking the right parameters. The issue of picking appropriate parameters in support machine vector approach can only be solved by using a genetic algorithm.

The purpose of (Gamal *et al.*, 2019) is to compare and contrast the algorithms used for sentiment analysis using different n-grams as a feature extraction approach, and to evaluate the collected tweets written in various Arabic dialects. Metrics like accuracy, precision, recall, and f-measure are utilised so that different algorithms' levels of effectiveness can be compared. Based on findings, using a unigram in conjunction with either PA (Passive Aggressive) or RR (Ridge Regression) results in maximum accuracy of 99.96percent.

With intention of aiding first-time users, (Alahmadi and Zeng, 2016) looked into the assessment and design of a personalised recommender system (RS) that makes use of unspoken social trust from OSNs (Online Social Networks). Taking into account things like retweet actions and followers/followings lists, proposed method exploits a user's social network on the popular social micro-blogger Twitter to generate implicit trust. Use the trust values you've calculated to cast your vote for the opinions of friends on Twitter regarding a particular product, like a movie. Higher confidence in a friend's judgement increases the likelihood that his or her suggestions will be taken seriously. To begin, we use a probabilistic sentiment analysis method to parse short tweets for friends' opinions expressed as ratings on a multi-point scale. In a second step, features of the user's trust relations are gleaned from their social networks. In addition, the characteristics of social trust are optimized using a genetic algorithm. Finally, the SVR algorithm is considered in this study to predict user ratings. According to the experimental findings we obtained using real-world data from Twitter, suggested method outperforms many other similar studies in terms of accuracy. These results indicate promise in addressing the "cold-start problem" experienced by new users of the systems by including the users' OSNs.

2.3 Research gap

(Pavitha *et al.*, 2022) used two ML models Naïve Bayes and SVM for analysing sentiment. When compared to NB, SVM offers higher accuracy. SVM achieved a 98.63% accuracy rate, while NB reached only 97.33 %. ML techniques including LR, k-nearest neighbor, SVM, and tf-idf and count vectorizer methods were used in (Jayakody and Kumara, 2021). Accuracy was best for LR using Count Vectorizer (88.26 percent). Finally, we'll know how many cases were false positives and how many were false negatives. In order to decipher the sentiment in this data, the authors of this study utilised ML classification methods in Apache Mahout. To improve the efficiency of sentiment analysis on Twitter data in near real time, a novel technique is presented that is based on a merging of NB and DT Algorithms. 86.44% accuracy compared to 81.11% for the NB Classifier (Rashid *et al.*, 2020). In addition to using TF-IDF vectorizers and Bag of Words independently, SVM, NB, DTs, RF and LR algorithms were employed in this study's model training and construction in order to classify customer twitter reviews into five opinion categories: Negative, Neutral, Positive, and the most positive statements. It has been found that SVM and LR algorithms achieve 88% accuracy when employing the Bag of Words vectorizer, and that SVM achieves 87% accuracy when employing the TF-IDF vectorizer (Arif and Binte Hossain, 2021). Six different ML methods were used in this research. Tweets could be categorised using the SVM, LR, RF, XgBoost (XGB), NB, or DT. SVM exhibited the best accuracy of any classifier tested, at 83.31 percent (Saad, 2020). (Gamal *et al.,* 2019) applied various algorithms and evaluated accuracy, f-measure, recall, and precision. Experiment showed that a unigram with Passive Aggressive (PA) or Ridge Regression (RR) yields the maximum accuracy of 99.96%. This study examined Twitter for sentiment analysis with a performance accuracy of 50% positive opinions, 20% negative opinions, and 30% neutral opinions. Sentiment analysis's end result is a positive/negative/neutral categorization of text expressed in natural language. Massive amounts of data are generated by social network sites; however, this data is unstructured and cannot provide useful insights unless it is examined. To make this massive amount of data relevant, we do sentiment analysis, which involves extracting and classifying features from the data. Sentiment analysis is crucial in the modern world because people are constantly influenced by the thoughts and views of others. Today, whether someone wants to purchase a product, vote for a candidate, or watch a movie, etc., they will first check social media websites such as Twitter, Facebook, and Tumblr for reviews, comments, and opinions about the product, candidate, or film. Consequently, a system that can automatically generate sentiment analysis from this vast amount of data is required (Mishra, Rajnish and Kumar, 2017).

2.4 Summary

Datasets containing Trump's scraped tweets from the web have been used to do research into sentiment analysis, which in this chapter is defined as the application of NLP (natural language processing) & ML classifiers. The host dataset has undergone the primary sentiment analysis processes following data preparation. There have been additional steps taken to prepare the dataset for text vectorization and other natural language processing methods. Stopword removal, lemmatization, regular expressions, and tokenization, along with any other necessary methods, have been used to the textual data in order to eliminate unwanted terms. Our efforts to make "content" feature of dataset smaller in order to use less storage space paid off. Especially since the rise of social-media platforms, hateful actions have been commonplace in the previous two decades, and it is now a significant challenge to understand the subjective polarities of each published article. As a result, each sentence has been evaluated based on its polarity, being labelled as either negative, positive, or neutral. Now that the data has been cleansed, it's been trained & tested with the help of ML methods including gaussian NB, SVM and RF classifier with a comparison showing 72%, 89% and 88% accuracy of prediction for each classifier.

CHAPTER- 3

RESEARCH METHODOLOGY

The third chapter examines research methodology and many methods to research. This chapter is important because it demonstrates that the suggested LSTM technique may be utilised to solve a problem using current framework or strategy. This chapter also contains the proposed algorithm and proposed flowchart. Afterward described the data collection, data preprocessing, wordcloud, data splitting, feature extraction, LSTM and lastly discussed proposed algorithm.

This section details the process used throughout this research to evaluate the effectiveness of the suggested machine-learning technique for sentiment analysis of customer reviews. Here, a method of analyzing feedback is developed based on the experiences of customers. Data for this study was taken from Kaggle repository. Every sentiment gave some information about the company and the product or service they provide, which allows us in gaining a better understanding of how consumers and the general public feel about a certain product or service provided by the company as well as the company itself. It may also assist the organization in determining the cause of consumer unhappiness, which will allow them to develop new features or enhance existing ones for their product or service.

A dataset is gathered with the use of Twitter API, and then the data are classified as either negative or positive tweets. NLTK (Natural Language Toolkit) corpora repository is where the dataset can be viewed by the general public. This repository is popular and extensively employed in a wide range of study fields. The total number of tweets in corpus is 10,000, with 5,000 tweets having a negative tone and 5,000 tweets having a positive tone. Then we used some libraries such as numpy, pandas, tensorflow, sklearn, genism and seaborn. The data was then preprocessed checking and removing null values, removing emojis, removing URLs, Twitter handles, punctuation, extra spaces, numbers and special characters from tweets. Before putting the data in the model, it is split in two datasets for training and testing in the ratio of 75 and 25, respectively. Then, the machine-learning models were applied and their performance was assessed utilizing accuracy, F1-Score, precision, and recall. The flowchart for the above-mentioned steps is shown in Fig 1.

