Advanced International Journal for Research
E-ISSN: 3048-7641 editor@aijfr.com
Volume 6, Issue 4 (July-August 2025)

Textual Sentiment Evaluation

Dr. Mohan Ravat[1]

[1]Professor, National Institute of Technology, Allahabad, India

Abstract

In today's fast-paced world, where life continues to grow more complex, staying updated with the latest trends and developments has become increasingly difficult. Sentiment Analysis offers valuable insights that can assist in decision-making by capturing how people feel about various topics. It involves monitoring human emotions and mapping opinions on specific subjects. Social media platforms have given people an open space to express their thoughts, opinions, and emotions on a wide range of issues, generating a vast amount of data that can be analyzed to gauge public sentiment. Text-based Sentiment Analysis evaluates emotions based on the text collected from various sources.

Businesses are now recognizing the value of providing a positive user experience. By utilizing Sentiment Analysis, companies can understand customer opinions about their products and services, which in turn can be used to enhance the customer experience.

The importance of Sentiment Analysis has surged in recent years. It helps in grasping public opinion by analyzing large datasets from diverse sources, using advanced techniques like Deep Learning Algorithms. Our goal is to develop a sentiment analysis model to track public opinion across various topics. For this study, we focus on data from social media, particularly YouTube comments. We aim to analyze the text gathered from these comments to understand the emotional tone, public sentiment, and feedback by collecting and processing the data, then applying algorithms for language analysis.

Once deployed, such models can prove to be highly useful and have a wide range of applications. These include understanding public opinion on political issues, tracking brand mentions, improving customer service, and more. Sentiment Analysis is a rapidly expanding field, though it remains challenging, as it requires understanding the complexities of human language, such as irony and context. Various approaches have been explored in previous research. Despite its challenges, Sentiment Analysis can be a powerful tool to understand the world better and help make well-informed decisions.

Keywords: Sentiment Analysis, Deep Learning, NLP, detection, CNN.

1. INTRODUCTION

In the present era, millions of users exchange their views, ideas, expressions, feelings, and opinions on social media like You-Tube, Twitter and Facebook. These websites have become a source of unlimited data as millions of users share their thoughts, beliefs and opinions openly.

Advanced International Journal for Research
E-ISSN: 3048-7641 editor@aijfr.com
Volume 6, Issue 4 (July-August 2025)

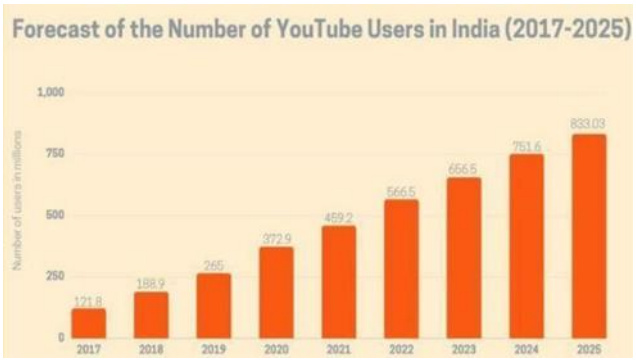

Fig.1: Forecast of the Number of You-Tube Users in India

We can apply Engineering on the collected data, convert it into very useful information, gather the zest of the data, draw the patterns and use the information to grow the businesses. Sentiment is the reflection of a thought, emotion, or judgment induced by feelings.

Whereas, Sentiment analysis is the procedure of analyzing textual data and determine if the sentiment of the message is neutral, negative, or positive. Various Sentiment Analysis models can evaluate the text and accordingly determine the customer's perception towards the topic. Therefore, it has become one of the most profound research areas with the increasing growth of social media on the web.

Sentiment analysis is also referred to as perception mining. These studies reveal people's sentiments towards a particular topic. With the knowledge of word vectors, deep learning algorithms develops speedily in Natural Language Processing(NLP). Therefore, textual sentiment analysis based on Deep Learning has also been broadly studied.

