International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

OPEN ACCESS

Available Online at :www.ijsrcseit.com

doi : https://doi.org/10.32628/IJSRCSEIT

Sentiment Analysis on YouTube Using Hybrid Model Deep Learning (LSTM)

Kumar Kundan[1], Vinod Mahor[2]

[1]M Tech Scholar, [2]Associate Professor

Department of Computer Science & Engineering, Millennium Institute of Technology and Science Bhopal, Madhya Pradesh, India

ARTICLEINFO

Article History:

Accepted: 10 Nov 2023

Published: 22 Nov 2023

Publication Issue

Volume 9, Issue 6

November-December-2023

Page Number

351-360

ABSTRACT

At the same time as information and communication technology (ICT) is expanding at the quickest rate, the amount of online material that is available on social media platforms is growing at an exponential rate. The study of sentiment derived from online evaluations is attracting the interest of researchers from a variety of organizations, including academic institutions, government agencies, and commercial businesses. Within the fields of Machine Learning (ML) and Natural Language Processing (NLP), sentiment analysis has emerged as a prominent area of study interest. In order to get outstanding outcomes in the field of sentiment analysis, Deep Learning (DL) approaches are now being utilized. A BiLSTM + WV model, which stands for a hybrid convolutional neural network and long short-term memory, was presented for the purpose of sentiment analysis in this study. In order to obtain results, our suggested model is now being utilized in conjunction with dropout, max pooling, and batch normalization. The datasets of airline sentiment on Twitter and Airline quality were subjected to experimental analysis. For this purpose, we utilized the Keras word embedding method, which transforms textual data into vectors of numeric values. This method ensures that words that are related to one another have short vector distances between them. In order to evaluate the effectiveness of the model, we computed a number of metrics, including accuracy, precision, recall, and F1-measure, among others. When it comes to sentiment analysis, these parameters for the suggested model are superior to those of the traditional machine learning models. 89.80% accuracy in sentiment analysis is demonstrated by our results analysis, which reveals that the suggested model surpasses competing models.

Keywords- DL, CNN, ML ANOVA ML, BiLSTM

I. INTRODUCTION

Human behavior is heavily influenced by emotions and views because we frequently base our choices on the knowledge and opinions of others. Since everyone's opinions are subjective and influenced by their own experiences and views, understanding them might help one see things from a wider angle. Politicians, businesses, and organizations have always been interested in public opinion. The collective viewpoint of people can be used to anticipate election results and to learn more about potential future trends and issues. Also, it aids in gathering public opinion about goods and services, which aids marketing teams in developing marketing strategies, optimizing the manufacturing of current goods or developing new ones, and enhancing customer service [1]. Consequently, it is crucial to recognize sentiment in numerous sectors.

At first, opinions were manually analyzed after being gathered through surveys or questionnaires. But, as internet usage grew, more people began to voice their opinions on the internet. The rise of social media platforms in recent years has given people a greater choice of options to exchange information and express their thoughts. Rich data sources, including blogs, discussion forums, reviews, comments, and microblogging services like Twitter and Facebook, are made up of audio recordings, videos, photos, and thoughts.

YouTube is one of the most widely used social media sites in the world today1. It enables users to browse and comment on videos submitted by other users or share their own videos. By giving a video a favourable or negative rating, a user can express support for or rejection of the film. Moreover, users can remark on videos with text in the form of comments. YouTube is a medium that text classification and categorization researchers are interested in because each video can include thousands of comments.

The construction of a YouTube comment dataset, the creation of a sentiment classifier, and sentiment analysis of the dataset are the main topics of this dissertation. The words sentiment analysis, text classification, and text sentiment classification are defined in the next section to assist readers better understand the principles of this subject

Sentiment analysis, often known as opinion mining, is a natural language processing (NLP) method for identifying the polarity of sentiment in textual data. It has grown to be one of the most well-liked NLP research areas over time [2]. It examines how people feel and behave towards many things, including products, digital information, businesses, people, and events [2]. Large amounts of opinion data take a long time to evaluate, but they can give an in-depth insight of general sentiment. It is time-consuming and difficult to review a lot of texts, and some texts may be extensive and complex and express a variety of sentiments, making it difficult to grasp the overall sentiment fast. The current character limit for comments on YouTube is 10,000, which is a significant number and can lead to lengthy messages. This volume of data logically necessitates the use of tools to streamline and automate the classification and sentiment analysis of text.

