Anbar Journal of Engineering Science (AJES)
P-ISSN: 1997-9428 ; E-ISSN: 2705-7440

Vol. 13 , No. 2 (2022) , pp. 14 ~ 21
https://doi.org/10.37649/aengs.2022.176353

Unviersty of Anbar

Anbar Journal of Engineering Science©

journal homepage: https://ajes.uoanbar.edu.iq/

Optimizing Sentiment Big Data Classification Using Multilayer Perceptron

Khalid Shaker [a]

[a] *Coleege of Computer Sceinces and IT, University of Anbar, Ramadi, Iraq raq*

PAPER INFO

Paper history:
Received 13/08/2022
Revised 14/10/2022
Accepted 25/10/2022

Keywords:
electrical power, optimum inclination angle, PV/T collector, thermal efficiency

ABSTRACT

Internet-based platforms such as social media have a great deal of big data that is available in the shape of text, audio, video, and image. Sentiment Analysis (SA) of this big data has become a field of computational studies. Therefore, SA is necessary in texts in the form of messages or posts to determine whether a sentiment is negative or positive. SA is also crucial for the development of opinion mining systems. SA combines techniques of Natural Language Processing (NLP) with data mining approaches for developing inelegant systems. Therefore, an approach that can classify sentiments into two classes, namely, positive sentiment and negative sentiment is proposed. A Multilayer Perceptron (MLP) classifier has been used in this document classification system. The present research aims to provide an effective approach to improving the accuracy of SA systems. The proposed approach is applied to and tested on two datasets, namely, a Twitter dataset and a movie review dataset; the accuracies achieved reach 85% and 99% respectively.

1. Introduction

Social media sites are popular destinations in the online world. Millions of users visit social networking sites, for example Twitter, YouTube, and Facebook [1]. A foundational shift is represented by online social med ia in how information is produced, transferred, and consumed. Due to the rapid expansion of social media, individuals and organizations use public opinions increasingly for their decision-making [2].

The public and private opinions are constantly expressed and diffused on different social media platforms about a wide range of topics. Social network sites are used by millions of people to share their opinion, express about their emotions and unfold details about what they do every day [3]. However, people write many things through online communities including social activities or commenting on products, and offer a place where people can talk to each other and share information. Moreover, businesses can directly communicate with their customers through social media in order to market their products and services [4,5].

SA (sentiment analysis) and opinion mining use information retrieval and computational linguistics to study people's perspectives, opinions, emotions and attitudes about entities, issues, individuals, topics, and events. The task is technically challenging and practical [6]. For example, businesses are constantly trying to gather information about the opinions of customers and the public. Furthermore, customers may want to find out what previous users have to say about a product or service before they purchase it [7].

In recent years, this field has grown quickly in parallel with the rapid growth of online social media. SA, which uses advanced techniques for web mining, machine learning, information retrieval and information extraction to derive information from natural language (NLP), has emerged as one of the most active research areas in the area of NLP since the early 2000s [8, 29, 30, 31].

Numerous techniques for supervised and unsupervised SA have been developed over the years. Throughout early research in supervised machine learning, all types of supervised machine learning methods were used, including neural networks, maximum entropy,

* Corresponding author: Khalid Shaker; khalidalhity@gmail.com ; +9647811061019

support vector machines (SVMs), and nave Bayes (NBs). In unsupervised learning algorithms, sentiment lexicons, grammar and syntactic patterns are exploited [9,10].

In addition, Studies have focused on various aspects of sentiment classification characterized by supervised and unsupervised methods as well as different goals. These models typically rely on machine and deep learning.

Duncan & Zhang in 2015, used a dataset gathered from Twitter API (Application Program Interface) for SA. The authors used a feedforward neural network to classify a Twitter dataset into positive or negative tweets. The accuracy obtained using the feedforward neural network was 74.15% [11].

Asghar et al. used supervised and unsupervised learning algorithms to develop a lexical resource for sentiment classification of health-related sentiments [13]. As well, Ramadhani et al. used the Twitter API to gather data. In addition, they compared the performance of NB smoothing methods at improving Tweet SA [14].