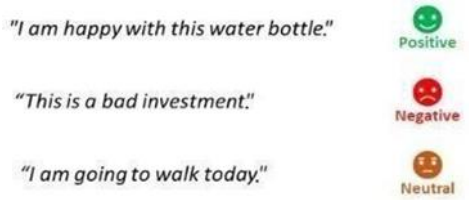

Fig.2: Example of Text-Mapped to Sentiment

The term Deep Learning is broadly used for various machine learning algorithms in regard to different properties of Deep Neural Networks(DNN) and self-directed learning. It has several advantages which include its solid capability to discriminate, and highlight self-learning ability. It's felicitous for unlabeled, high-dimensional, and big data features. Therefore, it is crystal clear that Sentiment Analysis seems to have a support of massive online data featuring a robust fundament. Deep learning techniques, which includes Convolutional Neural networks (CNNs), Recurrent Neural Networks (RNNs), and transformer models, have reshaped Sentiment Analysis by automatically learning intricate patterns and representations from large-scale text data. In this review paper we are going to discuss the methodologies of Sentiment Analysis on textual data using Deep Learning.

1.1 Objective of review

The main objective of this paper is to highlight cutting edge technologies which are used to draw and detect valuable patterns from the raw textual data. It emphasizes on the existing technologies, their shortcomings and the future scope of new evolving algorithms. It also points out the challenges and problems faced while analyzing text word by word.

2. LITERATURE REVIEW

A. Timeline of the reported problem

Earlier people did not have the privilege to share their emotions or thoughts as there was no platform to do so but with the development of technology it has been estimated that 329 million terabytes of data is introduced on the internet everyday. People now have the advantage to share there feelings and emotions using different ways and on different platforms like YouTube, Instagram, Twitter, etc. Well the sentiments of people based on visual data can be easily analyzed but the sentiments of text-based data needs proper analysis due to various factors. During the 1950's, early work on on public opinion analysis began, which laid the foundation of text-based sentiment analysis. Then from 2000 to 2020 sentiment analysis grew by research based on different technologies that were being developed on the way for example, computer linguistics, machine learning, deep learning, support vector machine, etc.

As we have evolved we have observed many challenges in this very field. Firstly there was no data at all and as there was no data the scope of data creation making analysis was limited to that very individual, then when there was proper data the challenge was having proper system for getting feedback from the people, but now when the data is in abundance and feedback is being provided directly or indirectly by the users, analyzing that data and drawing patterns from it has become the biggest challenge. In order to tackle this challenge several algorithms have been developed and mass research has been done. With this very idea Sentiment Analysis on text based data using deep leaning and KNN algorithm has been reviewed.

B. Existing Solutions

Sentiment Analysis is used to map phrases and words to sentiments to accept opinions and emotions from text. It is a significantly powerful tool that can be used to understand public opinion, track user satisfaction and identify treads in social media.

[1]**Lexicon-based Approaches:** These approaches are very common in models where mapping is required. In this dictionary of words or expressions which are mapped with a negative or positive emotion is used.

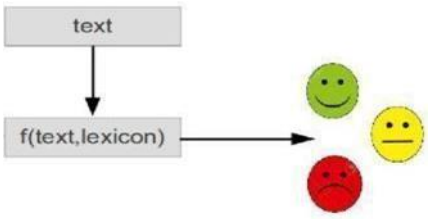

Fig.3: Lexicon-based Approach

For example, the phrase 'I am Happy' is mapped with a positive sentiment and the phrase 'I am Sad' is mapped with a negative sentiment while the phrase 'I am Fine' is mapped with a neutral sentiment.

It calculates the sentimental orientation of the complete text. Orientation can be neutral, positive or negative. Lexicon-based path for textual sentiment analysis is comparatively simple to execute.

[2]**Machine Learning Approaches:** Within the fields of Computer Science and Artificial Intelligence, Machine Leaning is the study in which we design an algorithm, train the model according to the algorithm and deploy the model.