Text classification involves grouping texts into different categories. It is used in a variety of fields, including the organization of news stories, the opinion mining of product reviews, spam filtering, the organization of documents in digital libraries, including books and social feeds, and more [3]. Text can be categorized in a variety of ways, including by topic, whether it satisfies a certain set of requirements, and sentiment [3].

The method of identifying the overall sentiment of a text is known as text sentiment classification [4].

Binary classification or multi-label classification might be the outcome of this issue. Binary classification generates two outputs, positive and negative, whereas multi-label classification generates more than two outputs, such as very negative, negative, neutral, positive, and very positive [4]. The majority of the time, deep learning-based methods or conventional machine learning algorithms are used to handle text sentiment categorization challenges. Text and sentiment classification tasks have been accomplished well by convolutional neural networks (CNN) and recurrent neural networks (RNN), such as Long Short-Term Memory (LSTM) and their bidirectional variations, such Bidirectional LSTM (BiLSTM) [5].

II. RELATED WORK

Traditionally, supervised machine learning methods such as Naive Bayes (NB), Support Vector Machine (SVM), or Logistic Regression (LR) have been used to tackle the text sentiment categorization problem [5].

Pang et al. [9] published one of the first articles to suggest the use of machine learning for text classification on online platforms based on sentiment. On the IMDb movie review dataset, Pang et al. assessed the Naïve Bayes, Maximum Entropy (ME), and Support Vector Machine classifiers [8]. SVM was able to achieve a respectable accuracy of around 83%. Since then, social media has become increasingly popular, and today, millions of people express their ideas online. There is interest in attaining automatic emotion classification due to the large amount of sentimental information that is easily accessible on the majority of social networking sites.

Twitter has been one of the social media platforms most thoroughly studied thus far in terms of sentiment research. Short communications, or tweets, are a popular way for people to express themselves on Twitter. The users are compelled to present their ideas in a succinct but direct manner due to the length restriction. As a result, the data is sentiment-rich and suited for NLP applications. While Neethu and Rajasree evaluated and evaluated the effectiveness of SVM, NB, and ME algorithms for classifying tweets about electronic products, reaching Agarwal et al. attained 75% accuracy while SVM and ME [11] yielded 93% accuracy with Binary classification using SVM for non-domain-specific data [10].A small amount of research has also been conducted using YouTube datasets, including the classification the comments of popular Arabic YouTube videos and SVM with the function of radial bases, which achieved an 0.88 F1-score [13] and achieved accuracy of 95.3% when classifying cooking videos on YouTube.

While these methods might result in excellent text sentiment prediction, they also have drawbacks. For example, deep learning has been outperforming classical machine learning methods in cross-lingual or cross-domain data sets [14].

Deep learning for text sentiment categorization

Owing to the limitations of conventional methods, researchers have begun to investigate more cutting-edge approaches to handling textual data. Deep learning has been used more frequently and successfully for natural language processing jobs due to its performance in these areas.

A subset of machine learning known as "deep learning" transforms the transformation of data from a low level to a higher level by using multiple non-linear layers in computational models [15]. Deep learning is frequently used for feature extraction, analysis, and classification [16]. The most popular deep learning designs include recurrent neural networks, feed-forward neural networks, and convolutional neural networks. Sentiment categorization is one NLP problem where CNNs, RNNs, and its variants have performed well [5].

RecurrentNeuralNetworks

An RNN is a particular kind utilising deep learning networks remembers details regarding input data and takes them into account when processing the following input [5]. Dependency capturing is the term used for this [5]. A series of data is used as the input for RNNs. For instance, each word in a sentence in the correct order would be the input for sentiment analysis. By analyzing each word separately and storing the "state vector," or history of the previous words in the phrase, it is possible to capture dependencies [15]. Figure 2.1 depicts the "unfolded" action-plan of an RNN, in which the data from one input is sent into the hidden layer along with the next input Xn. The preceding input X(n1) determines the output Yn. unfurled over time.