In this section, various strategies and methods are broken down in detail, as well as plans drawn up for their eventual application in an effort to accomplish the goals. In addition to that, this part elaborates about the outcomes that will result from the procedures. The parts that are to follow contain explanations of some of the strategies:

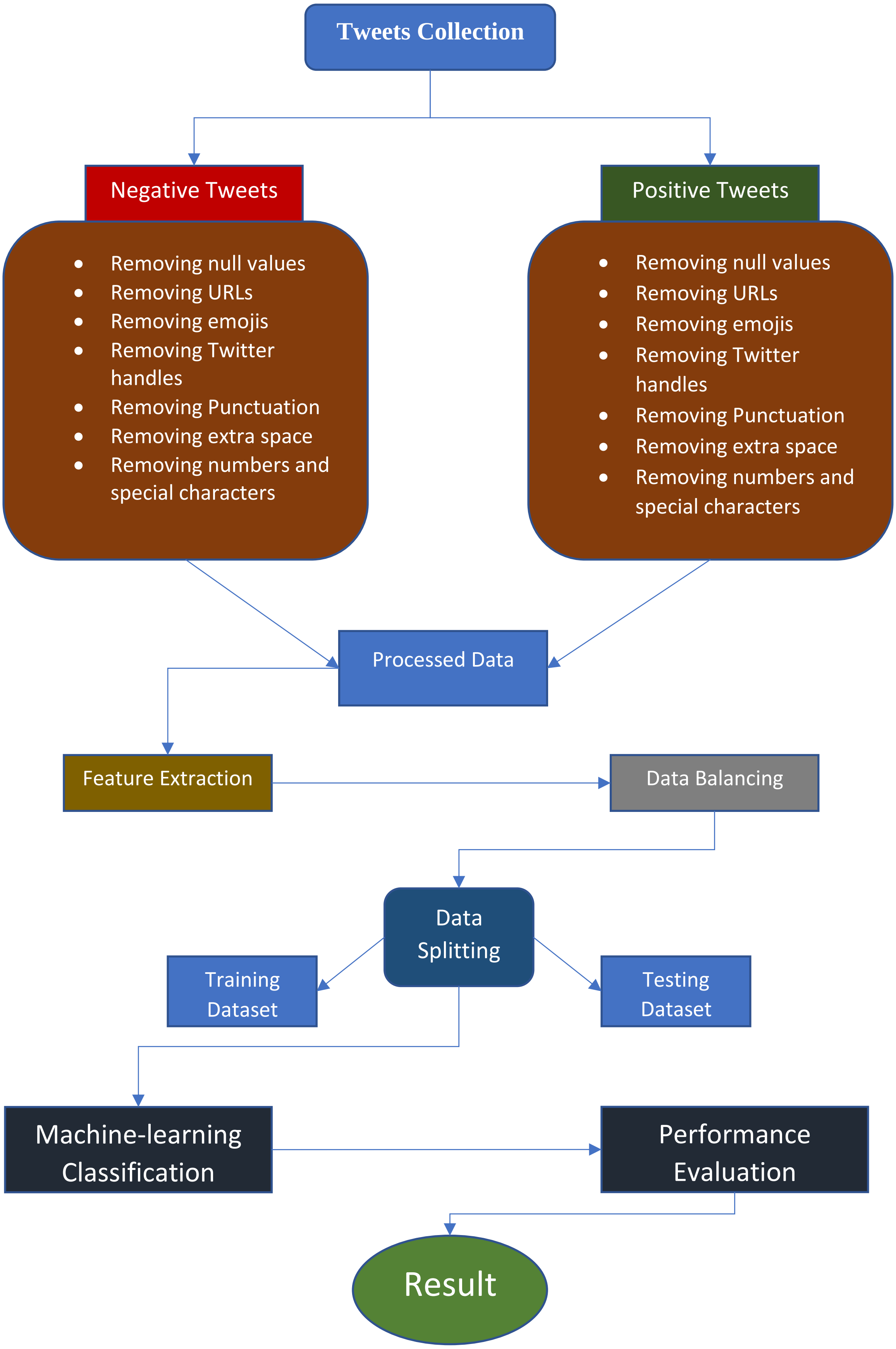

Figure 3.1: Proposed methodology flowchart

3.1 Data Collection

When talking about study, analysis, or assessment, data collection refers to the process of obtaining information from a variety of sources, measuring it, and recording it so that it can be used later. In order to answer a certain study questions or conduct an analysis of a certain occurrence, objective of data collection is to amass information which is accurate, representative, and pertinent to the issue at hand.

For purpose of this study, data were obtained from Kaggle repository. This repository has Twitter data that was gathered via the Twitter API, and it contains a total of 10,000 collected tweets. Of these tweets, 5,000 had a negative tone, and 5,000 had a positive tone.

3.2 Data Pre-processing

The process of cleaning, converting, and arranging raw data into a format that can be efficiently evaluated and used for a given purpose is referred to as data preprocessing. This process can also be referred to as data preparation. The purpose of preprocessing of data is to get it in a more accurate, consistent, and useable state so that it may be used later on in the analysis process. This step is critical for obtaining meaningful results from data analysis and machine learning algorithms. Following processes were done:

3.2.1 Removing null values

In the process of data pre-processing, it is usual practise to get rid of any values that are null. It entails finding and dealing with any data points in a dataset that are either missing or incomplete in some way. It is essential to get rid of null values in order to guarantee the correctness and dependability of the data.

3.2.2 Removing URLs

In NLP (natural language processing), one of the most common preprocessing steps involves removing URLs, also known as Uniform Resource Locators, from text input. URLs are not helpful for natural language processing activities since they do not provide any information that is meaningful for the analysis of the text. Eliminating them helps to reduce noise in the data, which in turn improves the NLP algorithms' overall performance..

3.2.3 Removing emojis

The elimination of emojis from text data is a common step in the pre-processing phase of natural language processing (NLP) projects. Emojis are graphical characters or icons that are commonly used in online forums, text messages, and social media platforms to communicate a range of emotions. Emojis are not considered significant text for natural language processing (NLP) tasks; hence, deleting them can help minimize noise in the data. Despite the fact that they can sometimes be helpful

3.2.4 Removing Twitter handles.

Preprocessing text data for use in natural language processing (NLP) tasks sometimes involves removing Twitter handles. Each user on Twitter is given a "handle," which is a special reference to their username. The "@" character, which is commonly used to preface handles, is not parsed as meaningful text by natural language processing tools. Data noise is reduced and NLP algorithm efficiency is increased when they are eliminated.

3.2.5 Removing Punctuations

Natural language processing (NLP) tasks sometimes necessitate the removal of punctuation from text data as a preprocessing step. NLP algorithms can be hampered by punctuation marks like commas, periods, and exclamation points because they are not part of the text's meaning. Data noise can be reduced by removing them.

3.2.6 Removing Extra space

Commonly performed in NLP preprocessing is removal of unnecessary spaces from text data. Because of formatting discrepancies, text data may contain unnecessary gaps, which might hinder the efficiency of NLP algorithms. Cleaning up the data and making it easier to process by NLP algorithms is facilitated by getting rid of unnecessary gaps.

3.2.7 Removing numbers and special characters

In NLP tasks, removing numbers and special characters from text data is a common preprocessing step. Numerals and special characters, such as symbols, punctuation marks, and numbers, are not regarded as meaningful text for NLP tasks and can cause NLP algorithms to malfunction. Eliminating them lessens data noise and enhances the effectiveness of NLP algorithms.

3.2.8 Removing stop words

The performance of NLP algorithms can be improved by removing stop words from the text input in order to make the data set smaller overall. Stop words are words which are utilized frequently in a language but don't have much of an impact on meaning of language as a whole. Examples of stop words are "a," "the," "and," "is," and others. These terms appear somewhat frequently in the text data, but they are not likely to convey any information that is relevant.

3.2.9 Removing single quotes new line characters

Text data may contain newline characters or single quotes as a result of inconsistent formatting, which might hinder the efficiency of natural language processing (NLP) algorithms. By removing them, we can clean up the data and boost the efficiency of algorithms.

3.2.10 Lemmatization

The process of reducing a word to its fundamental form, also known as the word's root or lemma, is referred to as lemmatization. In field of NLP, this is often done to simplify the processing of inflected (or derived) words by reducing them to their word stem, base, or root form.