Text-based Sentiment Analysis uses Machine Learning algorithms to train a model on textual data that has been labeled with sentiments like positive, negative or neutral. Once the model is trained the sentiment or emotion of the text under consideration can be predicted.

Machine Learning based approaches are more accuratethan Lexicon-based approaches, but at the same time they can be more complex to execute.

[3]**Deep Learning Approaches:** It is a further extension of Machine Learning in which the model tries to imitate human behavior.

Deep Learning based approaches use neural networks to learn from the textual data. Deep Learning approaches have proven to be the most accurate for sentiment analysis. While using Deep Learning-based Approaches one should be aware that they are computationally very expensive to train, test and deploy. Text based sentiment analysis has been a hot topic of research for many years.

Pioneers In Sentiment Analysis

In the early 1960's a computer program called PARRY which could simulate the behaviour of a paranoid patient. It was developed by Kenneth Colby. PARRY could identify and respond to the emotions expressed by patients such as anger, fear, sadness and happiness.

Kim(2014)

This paper proposed an approach based on Convolutional Neural Network(CNN) for analyzing the sentiment using movie reviews as the data set. It achieved 88.4% accuracy.

Convolutional Neural Networks in machine learning is an algorithm that is generally used for image classification but Kim proved CNNs can also be efficiently and effectively usedfor sentiment analysis of sentences.

Kim's Convolutional Neural Network model is comparatively simple. There was only one convolutional layer with a max pooling layer and a fully connected layer.

Amazon Comprehend

Amazon Comprehend is a service based on NLP (Natural Language Processing). It finds the insights and relationships in the textual data using machine learning. Language of the text can be identified using this. It can extract key phrases, places, brands, people and events.

It can understand wheer the text is positive or negative using tokenization and parts of speech. It can be used for identification of the sentiment of textual data

MonkeyLearn

It is a clould-based platform for sentiment and intent analysis based on NLP(Natural Language Processing). It can perform sentiment and intent analysis. It is used by businesses of all sizes to evaluate and understand customer opinion and track user satisfaction.

MonkeyLearn is an easy to use product and any prior coding knowledge is not required. Along with Sentiment Analysis it is capable of sentiment classification, intensity, aspect identification.

Lexalytics

It provides sentiment and intent analysis to many companies using Software as a Service(SaaS) and cloud-based technology. It was founded in 2003, by Catlin and Marshall. It’s one of the first products was a sentiment analysis engine called

SALIENCE.

Salience was created to help businesses understand the sentiment of their customers. It was one of the first sentiment analysis engine in the market and within no time it became popular.

C. Bibliometric Analysis

Text-based sentiment analysis is a rapidly growing field of research with a wide range of applications in business, government, and academia. sentiment analysis models are used to extract people's opinions and emotions from text data, such as product reviews, social media posts, and customer service interactions.

Bibliometric analysis of sentiment analysis research focuses on the quantitative aspects of the field, such as the growth of the field, the most active researchers and institutions, and the most influential publications.

There has been a lot of research and development in the field of sentiment analysis, as it can prove to be a very beneficial tool for growing businesses. Many scholars, scientists and researchers have tried to optimize the sentiment analysis tools and enhance the accuracy.

D. Review Summary

In the above literature review we have researched on how the data can be implemented, filtered, which technologies and algorithms were previously used in the recent years and which can be used now and we have created a model which is easily going to provide accuracy ignoring all nuances.

E. Problem Definition

Text-based sentiment analysis is a powerful tool for extracting insights from people's opinions and emotions expressed in text. It has a wide range of applications, including product reviews, social media, and customer service. However, existing sentiment analysis models often struggle to accurately capture the nuances of human language like irony and sarcasm, especially when dealing with complex or ambiguous texts.

One of the key challenges in sentiment analysis is the lack of labeled data. Most sentiment analysis models are trained on datasets of labeled reviews, but these datasets are often biased and limited in scope. This can lead to models that are unable to generalize to new data or accurately capture the sentiments of different groups of people.