RNNs are not need a fixed the length of is a variable the input and can accommodate inputs of any size without the model getting larger since after processing each input from a particular sequence, the input from the past will be added to the memory. A single output can be generated for many inputs because the output length can also change. While tackling text-related problems, RNNs can predict the following word in the text thanks to these characteristics.

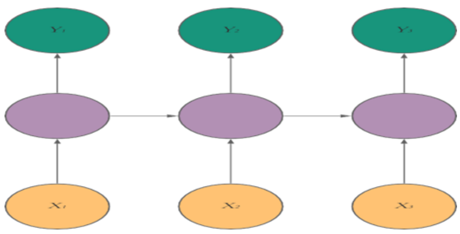

Figure 2.1 shows a recurrent neural network visually.

Although RNNs are great at solving problems involving sequential data, they theoretically have problems with bursting (too high) and disappearing (too low) gradients [17, 18]. The use of a Long Short-Term Memory alleviates these issues, particularly vanishing gradient [18].

CNN and LSTM have also been merged in other research to address the sentiment classification issue. Combining CNN and LSTM, Hassan and Mahmood [31] fed it word vectors that had already been learned. They used LSTM and convolutional layers to create a single model that outperformed CNNs, NB, and SVM algorithms while achieving the greatest accuracy for binary classification of 88.3%.

In order to merge Two LSTM layers, CNN and LSTM, were placed added to the CNN layer of by Huang et al[32] . embedding words model, which was trained beforehand with their own dataset. With an accuracy of 87.2%, their suggested model outperformed single-layer LSTM, standalone CNN, SVM, and a combination of CNN and LSTM.

Convolutional neural networks were developed by Cunha et al. [33] to categorize the sentiment of comments made on Brazilian political YouTube videos. By classifying user comments into three separate the user's perceptions of the information in categories producer and the video, the manner in which the subject is covered, and the relevance of the video they were being able to assess the efficacy of the suggested classifier. Theirclassifier achieved an accuracy of 84% when classifying comments from the first category.

With the aforementioned IMDB data, a sentiment classification method based on LSTMs has also been investigated. For sentiment detection, Yenter and Verma [34] integrated numerous CNN-LSTM kernels. The pre-trained word embedding was not used. Their multi-kernel CNN-LSTM model had an accuracy rate of 89.5%.

The effectiveness of LSTM, CNN, and CNN-LSTM among other deep learning approaches was examined by Mathapati et al. [35]. They utilized NB as a basis algorithm, but all deep learning methods outperformed it. Model CNN-LSTM outperformed the additional models for deep learning when trained on pre-learned word vectors, reaching an accuracy rate of 88.3%.

Building classifiers for a specific domain is typically the focus in the area of sentiment analysis research. These websites primarily contain reviews of different kinds, such as those for products [27, 28] or films [27, 30, 34]. Internet lingo as well as acronyms common comments on YouTube and frequently add to the majority the comment; nevertheless, this type of writing is rarely found in data from product or movie reviews. Reviews are typically written in a more formal style with formal language

III.PROPOSED WORK

The data collecting, cleansing, labelling, and processing study technique is covered in the section that follows. After that, it is explained how the model's application and assessment are carried out. Figure 3.6 shows the process in an illustration. The reliability, validity, and ethical issues surrounding the selected methodologies are discussed as the section comes to a close.

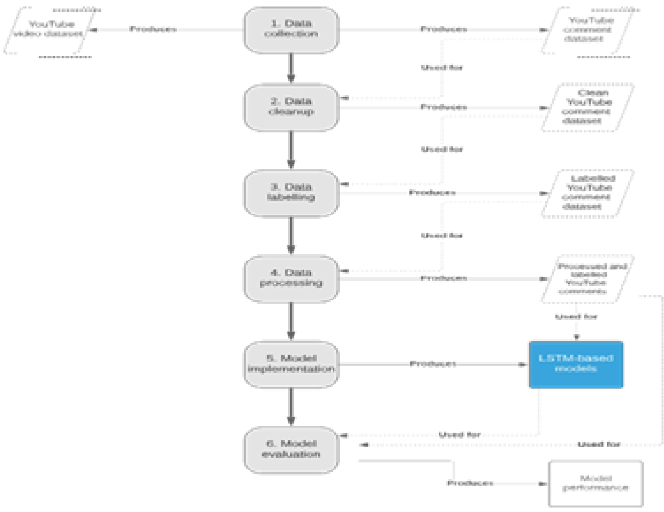

Figure3.1:Theresearchmethodology.