3.3 Word cloud

Word clouds are a popular way to display keyword information on websites or to graphically represent free-form text. When applied to a single word, the resulting "word cloud" provides a visual depiction of that term's frequency of use in a given context. This may be taken as evidence of the word's widespread usage in that field. Wordclouds are an easy approach to visualise the ideology supported by a big body of data (in this case, Twitter data) and provide insight into underlying philosophy of data's related textual conversation (Ahuja and Shakeel, 2017).

3.4 Data Splitting

Data splitting is the practise of breaking a large quantity of information into small, more manageable sections for testing and training machine learning models. The goal of this procedure is to make the training and testing processes more efficient. The most common approach involves separating the data into a test dataset and a training dataset. The model is trained with the help of training set, whereas performance of model in real-world scenarios is evaluated with the help of test set.

3.5 Feature Extraction

The process of transforming unstructured textual data into organised numerical or categorical representations that may be used with machine learning algorithms is referred to as "feature extraction." In NLP, feature extraction is one of the most significant phases because it converts unstructured text input into a format which can be used by machine learning algorithms. This step is also one of the most critical steps. It is the process of transforming a set of input data into a collection of features. Starting with a stable dataset, Feature Extraction in machine learning creates the borrowed values, or features, that are anticipated to be descriptive and non-redundant, streamlining subsequent processes of learning and observation (Dara and Tumma, 2018).

Tokenization is the process of splitting a sentence or document into smaller units, called tokens, usually words or phrases. Tokenization is the process of converting text input into numerical features that may then be fed into machine learning algorithms. This process occurs within framework of feature extraction. The tokens generated by the tokenizer are used as input features for the model.

In field of NLP, a technique known as "padding" is utilised to manage sequences of varying lengths by appending padding elements to the sequences in order to make them all the same length. This is essential due to the fact that several machine-learning algorithms count on receiving inputs of a fixed length and are unable to process sequences of different lengths. "Pad sequences" is a common technique used for padding in NLP. It involves adding a special padding token, usually a zero or a null value, to the beginning, end, or both of the shorter sequences so that all sequences have the same length.

3.6 Long Short-Term Memory (LSTM)

It can be challenging to train recurrent or very deep neural networks since these types of networks frequently experience a exploding/vanishing gradient problem. The LSTM design was developed in order to address this deficiency in the process of learning long-term dependencies (Van Houdt, Mosquera and Nápoles, 2020).

The LSTM (Long Short-Term Memory) network is a kind of recurrent neural network that avoids problem of vanishing gradients by being trained with backpropagation across time. An LSTM is shown in Fig.2. Memory blocks are used in LSTM networks in place of neurons, and these blocks are coupled to one another via layers. A block is equipped with components that give it greater intelligence than a conventional neuron as well as a memory for recent sequences. Gates are the components of a block that are responsible for controlling its output and state. A block is controlled by an input sequence, and each gate included inside a block makes use of the sigmoid activation units to determine whether or not it will be triggered. This makes the transition from one state to another and the addition of information flowing through the block conditional. Within each unit, there are three distinct kinds of gates: 1) Forget gate: This gate decides, on a case-by-case basis, which information in the block can be discarded. 2) Input gate: makes decisions based on a set of conditions to determine which values from the input will cause an update to the memory state. 3) Output gate: makes decisions regarding what should be output based on the input as well as what is stored in the memory of the block. Each unit functions similarly to a miniature state machine, with weights assigned to the gates of the unit, which are learned in the course of the training operation (Liu and Wang, 2021).

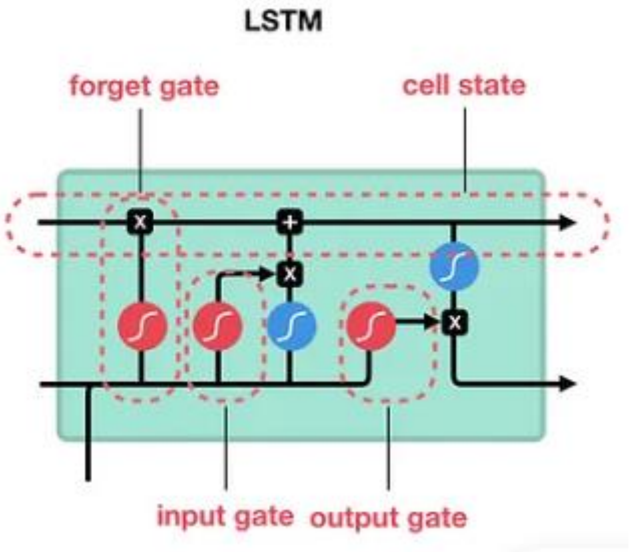

Figure 3.2: LSTM

3.7 Proposed Algorithm

Input: NLTK Twitter dataset

Output: Get classified outcomes

Strategy:

Step 1: Start implementation process

Step 2: Import set of data (NLTK)

Step 3: Processing of raw data before it is used

- Check null value
- Remove Punctuation

- Remove emoji
- Remove Twitter handles
- Remove newline character

Step 4: Apply feature extraction techniques

- Tokenizer

Step 5: Split the dataset into two parts training & testing set that divided into 75:25

- Training set (75%)
- Testing set (25%)

Step 6: Apply the proposed machine-learning model

- LSTM memory

Step 7: Evaluate Performance of the model using Performance metrics (Accuracy, Recall, F1 Score and Precision)

Step 8: Get the desired results.

Step 9: End

CHAPTER- 4

RESULTS

The implementation of this study is discussed in the fourth section, along with the presentation of dataset description, data visualization, and performance measures.

4.1 Dataset Description

Data collection is the process of acquiring data from a number of sources, quantifying it, and storing it in order to be utilized later on in the context of research, investigation, or evaluation. The goal of data collection process is to gather information that is accurate, representative, and relevant to the matter at hand. This is done with the intention of providing answers to specific study objectives or conducting an analysis of a particular incident. The Kaggle repository served as the source for the data that were used in this study. This repository has a total of 10,000 gathered tweets and contains Twitter data that was retrieved through the use of the Twitter API. There were a total of 10,000 tweets, 5,000 of which had a negative tone, and 5,000 of which had a positive tone.

4.2 Performance Evaluation

In machine learning, performance evaluation refers to the process of evaluating the accuracy and effectiveness of a trained model on a set of data. This is done to discover the model's robustness against out-of-sample data as well as any limits it may have. Precision, accuracy, F1 score and recall are typical measures of machine learning performance. Various performance metrics are explained below.

4.2.1 Confusion Matrix

The performance of a classifier can be measured with the use of a confusion matrix. The algorithm's projected class labels are compared to the actual class labels in the test data, yielding a classification accuracy score. The confusion matrix has four values: true positives, true negatives, false positive and false negatives, Metrics like accuracy, precision, recall, as well as F1 score can be determined from these statistics in order to assess the performance of algorithm.

4.2.2 Accuracy

The accuracy of the model can be evaluated by looking at how many predictions it gets right as a percentage of total predictions.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

4.2.3 Precision

The level of precision indicates what percentage of positive predictions turned out to be correct. It counts number of positive samples TP and divides by the overall number of positive predictions, whether correct or wrong TP, FP.

$$Precision = \frac{TP}{TP + FP}$$

4.2.4 Recall

The purpose of a recall test is to calculate the percentage of actual positives that were correctly detected. In order to accomplish this, it divides total number of positive samples by number of samples which were correctly projected as TP and those that were incorrectly forecasted as negative NP, FN.

$$Recall = \frac{TP}{TP + FN}$$

4.2.5 F1 Score

Historically, the F score has been represented by F1 score, which is harmonic mean of recall and precision:

$$F1\ Score = 2 * \frac{precision * recall}{precision + recall}$$

A model's accuracy and stability can be gauged by its F1-score. The optimal model has an F1 score of 1.