Another challenge in sentiment analysis is the complexity of human language. Sentiments can be expressed in a variety of ways, including explicit statements, sarcasm, and irony. Existing sentiment analysis models often have difficulty identifying and interpreting these implicit sentiments.
This review paper aims to contribute to the development of more accurate and reliable sentiment analysis models that can be used to extract valuable insights from text data in a variety of domains.

F. Goals/Objectives

Now a days, people can post their honest reviews and feedbacks anywhere directly or indirectly which is very essential for the developer as the developer can analyse and draw patterns based on those reviews and can understand what chances are required and what people are demanding from the him/her.

Therefor our model helps analyzing patterns based on textual data to understand the sentiments of the mass with accuracy.

3. DESIGN AND METHODOLOGIES

3. 1 Dataset Description

The dataset taken is the real-time comments on YouTube videos. We have to pass the link of the video we want to evaluate, the model analyzes the comments and provide us with the output.

3.2 Importing Libraries and Providing Dataset

Firstly, the Python libraries and packages like Pandas, Numpy, Pyplot from Matplotlib, and Seaborn were imported and used for data cleaning and preprocessing.

3.3 Implementing Data Visualization

Various data visualization graphs like Pie-chart, Histogram, Heat-map were plotted to highlight the feature distribution of the data using libraries Matplotlib and Seaborn.

3.4 Data Cleaning

It is the most important step in dealing with problems where we have to analyse a lot of data. We deal with the null values and eradicate insignificant words.

3.5 Implementation

We have used VADER(Valence Aware Dictionary and sEntiment Reasoner). It does not only tell the sentimental orientation of the text but also tells the value of the positive and negative score.

4. IMPLEMENTATION

The data flow diagram and control flow diagram for our model are represented below.

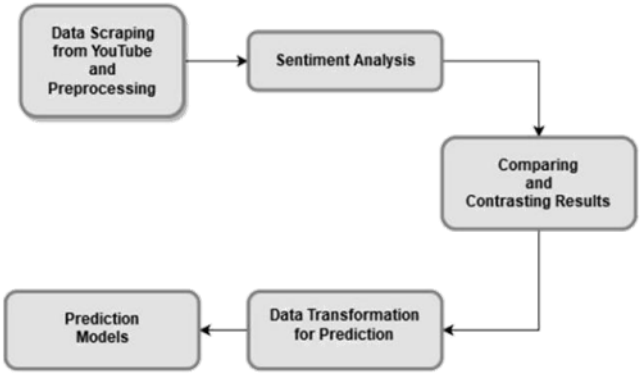

Fig.4: Data Flow Diagram

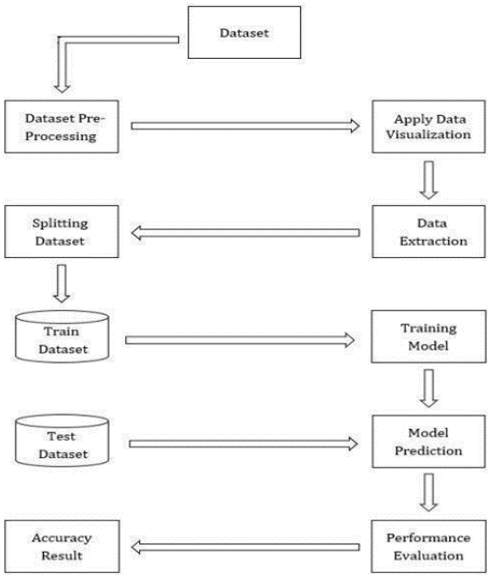

Fig.5: Control Flow Diagram

5. RESULT

Lastly, after working hard and putting in a lot of thoughts into our project we were able to make our model capable enough to accept the dataset very efficiently and import all the libraries required for the successful completion and implementation of sentiment analysis, the dataset then went through the process of cleaning to eliminate all the unimportant data, after the cleaning process the values were passed through the process of sentiment analysis which resulted in the following result categorizing the data as positive, negative and neutral.