Mehla et al. [15] used NB, K-nearest neighbor, and random forest techniques to analyze movie reviews. Based on 1,000 positive and 1,000 negative movie reviews archived on IMDb, the data were collected. In order to perform sentiment classification, a new model based on a three-layer neural network was developed by Han et al. [16].

To train and test deep learning models, Karakuş et al. [17] used Turkish movie reviews from www.beyazperde.com to perform sentiment classification. A total of 44,617 samples were analyzed, including both positive and negative reviews. CNN, LSTM, CNN–LSTM and MLP were employed, and CNN–LSTM achieved the best results and performance.

Reddy et al. categorized the tweets as either positive or negative. Instead of using conventional methods or preprocessing text data, the authors classified tweets using LSTM networks, CNNs, and ANN, which represent the distributed representations of words and sentences [18].

Mohamed et al. used MLP, CNN, and LSTM deep learning networks, as well as a hybrid CNN LSTM network, to develop a SA classifier for 50K movie reviews dataset from IMDb [19].

The study by Li et al. developed sentiment dictionaries for news sentiment model and built a LSTM with two-layer network for stock prediction. The results show that there is a clear advantage to using both technical indicators and news sentiment in an LSTM, rather than using just single of these sources of information at a time [20].

Gandhi et al. implemented SA model based on Deep Learning DL with NLP techniques and introduced new features that have a significant effect on the calculation of movie reviews. DL algorithms are designed to rate movie review tweets and recognize movie reviews with 87.74 and 88.02 percent testing accuracy, respectively [21].

Recently, Mahyarani et.al uses the Nave Bayes classification method, TF-IDF (Term Frequency – Inverse Document Frequency) feature extraction, and Information Gain for feature selection. The results of the evaluation show that Information Gain, TF-IDF, and Naive Bayes are capable of producing adequate precision, recall, and f1-score. [22].

Pota et al. developed a two-step method for analyzing sentiment on Twitter. Twitter jargon, like emoticons and emojis, is first transformed into plain text using language-independent or easily applicable methods [23].

The Universal Language Model Fine-Tuning ULMFiT with SVM were combined by Albadani et al. to create a new effective model of SA based on deep learning architectures to improve detection efficiency and accuracy. When the model was tested on the Twitter US Airlines dataset, the accuracy rate was 99.78% [24].

Hawever, Swathi et al. developed a new model for predicting stock price movements to understand how tweets and stock market prices are linked, the LSTM model has been used to classify the tweets into two kinds, positive and negative sentiments. Improvement the LSTM model's recall and accuracy are 90% and 95.33 percent, respectively [25].

This paper is structured as follows. Section 1 presents the introduction of sentiment Analysis and the related work done in this field. Section 2 discusses the proposed system for sentiment analysis and the different steps of this system. Section 3 discusses the results of the work. Finally, Section 4 conclusion and provide recommendations for future work.

2. The Improvement appraoch

In our proposed approach, as shown in Figure 1, there are two phases: training and testing. The training phase is designed to develop a classification model that can distinguish positive from negative documents based on input labeled documents. To test the training model, unlabeled documents are assigned positive or negative labels. Five steps are involved in the system: pre-processing, feature extraction, feature selection, classification model for SA, and system evaluation.

In the first step, the main goal is to use NLP techniques to process the document text so that it can be used in the next step for accurate feature extraction. In the second step, several features from the text of the input document need to be pulled out. In this step, we use TF-IDF to show the different kinds of features that are used in the proposed system. As a third step, we use the chi-square test to choose which features to use.