4.3 Experiment Results

4.3.1 Data Visualization

Putting information or data into a visual representation is known as data visualisation. It is used to help people understand and communicate patterns, trends, and insights hidden within data.

Popular forms of data visualization include line graphs, bar charts, pie charts, scatter plots, and heat maps. Data visualization's goal is to simplify & clarify complicated datasets for human understanding.

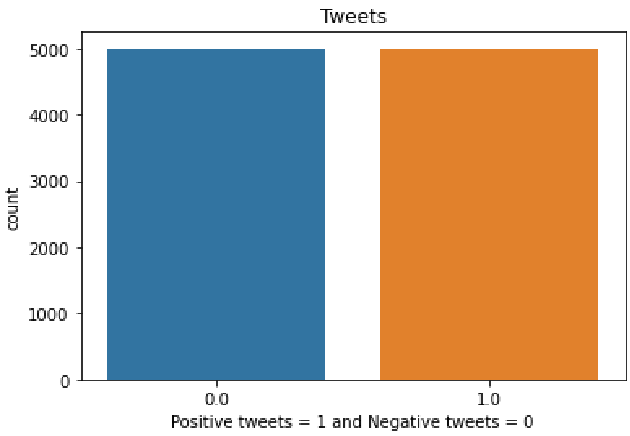

Figure 4.1: Count of Tweets

Fig. 4.1 shows the count of negative and positive tweets. The negative and positive tweets are both 5000 in number.

Figure 4.2: Word cloud for tweets

Fig. 4.2 shows the word cloud for tweets. It shows the most frequent words in a tweet with the largest size, in other words, the frequency of a particular word is proportional to its size in a word cloud.

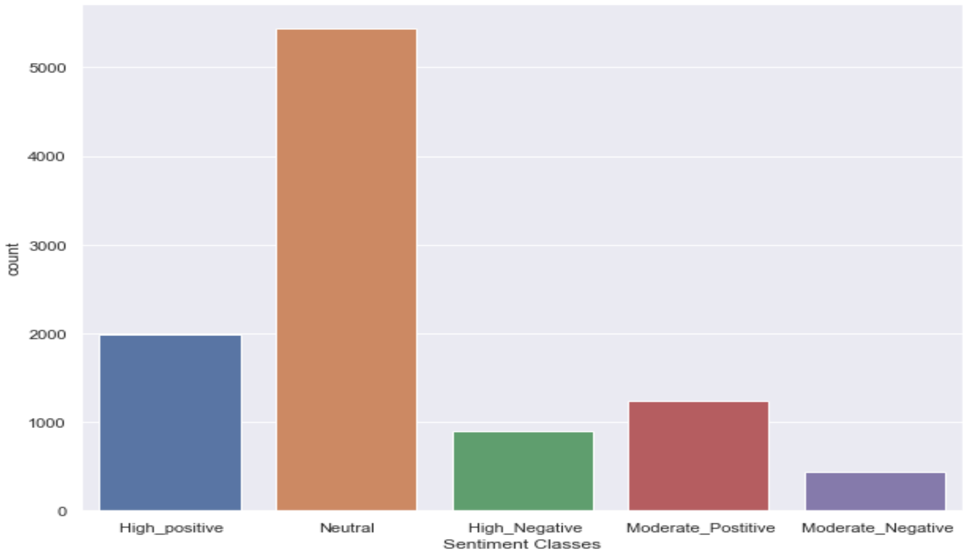

Figure 4.3: Data Distribution graph

Fig. 4.3 shows the sentiment distribution of tweets in various categories of high positive, neutral, high negative, moderate positive and moderate negative. The x-axis shows sentiment categories, and y-axis shows the number of tweets with that sentiment.

Figure 4.4: Word cloud of highly positive tweets

Fig. 4.4 shows the word cloud of highly positive tweets. "great", "good", "happy", "love" among others appear to be most frequent in the highly positive tweets category.

Figure 4.5: Word cloud of neutral tweets

The word cloud representing neutral tweets is displayed in Fig. 4.5. The words "follow", "thank", "smiley", "want" among others appear to be the most frequent in the neutral tweets category.

Figure 4.6: Word cloud of highly negative tweets

Fig. 4.6 shows the word cloud of highly negative tweets. "sorry", "hate", "feel", "tired" among others appear to be the most frequent in the highly negative tweets category.

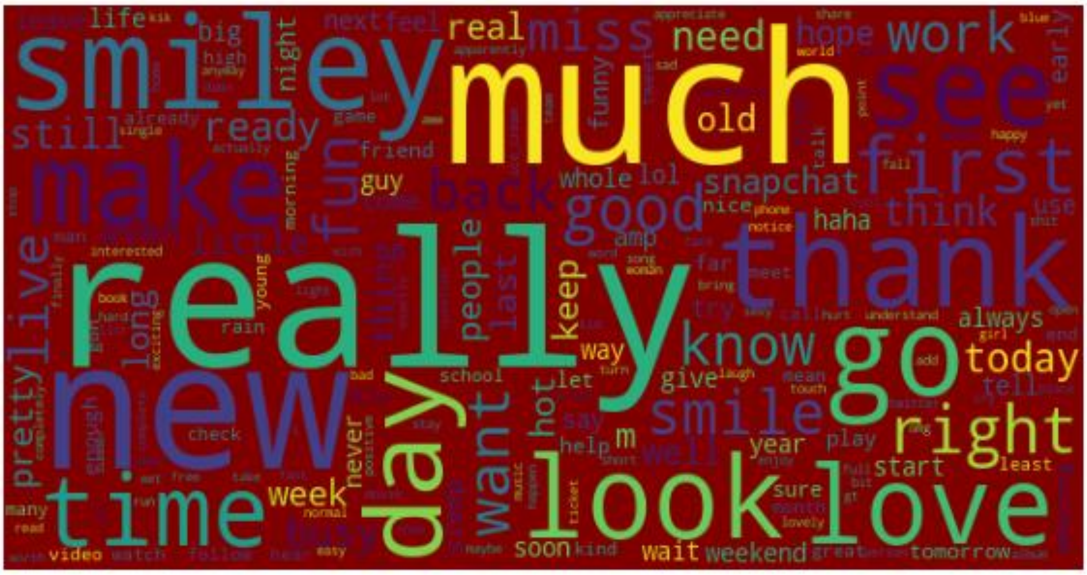

Figure 4.7: Word cloud of moderately positive tweets

Fig. 4.7 shows the word cloud of moderately positive tweets. The words "really", "new", "much" among others appear to be the most frequent in the moderately positive tweets category.