Datacollection

Gathering the information that will be employed in the dissertation's practical section's later stages is the initial step. Two categories of data must be gathered: YouTube video statistics and YouTube video comments, in order to amass sufficient data to meet the study questions. The section that follows provides a description of the approach used to collect this data..

A list of videos needs to be defined before the YouTube video comments can be gathered. These videos gathered based on the date of the upload and the range of classes because it is unclear whether the general emotion of the video comments differs from that of the video ratings or has any relationship to the number of views. A preliminary list of videos is compiled using a combination of user-made playlists of popular videos and a general public collection of YouTube video statistics that are in demand [36]. The initial video collection is then honed using the following inclusion standards:

- A video's access must be public (not unlisted or private);
- A video must support ratings and comments;
- A video needs to be published between 2016 and 2018.
- A video must have at least 100,000 views and 2,000 comments in order to qualify.
- Its purpose cannot be to solicit feedback from the viewers, and its creator must not be pressuring them to give it a low rating.
- It's possible that the video's creator is still alive.

To guarantee that a video has all the necessary data, It must be accessible to the public, allow comments and ratings, and include a specific number comments and opinions. The date requirements make sure the video is statistics (views, ratings), are not being actively changed, and they provide a time limit on the collection of videos in case the methods for counting views and ratings on the videos change. This approach aims to lessen the impact on the data of modifications to platform implementation specifics on such as the view count algorithms employed on YouTube. The video is then scrutinized to remove any that prompt viewers to respond in the comments section or those in which the author urges viewers to give the film a poor rating in order to prevent data manipulation [37]. The developer of the chosen video is also checked to make sure they have not recently developed a sickness, as this could lead to the most recent comments focusing on the author's demise rather than the video's substance.

Video gathering is done manually because YouTube does not offer the ability to search for videos based on the year they were uploaded or the quantity of comments and views.

The selection of videos is important since it influences how the YouTube Comment dataset is later built. There must be an equal number of comments with positive and negative sentiment in the final YouTube Comment dataset. The LSTM-based model may favour one of the sentiments if it makes up a larger proportion of either sentiment.

A video collection of 49 videos is the outcome of the video selection using the aforementioned criteria, which was completed the 2nd of February 2021. The chosen videos fall under the following groups:

a) Technology & Science
b) Entertainment
c) People and Blogs
d) How to & Style
e) Music
f) News and Politics
g) Gaming
h) Comedy
i) Film and Animation
j) Sports

The process of collecting videos is completed by preserving the following data for each one: title, video ID (or URL), votes for and against, and the amount of upvotes, views, date of upload, and number of comments.

Datacleanup

The dissertation's next step entails cleaning the data with the goal of creating a clean dataset that can then be used for labelling the data. To make the labelling process easier and get the data ready for the final YouTube comment dataset, the data cleanup procedure entails finding, removing, or correcting the collected comment texts.

Data labelling

The data labelling can be initialized with cleaned data. In order to use the data for model training and evaluation, labelling must be done. Each comment is read, and a sentiment is manually assigned as part of the data labelling process.

Data processing

To guarantee that sentiment classification model works properly, the comment data must be standardized and condensed. Just patterns in text and text reveal each comment's sentiment and are clear to the example should remain in the remaining data. Processing refers to the standardization and simplification of comments

Model implementation

Defining a model's architecture and then training it using the validation data and the training subset of the data constitute the model implementation. The YouTube comment dataset is divided into three sections for this step: training, validation, and testing. Four LSTM-based models' architectures are also created.

Three sub-sets of the labelled and processed YouTube comment dataset are created, with a ratio of 70:10:20. 80% of the dataset is allocated to the training subset, 10% to the validation subset, and 20% to the testing subset.

The ratio of correct forecasts to total predictions is called accuracy, and it can be calculated as follows:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

The effectiveness of the model's labelling of the sample as positive when it is actually negative is shown by Precision [44]:

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

Recall reveals the proportion of expected positives that turned out to be positive [45]:

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

A weighted average of recall and precision is the F1 score [46]. The value ranges from 0 to 1, and higher numbers denote improved efficiency:

$$\text{F1score} = \frac{2\times(\text{Precision}\times\text{Recall})}{(\text{Precision}+\text{Recall})}$$

Confusion matrix shows how many Predictions of the TP, TN, FP, and FN were made.