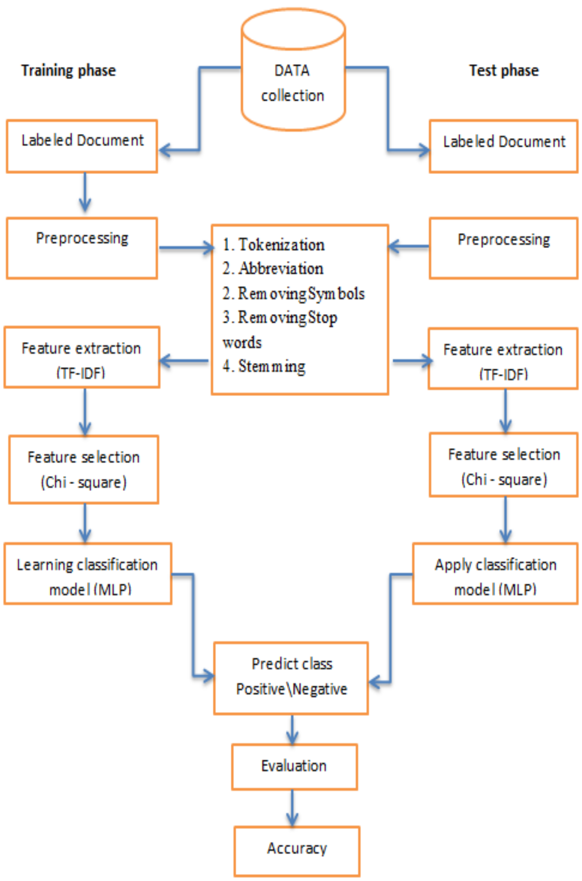

Figure 1. The Proposed Sentiment Analysis approach

The goal is to choose a subset of the original set of features that keeps the same meaning and doesn't lose any information. In the step, we build a classification model that can tell the difference between documents that have been labeled positively and documents that have been labeled negatively. Creating such a model can be done with a number of neural network algorithms.

In the proposed method, an MLP classifier is used. Text classification problems are often solved with these algorithms, which work very well. In the fifth and final step, the accuracy, recall, precision, and F-score have been selected to measure how well the system works.

2.1. Dataset Description

For training and testing of the system, two different dataset corpora have been used. The main objective for using two data sets in this paper was to show the difficulties of guessing the sentiment in short and often ungrammatical English texts

In the Twitter dataset, as opposed to relatively long and well-established English texts as in the movie reviews dataset. These two dataset corpora are described below:

A. Twitter Dataset

In paper, ML is used to classify sentiments and Twitter datasets from Kaggle [26] which were crawled and labeled as positive/negative. The data are in the form of CSV files containing 'tweets' and their corresponding sentiments. The training set that contain item_id, sentiment, and sentiment text CSV file has been used to train the model. An item id identifies a tweet, a sentiment is either 1 that represent (positive) or 0 that represent (negative), and a sentiment text is the actual tweet. The same thing for the test dataset that is consist of a CSV file of type item_id and sentiment text.

As seen on Twitter, the dataset is made up of text, symbols, emoticons, URLs, and references to specific individuals or groups. Sentiments can be predicted using words and emoticons, but URLs and references to other do not do this. As a result, URLs and references are ignored. Mistaken or incorrect words, extra punctuation, and repeated letters are also present in the text. It is therefore necessary to preprocess tweets in order to ensure a consistent dataset is generated. This dataset contains 1,494 tweets split into 749 negative tweets and 745 positive tweets, and the data are split into 70% training dataset and 30% test dataset.

B. Movie Reviews Dataset

The dataset of movie reviews was collected from the archives of the IMDb web portal [27]. In this dataset, we also apply all the pre-processing steps applied to the Twitter dataset. This dataset contains 400 tweets split into 200 negative tweets and 200 positive tweets, and the data are split into 70% training dataset and 30%

test dataset. The length factor of feature selection is 7,511 for the dataset of movie reviews.

2.2. Pre-processing

Pre-processing is the first step of the proposed system. Text pre-processing is an important stage in the SA of the data on Twitter and movie reviews. When it comes to movie review data, there is a lot of noise in a tweet. This includes unnecessary symbols like emoticons, misspellings, stop words and slang words.

The performance of SA approaches is frequently hampered by such noisy characteristics. Thus, in this work, pre-processing methods are applied prior to feature extraction and selection using specific machine learning algorithms.