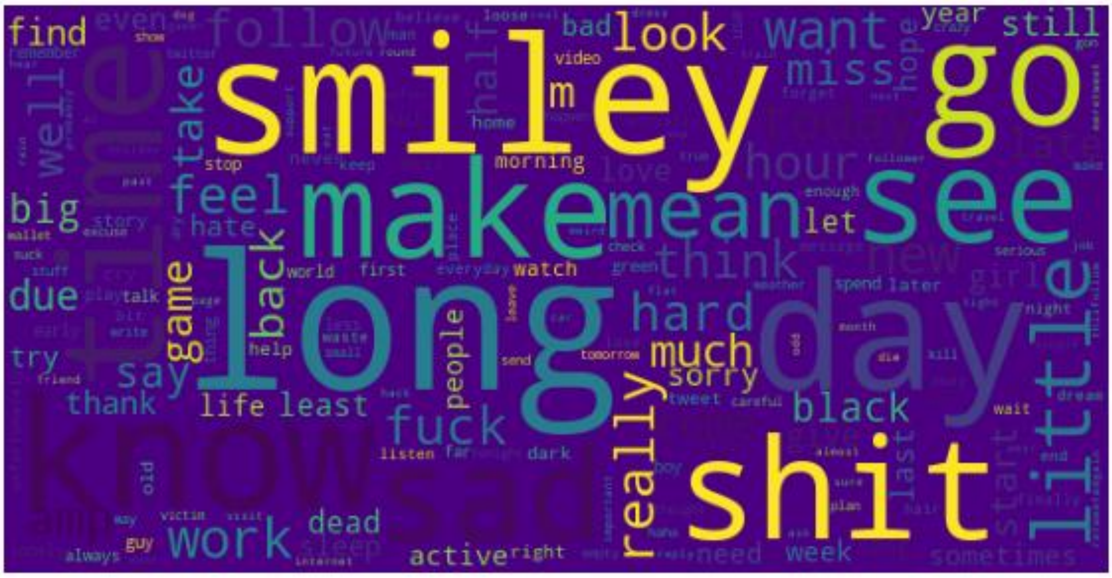

Figure 4.8: Word cloud of moderately negative tweets

Fig 4.8 shows the word cloud of moderately negative tweets. "smiley", "make", "long" and "day" among others appear to be the most frequent in the moderately negative tweets category.

4.3.2 Results of Simulation

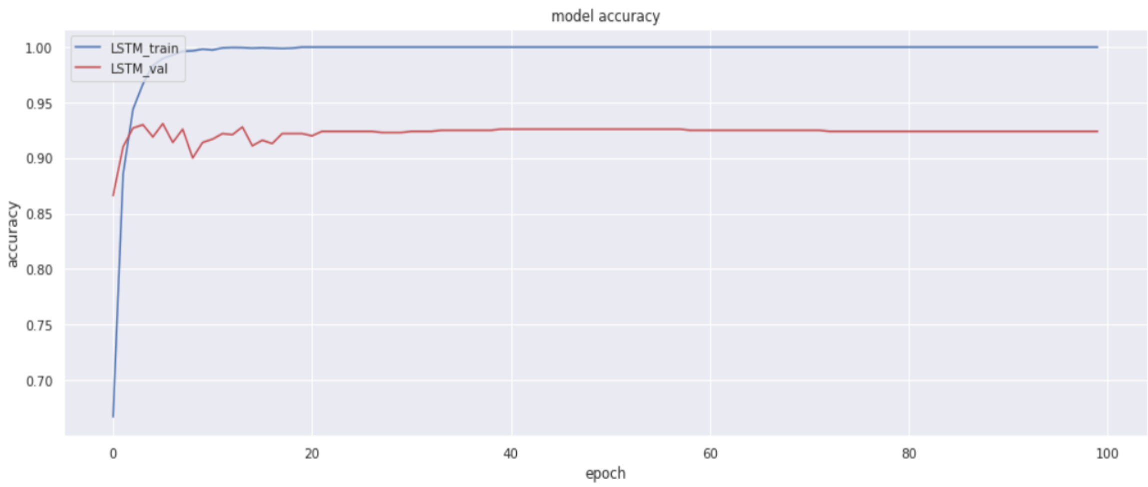

Figure 4.9: Accuracy graph for LSTM model

Figure 4.9 illustrates the LSTM model's accuracy graph. Training LSTM is shown by the blue line and testing or validation LSTM by the red line on a graph of epochs (x-axis) and accuracy scores (y-axis).

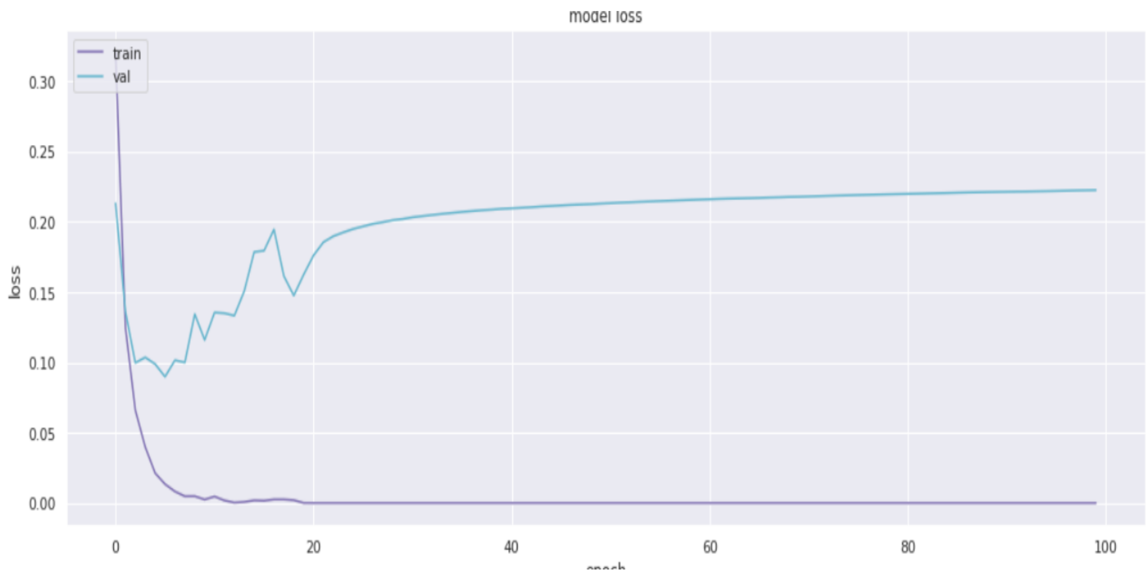

Figure 4.10: Loss graph for LSTM model

The loss graph for the LSTM model is depicted in Figure 4.10. On a graph with epochs shown on x-axis and loss values shown on y-axis, blue line represents LSTM validation or testing, while purple line shows training.

	precision	recall	f1-score	support
0	0.91	0.94	0.93	181
1	0.99	0.96	0.98	568
2	0.92	0.85	0.88	94
3	0.80	0.90	0.84	118
4	0.69	0.72	0.71	40
accuracy			0.93	1001
macro avg	0.86	0.88	0.87	1001
weighted avg	0.93	0.93	0.93	1001

Figure 4.11: LSTM model classification report.

The LSTM model classification report is displayed in Fig. 4.11. It shows the precision, recall, f1 scores values classification report for all five classes of sentiment categories. Additionally, it displays macro average as well as weighted average for accuracy values.