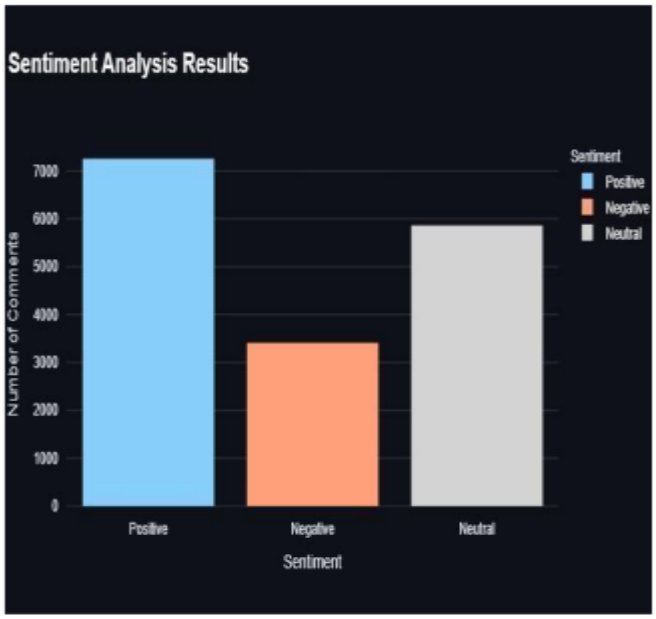

Fig.6: Result Visualized as Bar Graph

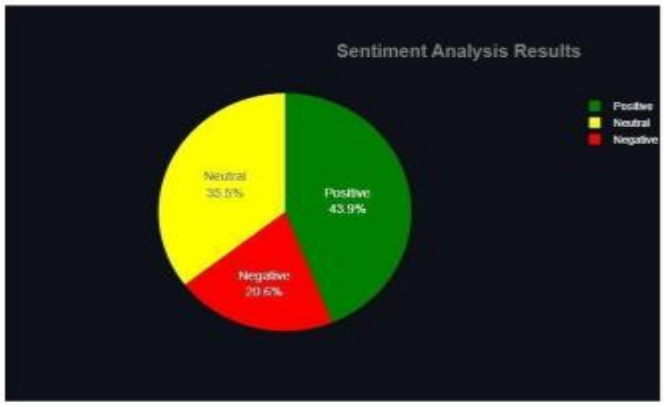

Fig 7. Result Visualized as Pie-chart

6. KNN ALGORITHM

K-Nearest Neighbors (KNN) is a simple but powerful supervised learning algorithm that can be used for both classification and regression tasks. KNN works by finding the K most similar training examples to a new data point and then using the labels of those training examples to predict the label of the new data point.

For classification tasks, KNN simply predicts the most common label among the K-nearest neighbors. For regression tasks, KNN averages the values of the K nearest neighbors.

KNN is a non-parametric algorithm, which means that it does not make any assumptions about the underlying distribution of the data. This makes KNN a versatile algorithm that can be used with a wide variety of data types.

KNN is also a lazy learning algorithm, which means that it does not train a model in the traditional sense. Instead, KNN simply stores the training data and then finds the K nearest neighbors to a new data point at prediction time.

This makes KNN a computationally expensive algorithm, especially when dealing with large datasets. However, there are a number of techniques that can be used to speed up KNN, such as using approximate nearest neighbor search algorithms.

To use KNN for text-based sentiment analysis, we first need to preprocess the text data. This includes cleaning the text, removing stop words, and stemming or lemmatizing the words. Once the text data has been preprocessed, we need to represent it as vectors. This can be done using a variety of methods, such as bag-of-words, TF-IDF, or word embeddings then we need to choose a distance metric. This is the metric that will be used to measure the similarity between two text vectors. Common distance metrics for text-based sentiment analysis include Euclidean distance, cosine similarity, and Jaccard similarity.

Finally, we would need to choose a value for K. This is the number of nearest neighbors that will be used to predict the sentiment of a new text sample. A higher value for K will result in more stable predictions, but it will also be more computationally expensive.