After collecting many raw datasets, we process the document text. Preprocessing is implemented using Python, and a number of steps involved in the pre-processing is applied on the dataset for standardize it and filter the noise from the data to reduce its size. We initially perform pre-processing on a number of documents which is as follows:

A. Tokenization

This step is important in the pre-processing of data. Each word in the document is separated from the next by a space, a process known as tokenization. Symbols, numbers, and words can all be used as tokens. For the tokenization of words, the the natural language toolkit (NLTK) package is used [28]. Tokenizer selection depends on the characteristics of the data and the language being used to analyze it. In this work, a tokenizing method is created to tokenize the words using the document Tokenizer module for processing English language terms.

B. Abbreviation

These platforms are based on short and snappy communication. The number of abbreviations that people use on popular social media channels especially for common phrases is astonishing. People tend to use and create new abbreviations on Twitter since it has a 140-character limit. These platforms are based on short and snappy communication. The abbreviations dictionary has been developed manually and contains a large number of abbreviations. It helps expand the tweets and improve the overall sentiment score.

C. Removing Stop Words

Words like "on," "am," "the," and "are," which appear frequently in documents, are considered stop words in Natural Language Processing (NLP). They do not affect opinion or sentiment scores when applied to lexical resources because of their low emotional significance.

To identify and remove these words from any document a predetermined stop words are employed. The list contains all possible stop words in the English language. Therefore, the usage of stop word removal in the SA process is necessary in any system. Removing stop words from a document leads to improved accuracy.

D. Stemming

The stemming and lemmatizing of words produce a normalized form in a word of the text. To find the root (basis) of a word in a text, the stemming method is used. By removing the suffix using normalizes, the root meaning of the word is revealed. It is possible to use a variety of stemming algorithms to perform word stemming. The Porter Stemmer algorithm has been used to remove suffixes from words in this data pre-processing approach in order to retrieve the appropriate meaning from the text.

As mention above, the Porter Stemmer algorithm from the Python NLTK stem package was used to implement this stemming functionality. Stemming the word character-by-character, Porter Stemmer removes the suffix and provides the root meaning. As a result of this work's stemming, the word is returned to its root meaning.

E. Removing Symbols and Repeated Characters

We discovered some features that could affect the experimental results, such as "http://t.co/NxSE1HYTHj," or (: | [] ;: - + () >?! @ # percent *,). Links to other websites, such as "@" or "#," can be found throughout a lot of documents. An @ symbol immediately following a word denotes a username, which is omitted entirely from our search results. Usernames are deemed to have no bearing on the content. People frequently use multi-character words to convey their deepest emotions. For instance, "LOVEEEEEEEEE." Due to the excessive number of 'Es' in this word, it is imperative that they be removed.

2.3 Feature Extraction Using TF-IDF

A feature vector is generated for each pre-processed document. Full-text documents should be converted to document vectors in order to represent a document's features. It is possible to determine the frequency of individual words or features in documents by utilizing the TF-IDF model. In this way, a set of documents can be displayed in a matrix that shows the weighted relationship between the words in each document. It is the weight of a feature in a document that determines how important it is. This is the case with documents like D, where each feature's importance is determined by the weight it has been assigned.

2.4 Feature Selection Using Chi Square

When the pre-processing and feature extraction stages have been done, the next critical step in text classification is feature selection. Attribute selection is another name for it. The main aim of feature selection is to choice a subset of features from the dataset without affecting the classifier performance by removing redundant and irrelevant features. Both the classification algorithm and the feature selection method play a role in how accurate the sentiment analysis is.

Selecting features that are both irrelevant and inappropriate can lead to unexpected outcomes. Sentiment analysis can be made more accurate and efficient with the help of Feature Selection. The dataset's dimensionality is reduced, learning accuracy is improved, and the results are easier to understand.

Preprocessing and feature extraction still produce a large amount of data, even after reducing the number of features. Several methods have been devised to improve classification accuracy by limiting the number of features used as input.