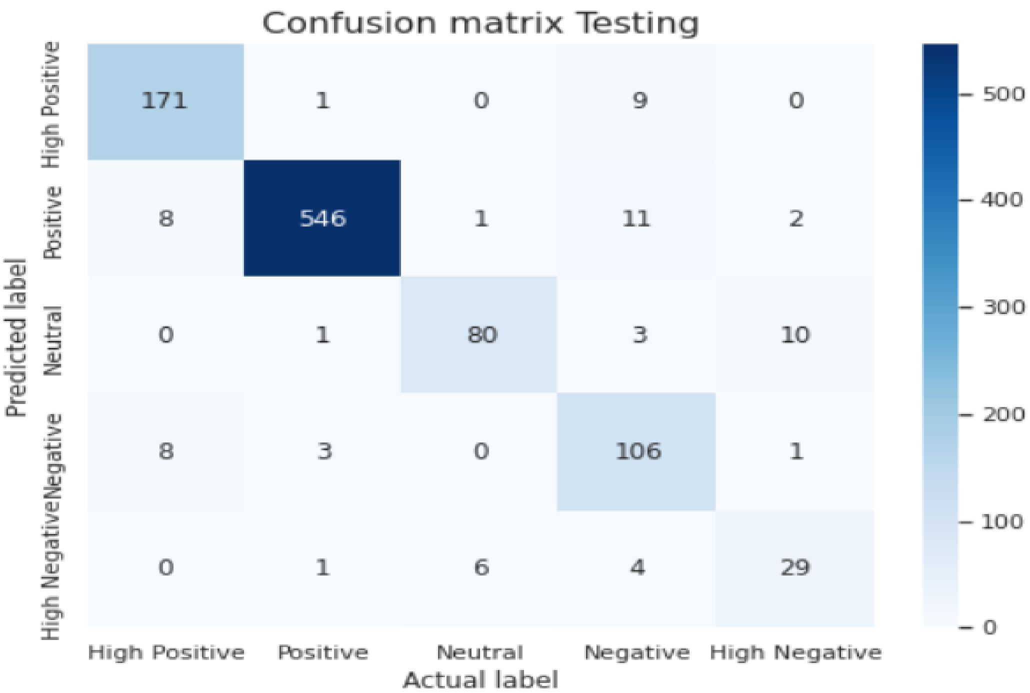

Figure 4.12: Confusion matrix for testing LSTM

The confusion matrix that was utilized to test LSTM may be seen in Figure 4.12. It displays the confusion matrix for all 5 categories of sentiment categorization, which are positive, neutral, negative, and high positive, respectively. High negative is also displayed. The values shown along the diagonal are correctly predicted.

Table 4.1: Performance evaluation metrics for the proposed LSTM model

Model	Test Accuracy	F1 Score	Precision	Recall
LSTM (Long Short-Term Memory)	0.93	0.93	0.93	0.93

Metrics for gauging the LSTM model's efficacy are laid forth in Table 4.1 below. The model has recall, precision, F1 score, and accuracy of 93percent.

Table 4.2: Comparison of proposed and base model performance

Model	Accuracy	F1 Score	Precision
Random Forest	86	68.99	82.92
Decision Tree	84.17	80.67	80.63
Multinomial Linear Regressor	80.71	71.45	75.87
LSTM (Proposed model)	93	93	93

The contrast between suggested model and base model's performance evaluation indicators is laid forth in Table 4.2. LSTM is the proposed model, whereas Random Forest, Decision Tree, & Multinomial Linear Regressor were the base models.

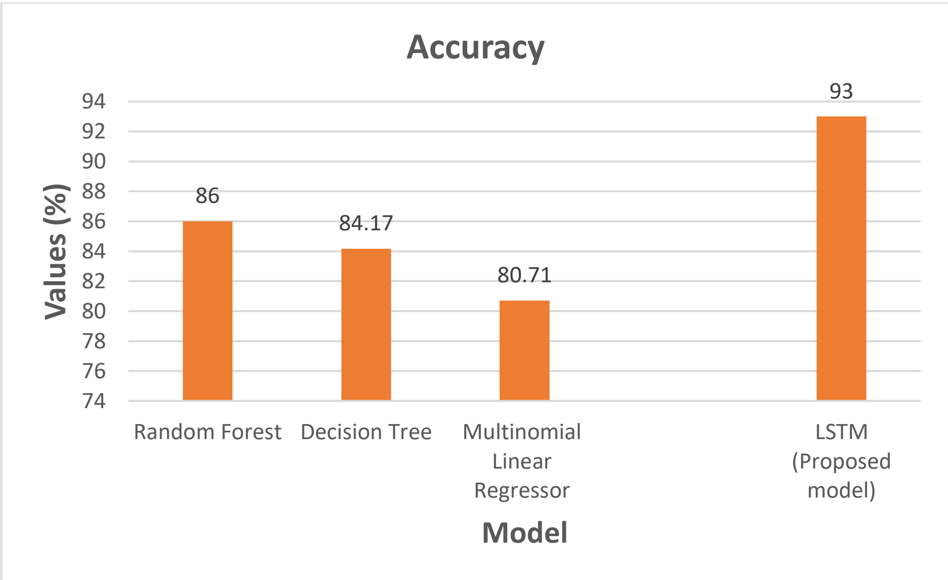

Figure 4.13: Accuracy value for the proposed and base models

Fig. 4.13 shows accuracy values for proposed and base models. The base models of Random Forest, Decision Tree and Multinomial Linear Regressor have accuracy values of 86, 84.17 and 80.7, respectively. The suggested LSTM model is 93% accurate.

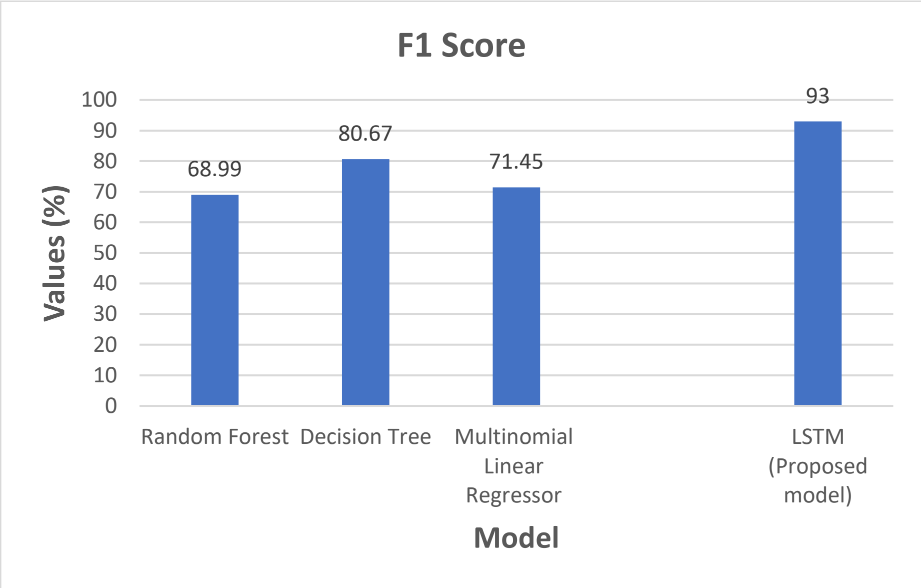

Figure 4.14: Base & proposed model F1 scores

Fig. 4.14 shows the base and proposed model F1 scores. The base models of Random Forest, Decision Tree and Multinomial Linear Regressor have F1 scores of 68.99, 80.67 and 71.45, respectively. The proposed LSTM model has a F1 score of 93.

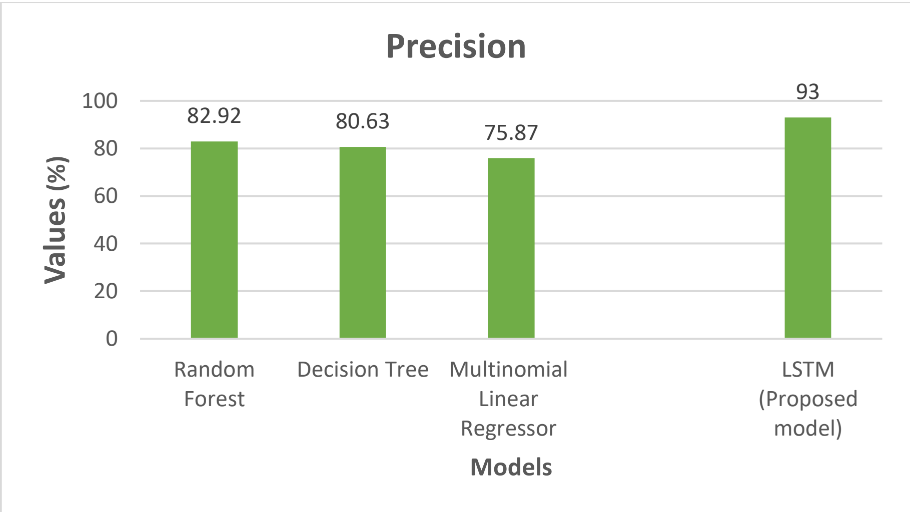

Figure 4.15: Precision values of the proposed and base models

Fig. 4.15 shows the precision values for the proposed and base models. The base models of Random Forest, Decision Tree and Multinomial Linear Regressor have precision values of 82.92, 80.63 and 75.87, respectively. A precision of 93% is achieved by the proposed LSTM model.