IV.RESULT ANALYSIS

This section introduces the video statistics before describing the ultimate dataset of labelled YouTube comments (Section 5.1) the Section 5.2. The results of the performance of the model on the dataset of YouTube comments are then shown (Section 5.3). Finally, performance of the model on the dataset of IMDB movie reviews is shown (Section 5.4).

4.1 YouTubecommentdataset

A total of 19,951 comments were tagged throughout the dissertation, and the final dataset had 4,315 comments with positive sentiment, 5,144 with negative sentiment, and the remaining comments with neutral sentiment. The majority of comments are neutral, but only comments with Positive and negative emotions are considered for the evaluation and training of models. Figure 5.1 displays the sentimental distribution in the dataset of YouTube comments.

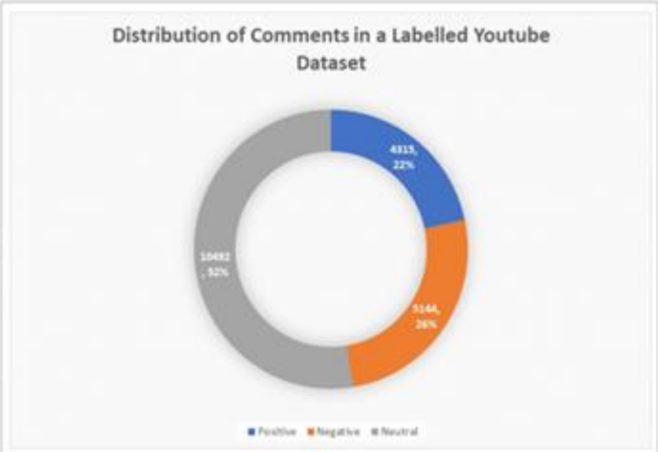

Figure 5.1.The finalized YouTube comment dataset's sentiment distribution is shown

4.2 YouTubevideostatistics

The video statistics were compiled at the outset of the dissertation. These data points include the video ID, the quantity of favorable and unfavorable ratings, the number of comments, and the number of views. A table comprising rating counts, both positive and negative, view counts, and anticipated sentiment was created because the goal of this dissertation is to investigate the connection between videos sentiment based on comments and different user input. The emotion value is defined as being positive if the value is equal to or greater than 0.5. Otherwise, they are not feeling good. Five sample rows from the sentiment forecast table and video statistics are shown in Table 5.1. The entire table is displayed in A.1 Appendix.

Table 5.1: An illustration of compressed video information (date, number of comments, and title deleted). The video data was collected on 2021 December 4.

Video ID	Upvotes	Downvotes	Views	Sentiment
xuCn8ux2gbs	4467927	71547	112060703	0.869
WeA7edXsU40	129840	2236	2826612	0.766
HVjlcUtuENM	25551	31465	1741401	0.123
u_J0Ng5cUGg	57341	642191	4679316	0.062
ffxKSjUwKdU	7036447	316928	977824338	0.953

4.3 Modelperformance

The findings of two assessments of Cross-validation on the training and validation subsets for the models and model assessment on an unknown subset testing —are presented in this section. performance of a model and without usage vectors words is included in the results for both evaluations. The suffix + WV is used in this section if the model was trained using word vectors. accuracy over all iterations and on average are included in cross-validation findings. Recall, accuracy, precision, and F1 score—which are given as a measure between 0 and 1—along with a ROC curve and a confusion matrix make up the evaluation of findings from unseen data.

Table 5.2: Scores for precision on an unknown test subset.

MODEL	PRECISION
LSTM	**0.9021**
LSTM+ WV	0.8931
BiLSTM	0.8753
BiLSTM+WV	0.8937
CNN-LSTM	0.8648
CNN-LSTM+WV	0.8675
CNN-BiLSTM	0.8539
CNN-BiLSTM+WV	0.883

score F1 may be observed in Table 5.2 for each of the four models, both word vectors and without.