Chi-square statistics (X2), one of the most commonly used metrics for feature selection, are used in our approach. Since its introduction, the Chi-square (x2) statistical test has been widely accepted as a hypothesis test for assessing the interdependence of two variables. The chi-square test is frequently used in natural language processing to test the independence between the term and the class. In NLP, it is a common method for identifying the most relevant features. The CHI2 test is a statistical method for determining the independence of two independent events.

2.5 Classification Model for Sentiment Analysis

Preparing a classification model that can effectively distinguish between tweets with positive and negative labels during training is the most important step in the proposed system. The MLP network is a popular neural network for use in SA. A set of data samples is used to build the network's predictive model.

Three layers make up the MLP classifier network. The weights we used to connect the various layers are typically set to a value ranging from 0 to 1. When MLP is used, the error function is minimized to the greatest extent possible. The best weights are determined by minimizing this error. Table 1 represents the number of fully connected layers in the model, which comprises three layers.

Table 1 Three-Layer Model Using All Features

No.	**No. of Layers**	**No. of Nodes**	**Accuracy**	**Precision**	**Recall**	**F-score**
1	3	2048-2048-2048	0.6643	0.67	0.66	0.66
2	3	1024-1024-1024	0.6643	0.67	066	0.66
3	3	512-512-512	0.6620	0.67	0.66	0.66
4	**3**	**256-256-256**	**0.6875**	**0.69**	**0.69**	**0.69**
5	3	128-128-128	0.6736	0.68	0.67	0.67
6	3	64-64-64	0.6712	0.68	0.67	0.67
7	3	32-32-32	0.6643	0.70	0.66	0.65
8	**3**	**16-16-16**	**0.6481**	**0.65**	**0.65**	**0.64**
9	3	2048-1024-512	0.6666	0.67	0.67	0.67
10	3	1024-512-256	0.6759	0.69	0.68	0.67
11	3	512-256-128	0.6851	0.69	0.69	0.68
12	3	256-128-64	0.6851	0.69	0.69	0.68
13	3	128-64-32	0.6620	0.67	0.66	0.66
14	3	64-32-16	0.6736	0.68	0.67	0.67

Classification is the process of determining to which group an item belongs. We're going to use a set of training data and labels to create a classification algorithm. A classifier that is trained on a set of documents labeled "positive" or "negative" is what we need in this case. Based on a document's attributes, the classifier can label future documents as either positive or negative. Models for both training and testing datasets are constructed using an MLP classifier.

We divided the data into two sets: one set for training and the other set for testing purposes. As part of the testing process, we also keep separate positive and negative relabeled datasets. These data are then fed into the MLP classifier, which uses back propagation to refine its weights as it gains classification experience.

In this work, the MLP classifier based on a feedforward ANN in the current implementation uses backpropagation to learn the model and modify the weights. Each layer consists of weight matrix W, a bias vector b, and an output vector. Inputs (x1, x2, x3..xm) and connection weights (w1, w2, w3..wm) are commonly real numbers. Single value has been passed through the nonlinear ReLU activation function to squish it into the range [0, 1]. Based on the neural network's output, the ReLU function returns a probability of 0 (negative) and 1 (positive).

3. Experimental Results

The results of a Twitter and movie reviews SA using ML are demonstrated that the proposed method achieved the best results and highest scores on the twitter dataset when use three layers as shown in Figure 2. The best result of accuracy is 84.95%, the best result of Precision is 86%, the recall 85% and F1-score was 84%.

Figure 2 Evaluation Metrics for Three-Layers

From results in Table 2, it can be seen that the number of nodes in each layer has a major effect on the performance of the model. In our experiment, different number of nodes have been selected for detect the best number that provide the better results, and found that 16 nodes in each layer gave the best performance in Accuracy, Precision, Recall, and F1-score.