CHAPTER- 5

CONCLUSION

This chapter presents the findings and recommendations for the future relating to the analysis of opinions expressed by customers in their reviews.

The goal of Twitter sentiment analysis is to ascertain general tone of a tweet by identifying and extracting subjective information utilizing NLP and ML methods (positive, negative, neutral). This work has a lot of benefits, most notably time savings while solving equations, which will be valuable for many fields whose timetables include time-consuming equation solving and analysis. This study effort concludes the whole suggested study on Twitter spam streaming analysis and grouping of emotions. A feature extraction technique based on a blend of linear regression, random forest, and PCA (principal component analysis) is presented for extracting specific feature sets for boosting classification accuracy and using machine learning classifiers to expose spam tweets. When compared to other existing works, simulation outcomes suggest that proposed work has a higher detection ratio. When employing a vast quantity of data, the results achieved in this suggested study show a very big difference in terms of F1 score, recall. precision and accuracy when compared to other classifiers. Furthermore, this hybrid technique is applied for sentiment classification of tweets with modest alterations in the suggested algorithm in terms of negative & positive tweets, resulting in good classification accuracy Sentiment analysis on Twitter is a subfield of text and opinion mining. It analyses tweet emotions and trains a machine-learning model to measure its accuracy so we can utilize it in future. Sentiment detection, text pre-processing, Data collection, sentiment classification, model training, and testing are all a part of the process. Over past decade, advancements in this area of study have led to model efficiencies of 85-90 percent. Diverse information is still missing, however. It also has a lot of problems in practical use due to slang and abbreviations. The model's applicability to areas other than the one under discussion has not yet been thoroughly examined. Thus, sentiment analysis offers tremendous potential for growth in the years to come.

REFERENCE

Abid, F. *et al.* (2019) ‘Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter’, *Future Generation Computer Systems*, 95, pp. 292–308. doi: https://doi.org/10.1016/j.future.2018.12.018.

Ahmed, M. *et al.* (2021) ‘Sentiment Analysis on Twitter using Ordinal Regression’, in *2021 International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON 2021*. doi: 10.1109/SMARTGENCON51891.2021.9645751.

Ahuja, V. and Shakeel, M. (2017) ‘Twitter Presence of Jet Airways-Deriving Customer Insights Using Netnography and Wordclouds’, *Procedia Computer Science*, 122, pp. 17–24. doi: https://doi.org/10.1016/j.procs.2017.11.336.

Alahmadi, D. H. and Zeng, X. J. (2016) ‘Twitter-based recommender system to address cold-start: A genetic algorithm based trust modelling and probabilistic sentiment analysis’, in *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*. doi: 10.1109/ICTAI.2015.149.

Alicioglu, G., Sun, B. and Ho, S. S. (2020) ‘Assessing Accident Risk using Ordinal Regression and Multinomial Logistic Regression Data Generation’, in *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. doi: 10.1109/IJCNN48605.2020.9207105.

Alkhaldi, S. *et al.* (2020) ‘Twitter Sentiment Analysis on Activities of Saudi General Entertainment Authority’, in *ICCAIS 2020 - 3rd International Conference on Computer Applications and Information Security*. doi: 10.1109/ICCAIS48893.2020.9096738.

Alsalman, H. (2020) ‘An Improved Approach for Sentiment Analysis of Arabic Tweets in Twitter Social Media’, in *ICCAIS 2020 - 3rd International Conference on Computer Applications and Information Security*. doi: 10.1109/ICCAIS48893.2020.9096850.

Appel, O. *et al.* (2016) ‘A hybrid approach to the sentiment analysis problem at the sentence level’, *Knowledge-Based Systems*. doi: 10.1016/j.knosys.2016.05.040.

Arif, S. A. and Binte Hossain, T. (2021) ‘Opinion Mining of Customer Reviews Using Supervised Learning Algorithms’, in *2021 5th International Conference on Electrical Information and Communication Technology (EICT)*, pp. 1–6. doi: 10.1109/EICT54103.2021.9733435.

Aydogan, E. and Ali Akcayol, M. (2016) ‘A comprehensive survey for sentiment analysis tasks using machine learning techniques’, in *Proceedings of the 2016 International Symposium on INnovations in Intelligent SysTems and Applications, INISTA 2016*. doi: 10.1109/INISTA.2016.7571856.

Barnaghi, P., Ghaffari, P. and Breslin, J. G. (2016) ‘Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment’, in *Proceedings - 2016 IEEE 2nd International Conference on Big Data Computing Service and Applications, BigDataService 2016*. doi: 10.1109/BigDataService.2016.36.

Biradar, S. H., Gorabal, J. V and Gupta, G. (2021) ‘Machine learning tool for exploring sentiment analysis on twitter data’, *Materials Today: Proceedings*, 56, pp. 1927–1934. doi: https://doi.org/10.1016/j.matpr.2021.11.199.

Dara, S. and Tumma, P. (2018) ‘Feature Extraction By Using Deep Learning: A Survey’, in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1795–1801. doi: 10.1109/ICECA.2018.8474912.

Gamal, D. *et al.* (2019) ‘Implementation of Machine Learning Algorithms in Arabic Sentiment Analysis Using N-Gram Features’, *Procedia Computer Science*, 154, pp. 332–340. doi: https://doi.org/10.1016/j.procs.2019.06.048.

Gandhi, R. (2018) ‘Naive Bayes Classifier’.

de Godoi Brandão, J. and Calixto, W. P. (2019) ‘N-Gram and TF-IDF for Feature Extraction on Opinion Mining of Tweets with SVM Classifier’, in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, pp. 1–5. doi: 10.1109/IDAP.2019.8875900.

Hamza, M. and Gupta, S. (2022) ‘a Comparison of Sentimental Analysis Algorithms on Twitter Data Using Machine Learning’.

Van Houdt, G., Mosquera, C. and Nápoles, G. (2020) ‘A review on the long short-term memory model’, *Artificial Intelligence Review*, 53(8), pp. 5929–5955. doi: 10.1007/s10462-020-09838-1.

Jain, A. P. and Dandannavar, P. (2017) ‘Application of machine learning techniques to sentiment analysis’, in *Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016*. doi:

10.1109/ICATCCT.2016.7912076.

Jasneet Kaur, D. N. D. (2018) ‘A survey on sentiment analysis algorithms and techniques’, *Journal of Advanced Research in Dynamical and Control Systems*, 6(2), pp. 2320–2882. doi: 10.5373/JARDCS/V12SP4/20201676.

Jayakody, J. P. U. S. D. and Kumara, B. T. G. S. (2021) ‘Sentiment analysis on product reviews on twitter using Machine Learning Approaches’, in *2021 International Conference on Decision Aid Sciences and Application, DASA 2021*. doi: 10.1109/DASA53625.2021.9682291.

Jumadi *et al.* (2016) ‘Opinion mining on Twitter microblogging using Support Vector Machine: Public opinion about State Islamic University of Bandung’, in *2016 4th International Conference on Cyber and IT Service Management*, pp. 1–6. doi: 10.1109/CITSM.2016.7577569.

Juneja, P. and Ojha, U. (2017) ‘Casting online votes: To predict offline results using sentiment analysis by machine learning classifiers’, in *8th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2017*. doi: 10.1109/ICCCNT.2017.8203996.