Table 5.3 Confusion matrix for CNN-BiLSTM Model

True

Predicted	Negative	Positive
Negative	784	79
Positive	135	728

(a) Trained without word vectors

True

Predicted	Negative	Positive
Negative	773	90
Positive	111	752

(b)Trained with word vectors

Table 5.9: Confusion matrix for the LSTM Model

True

Predicted	Negative	Positive
Negative	755	108
Positive	105	758

(a) Trained without word vectors

True

Predicted	Negative	Positive
Negative	772	91
Positive	98	765

(b)Trained with word vectors

Table 5.10: Confusion matrix for the BiLSTM Model

True

Predicted	Negative	Positive
Negative	745	118
Positive	108	755

(a) Trained without word vectors

True

Predicted	Negative	Positive
Negative	744	119
Positive	84	779

(b)Trained with word vectors

Table 5.11: Confusion matrix for CNN-LSTM Model

True

Predicted	Negative	Positive
Negative	734	129
Positive	109	757

(a)Trained without word vectors

True

Predicted	Negative	Positive
Negative	762	101
Positive	101	762

(b)Trained with word vectors

Figures 5.2, 5.3, 5.4, and 5.5 show ROC curve graphs, which show the model's predictive capabilities. The diagonal line that is dashed represents predictions at random.

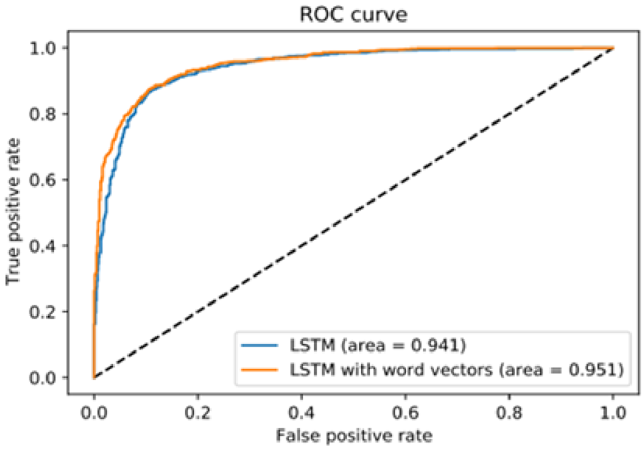

Figure 5.2 depicts the ROC curve of an LSTM model trained with and without word vectors.5.4 Model performance based on LSTM

The model's likelihood of success categorizes a unexpectedly positive comment higher than an arbitrary negative comment is shown by the worth of the area, or AUC is the area under the ROC curve.

As a result, goal O5 was met, and the following provides the response to the fourth and final research question: If a video is well received, viewers are more likely to vote it up and watch it more often. Videos that have been typically unpopular will receive fewer votes and views. Because there is no statistically significant difference between the sentiment of the video and the number of downvotes, users may express their displeasure with it in comments rather than by giving it a low rating.

V. CONCLUSION

this dissertation set out to create a dataset of YouTube comments and develop LSTM-based models capable of categorizing remarks as positive or negative. Through manual classification, nearly half of the 19,951 comments were identified as either positive or negative. Four LSTM-based models were then trained and evaluated using cross-validation and assessment on previously unknown data. The BiLSTM model utilizing pre-trained word vectors achieved the highest accuracy at 90%, along with superior performance in other metrics such as recall, precision, F1 score, and AUC.

Furthermore, the models were applied to the IMDB dataset to demonstrate their versatility in analyzing sentiment across different types of text, yielding promising results with an accuracy of 89.87% for the BiLSTM model. This suggests that the models are not only effective for analyzing YouTube comments but also applicable to other forms of sentiment data like movie reviews. Additionally, the dissertation explored the relationship between sentiment in YouTube comments and video metrics, revealing that videos with positive sentiment tend to receive more views and upvotes. However, no significant correlation was found between sentiment and the quantity of downvotes, indicating a tendency for viewers to express negative opinions through comments rather than downvoting.Looking ahead, there are several avenues for future research to enhance the performance of sentiment analysis in YouTube comments. This includes updating and expanding the comment dataset, incorporating longer comments, and exploring more sophisticated model architectures such as multi-kernel CNN-LSTM. Moreover, utilizing word vectors pre-trained specifically on YouTube comment data could potentially further improve model accuracy and effectiveness in capturing the nuances of sentiment expressed in these comments.