Once we would have chosen a value for K, we can use KNN to predict the sentiment of a new text sample by following these steps:

Finding the K most similar training examples to the new text sample using the distance metric you chose.

Then predict the sentiment of the new text sample based on the sentiments of the K most similar training examples.

7. CONCLUSION AND FUTURE WORK

Sentiment analysis based on the text is a project that helps us to understand people's mentality, attitude and thinking based on the words and phrases they use in their comments. It helps us to understand the motion of the mass towards or against certain things and which in turn can help in understanding and noting down the indirect feedback about certain entities, lookout for an opportunity in the market or anything which becomes clear after analyzing the texts. The main points in the conclusion are:

1. We explored various methodologies employed in sentiment analysis from old lexicon-based approaches to the latest deep learning models.

2. Sentiment Analysis has a huge range of applications from social media monitoring to consumer feedback analysis.

3. There were persistent challenges including handling of sarcasm, ignoring politically spam data.

To summarize this paper serves as a valuable reference for research scholars and scientists. Sentiment Analysis is very crucial in this Data-Driven world, we also highlight the ethical and moral considerations that must be taken care of during its advancement. As the field keeps on evolving this paper appreciates and motivates the ongoing exploration and research work in the branch of Sentiment Analysis.

In addition to the current work we have done, in future we can add more functionalities to our model. This will make it more efficient and useful. Increasing the result array from 3 values to 5 values can be a good change.

[Positive, Neutral, Negative]
=> [Excellent, Positive, Neutral, Negative, Worst] It will help to understand the feedback in a better and more minute manner.

REFERENCES

[1] Basant Agarwal, Richi Nayal, Namita Mittal and Srikant Patnaik- "Deep Learning-Based Approaches for Sentiment Analysis".
[2] Vadim Kagan, Edward Rossini and Demetrios Sapounas - "Sentiment Analysis for PTSD Signals".
[3] Mark J S Keenan - "Advanced Positioning, Flow, And Sentiment Analysis".
[4] Zulfadzli Drus and Haliyana Khalid - "Sentiment Analysis In Social Media And Its Applications".
[5] Vishal A. Kharde and S.S. Sonawane - "Sentiment Analysis of Twitter Data".
[6] Pablo Gamallo, Marcos Garcia- "Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets".
[7] Neethu MS and Rajashree R- "Sentiment Analysis in Twitter using Machine Learning Techniques".
[8] YouTube Data API Overview — Google Developers: https://developers.google. com/youtube/v3/getting-started.
[9] Sentiment Analysis Explained: https://www.lexalytics.com/technology/sentiment-analysis.
[12] Multilayer perceptron : https://en.m.wikipedia.org/wiki/Multilayer_ perceptron.
[13] About Keras Models : https://keras.rstudio.com/articles/about_keras_mode_ls.html.
[14] Susan Li. Time series analysis, visualization forecasting with lstm, May 2019.
[15] Basant Agarwal, Richi Nayal, Namita Mittal and Srikant Patnaik- "Deep Learning-Based Approaches for Sentiment Analysis".
[16] Vadim Kagan, Edward Rossini and Demetrios Sapounas - "Sentiment Analysis for PTSD Signals".
[17] Mark J S Keenan - "Advanced Positioning, Flow, And Sentiment Analysis".

[18] Zulfadzli Drus and Haliyana Khalid - “Sentiment Analysis In Social Media And Its Applications”.
[19] Vishal A. Kharde and S.S. Sonawane - “Sentiment Analysis of Twitter Data”.
[20] Pablo Gamallo, Marcos Garcia- “Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets”.
[21] Neethu MS and Rajashree R- “Sentiment Analysis in Twitter using Machine Learning Techniques”.
[22] YouTube Data API Overview — Google Developers: https://developers.google. com/youtube/v3/getting-started.
[23] Sentiment Analysis Explained: https://www.lexalytics.com/technology/sentiment-analysis.