Katsurai, M. and Satoh, S. (2016) ‘Image sentiment analysis using latent correlations among visual, textual, and sentiment views’, in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. doi: 10.1109/ICASSP.2016.7472195.

Kaur, P. (2022) ‘Sentiment analysis using web scraping for live news data with machine learning algorithms’, *Materials Today: Proceedings*, 65, pp. 3333–3341. doi: https://doi.org/10.1016/j.matpr.2022.05.409.

Khan, H. U. *et al.* (2021) ‘Twitter trends: A ranking algorithm analysis on real time data’, *Expert Systems with Applications*, 164, p. 113990. doi: https://doi.org/10.1016/j.eswa.2020.113990.

Leow, E. K. W., Nguyen, B. P. and Chua, M. C. H. (2021) ‘Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimisation’, *Expert Systems with Applications*, 179, p. 115060. doi: https://doi.org/10.1016/j.eswa.2021.115060.

Lin, H. T. and Li, ng (2012) ‘Reduction from cost-sensitive ordinal ranking to weighted binary classification’, *Neural Computation*. doi: 10.1162/NECO_a_00265.

Liu, C. L. *et al.* (2012) ‘Movie rating and review summarization in mobile environment’, *IEEE*

Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews. doi: 10.1109/TSMCC.2011.2136334.

Liu, J. and Wang, R. (2021) 'Research performance prediction for scientists using LSTM neural networks', *CEUR Workshop Proceedings*, 2871, pp. 194–201.

Medhat, W., Hassan, A. and Korashy, H. (2014) 'Sentiment analysis algorithms and applications: A survey', *Ain Shams Engineering Journal*. doi: 10.1016/j.asej.2014.04.011.

Mishra, P., Rajnish, R. and Kumar, P. N. (2017) 'Sentiment analysis of Twitter data: Case study on digital India', in *2016 International Conference on Information Technology, InCITe 2016 - The Next Generation IT Summit on the Theme - Internet of Things: Connect your Worlds*. doi: 10.1109/INCITE.2016.7857607.

Mohammad, S. M. and Kiritchenko, S. (2015) 'Using hashtags to capture fine emotion categories from tweets', *Computational Intelligence*. doi: 10.1111/coin.12024.

Mohammed, M., Khan, M. B. and Bashie, E. B. M. (2016) *Machine learning: Algorithms and applications, Machine Learning: Algorithms and Applications*. doi: 10.1201/9781315371658.

Mridula, A. and Kavitha, C. R. (2018) 'Opinion Mining and Sentiment Study of Tweets Polarity Using Machine Learning', in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 621–626. doi: 10.1109/ICICCT.2018.8473079.

Nennuri, R. *et al.* (2021) 'Twitter Sentimental Analysis based on Ordinal Regression', in *Journal of Physics: Conference Series*. doi: 10.1088/1742-6596/1979/1/012069.

Pavel, A. *et al.* (2017) 'Using Short URLs in Tweets to Improve Twitter Opinion Mining', in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 965–970. doi: 10.1109/ICMLA.2017.00-28.

Pavitha, N. *et al.* (2022) 'Movie recommendation and sentiment analysis using machine learning', *Global Transitions Proceedings*, 3(1), pp. 279–284. doi: https://doi.org/10.1016/j.gltp.2022.03.012.

El Rahman, S. A., AlOtaibi, F. A. and AlShehri, W. A. (2019) 'Sentiment Analysis of Twitter Data', in *2019 International Conference on Computer and Information Sciences (ICCIS)*, pp. 1–4. doi: 10.1109/ICCISci.2019.8716464.

Rashid, M. *et al.* (2020) ‘Novel machine learning approach for sentiment analysis of real time twitter data with apache flume’, in *PDGC 2020 - 2020 6th International Conference on Parallel, Distributed and Grid Computing*. doi: 10.1109/PDGC50313.2020.9315782.

Rathi, M. *et al.* (2018) ‘Sentiment Analysis of Tweets Using Machine Learning Approach’, in *2018 11th International Conference on Contemporary Computing, IC3 2018*. doi: 10.1109/IC3.2018.8530517.

Saad, A. I. (2020) ‘Opinion Mining on US Airline Twitter Data Using Machine Learning Techniques’, in *2020 16th International Computer Engineering Conference (ICENCO)*, pp. 59–63. doi: 10.1109/ICENCO49778.2020.9357390.

Saad, S. E. and Yang, J. (2019) ‘Twitter Sentiment Analysis Based on Ordinal Regression’, *IEEE Access*. doi: 10.1109/ACCESS.2019.2952127.

Sánchez-Monedero, J. *et al.* (2013) ‘Exploitation of pairwise class distances for ordinal classification’, *Neural Computation*. doi: 10.1162/NECO_a_00478.

Sarker, I. H. *et al.* (2020) ‘Cybersecurity data science: an overview from machine learning perspective’, *Journal of Big Data*. doi: 10.1186/s40537-020-00318-5.

Shamantha Rai, B., Shetty, S. M. and Rai, P. (2019) ‘Sentiment analysis using Machine learning classifiers: Evaluation of performance’, in *2019 IEEE 4th International Conference on Computer and Communication Systems, ICCCS 2019*. doi: 10.1109/CCOMS.2019.8821650.

Sidey-Gibbons, J. A. M. and Sidey-Gibbons, C. J. (2019) ‘Machine learning in medicine: a practical introduction’, *BMC Medical Research Methodology*. doi: 10.1186/s12874-019-0681-4.

Singh, S., Kumar, K. and Kumar, B. (2022) ‘Sentiment Analysis of Twitter Data Using TF-IDF and Machine Learning Techniques’, in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, pp. 252–255. doi: 10.1109/COM-IT-CON54601.2022.9850477.

Srivastava, S., Singh, J. P. and Mangal, D. (2020) ‘Time and Domain Specific Twitter Data Mining for Plastic Ban based on Public Opinion’, in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 755–761. doi: 10.1109/ICIMIA48430.2020.9074935.

Sufi, F. K. (2022) 'Identifying the drivers of negative news with sentiment, entity and regression analysis', *International Journal of Information Management Data Insights*, 2(1), p. 100074. doi: https://doi.org/10.1016/j.jjimei.2022.100074.

Tu, H. H. and Lin, H. T. (2010) 'One-sided support vector regression for multiclass cost-sensitive classification', in *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*.

Ullah, M. A. *et al.* (2020) 'An algorithm and method for sentiment analysis using the text and emoticon', *ICT Express*, 6(4), pp. 357–360. doi: https://doi.org/10.1016/j.icte.2020.07.003.

Verwaeren, J., Waegeman, W. and De Baets, B. (2012) 'Learning partial ordinal class memberships with kernel-based proportional odds models', *Computational Statistics and Data Analysis*. doi: 10.1016/j.csda.2010.12.007.

Windha Mega, P. D. and Haryoko (2019) 'Optimization of parameter support vector machine (SVM) using genetic algorithm to review go-jek's services', in *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2019*. doi: 10.1109/ICITISEE48480.2019.9003894.

Wongkar, M. and Angdresey, A. (2019) 'Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter', in *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*. doi: 10.1109/ICIC47613.2019.8985884.

You, Q. *et al.* (2015) 'Joint visual-Textual sentiment analysis with deep neural networks', in *MM 2015 - Proceedings of the 2015 ACM Multimedia Conference*. doi: 10.1145/2733373.2806284.

Zervoudakis, S. *et al.* (2021) 'OpinionMine: A Bayesian-based framework for opinion mining using Twitter Data', *Machine Learning with Applications*, 3, p. 100018. doi: https://doi.org/10.1016/j.mlwa.2020.100018.

Zhao, H. *et al.* (2021) 'A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach', *Information Processing & Management*, 58(5), p. 102656. doi: https://doi.org/10.1016/j.ipm.2021.102656.