VI. REFERENCES

[1]. Akshi Kumar and Geetanjali Garg. 2019. Systematic literature review on context-based sentiment analysis in social multimedia. Multimedia Tools and Applications 79 (2019), 15349–15380.

[2]. Antonio Gulli and Sujit Pal. 2017. Deep Learning with Keras. Packt Publishing Ltd.

[3]. Ashraf Elnagar, Ridhwan Al-Debsi, and Omar Einea. 2020. Arabic text classification using deep learning models. Information Processing & Management 57, 1 (2020), 102121.

[4]. Beakcheol Jang, Inhwan Kim, and Jong Wook Kim. 2019. Word2Vec convolutional neural networks for classification of news articles and tweets. PLoS One 14, 8 (2019), e0220976.

[5]. Charles F. Hofacker, Edward Carl Malthouse, and Fareena Sultan. 2016. Big data and consumer behavior: Imminent opportunities. Journal of Consumer Marketing (2016).

[6]. Erik Cambria. 2016. Affective computing and sentiment analysis. IEEE Intelligent Systems 31, 2 (2016), 102–107.

[7]. Fatemeh Hemmatian and Mohammad Karim Sohrabi. 2019. A survey on classification techniques for opinion mining and sentiment analysis. Artificial Intelligence Review 52 (2019), 1495–1545.

[8]. FennaMiedema. 2018. Sentiment Analysis with Long Short-Term Memory Networks. Vrije Universiteit Amsterdam.

[9]. Ibrahim Said Ahmad, Azuraliza Abu Bakar, and MohdRidzwanYaakub. 2020. Movie revenue prediction based on purchase intention mining using YouTube trailer reviews. Information Processing & Management 57, 5 (2020), 102278.

[10]. ImeneGuellil and Kamel Boukhalfa. 2015. Social big data mining: A survey focused on opinion mining and sentiments analysis. In Proceedings of the 2015 12th International Symposium on Programming and Systems (ISPS'15). IEEE, Los Alamitos, CA, 1–10.

[11]. Iti Chaturvedi, Erik Cambria, Roy E. Welsch, and Francisco Herrera. 2018. Distinguishing between facts and opinions for sentiment analysis: Survey and challenges. Information Fusion 44 (2018), 65–77.

[12]. Jiachen Du, Lin Gui, Ruifeng Xu, and Yulan He. 2017. A convolutional attention model for text classification. In Proceedings of the National CCF Conference on Natural Language Processing and Chinese Computing. 183–195.

[13]. Jin-Ah Kwak and Sung Kyum Cho. 2018. Analyzing public opinion with social media data during election periods: A selective literature review. Asian Journal for Public Opinion Research 5, 4 (2018), 285–301.

[14]. KigonLyu and Hyeoncheol Kim. 2016. Sentiment analysis using word polarity of social media. Wireless Personal Communications 89, 3 (2016), 941–958.

[15]. Laith Abualigah, Hamza Essam Alfar, Mohammad Shehab, and Alhareth Mohammed Abu Hussein. 2020. Sentiment analysis in healthcare: A brief review. In Recent Advances in NLP: The Case of Arabic Language. Springer, 129–141.

[16]. Liang-Chu Chen, Chia-Meng Lee, and Mu-Yen Chen. 2019. Exploration of social media for sentiment analysis using deep learning. Soft Computing 24 (2019), 8187–8197.

[17]. Lopamudra Dey, Sanjay Chakraborty, Anuraag Biswas, Beepa Bose, and Sweta Tiwari. 2016. Sentiment analysis of review datasets using naive Bayes and K-NN classifier. arXiv:1610.09982.

[18]. Maryam Esmaeili and Alberto Vancheri. 2010. 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. IEEE/WIC/ACM.

[19]. Mika V. Mäntylä, Daniel Graziotin, and MiikkaKuutila. 2018. The evolution of

sentiment analysis—A review of research topics, venues, and top cited papers. Computer Science Review 27 (2018), 16–32.

[20]. Mohammad Al-Smadi, Bashar Talafha, Mahmoud Al-Ayyoub, and YaserJararweh. 2019. Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. International Journal of Machine Learning and Cybernetics 10, 8 (2019), 2163–2175.