Table 2: Three-Layers Model

No	No. of Layers	No. of Nodes	Accuracy	Precision	recall	F-score
1	3	2048-2048-2048	0.8310	0.83	0.83	0.83
2	3	1024-1024-1024	0.8310	0.83	0.83	0.83
3	3	512-512-512	0.8356	0.84	0.84	0.83
4	3	256-256-256	0.8333	0.84	0.83	0.83
5	3	128-128-128	0.8379	0.84	0.84	0.84
6	3	64-64-64	0.8379	0.84	0.84	0.84
7	3	32-32-32	0.8402	0.85	0.84	0.84
8	**3**	**16-16-16**	**0.8495**	**0.86**	**0.85**	**0.84**
9	3	2048-1024-512	0.8310	0.84	0.83	0.83
10	3	1024-512-256	0.8333	0.84	0.83	0.83
11	3	512-256-128	0.8356	0.84	0.84	0.83
12	3	256-128-64	0.8379	0.84	0.84	0.84
13	3	128-64-32	0.8356	0.84	0.84	0.83
14	3	64-32-16	0.8425	0.85	0.84	0.84

The results of appling our proposed Multilayer Perceptron (MLP) model for SA on movie reviews dataset have been shown in Table 2. Two layers have been selected to build our model, and different number of nodes have been tested for detect the best results. It can be seen in Table 2 the best result gained when 16 nodes have been selected for each layer. The accuracy was 99.16%, the best result of Precision was 99%, the recall 99% and F1-score was 99%. We believe that by employing the MLP helps significantly the algorithm to explore more regions in the search space and obtain better results,

Table 3: Two-Layers Model

No	No. of Layers	No. of Nodes	Accuracy	Precision	Recall	F-score
1	2	2048-2048	0.989	0.99	0.99	0.99
2	2	1024-1024	0.9833	0.98	0.98	0.98
3	2	512-512	0.9916	0.99	0.99	0.99
4	2	256-256	0.9833	0.98	0.98	0.98
5	2	128-128	0.9916	0.99	0.99	0.99
6	2	64-64	0. 9916	0.99	0.99	0.99
7	2	32-32	0. 0.9916	0.99	0.99	0.99
8	**2**	**16-16**	**0. 0.9916**	**0.99**	**0.99**	**0.99**

Note that the best results are presented in bold. The results in Table 3, indicate that the accuracy and other performance parameters obtained when using the movie review dataset are higher than those obtained when using the Twitter dataset because of the following reasons:

1. Limited tweet size: With only 140 characters on hand, compact and abbreviated statements are generated, resulting in a sparse set of features.
2. Use of slang in Twitter: These words are different from English words, thereby resulting in an outdated approach because of the casual nature of the usage of social media by people. Thus, a noisy and obscure dataset is obtained.
3. Twitter features: Tweets are short messages that are riddled with irrelevant symbols and other noise, hashtags, user references, URLs, misspellings, emoticons and stop words. It is common for SA approaches to suffer from the effects of such noise.
4. User variety: Users express their opinions in various ways because of the wide popularity of Twitter; some use different languages in between,

and others use repeated words or symbols to convey an emotion.

4. Conclusions

In this study, it's been noticed different steps and different available tools used to perform sentiment analysis. A novel model for SA was designed. The model captures the polarity of positive or negative sentiments from Twitter and movie review datasets. Multilayer Perceptron Neural Networks (MLP) for the purpose of classification the dataset to positive or negative was used. On both twitter dataset and Movie reviews dataset the proposed approch obtained 85% sentiment analysis accuracy. We believe that the pre-processing stage plays an important role in data filtering or noise removal, and demonstrate the pre-processing linguistic data using NLP techniques. The TF-IDF technique is used in the feature extraction stage to dimensionally reduce a primary set of raw data to manageable groups for processing. Also, the results show that using chi-square feature selection will meaningfully improve the accuracy of classification for both datasets. for future work, the model can be trained and improved to handle a range of sentiments. Tweets do not consistently have positive or negative sentiments. At times, they may have no sentiment, i.e. they are neutral. In addition, another social media dataset may be used, and the data size may be increased.

References

[1] K. Francisco and M. Panguila, “Sentiment Analysis on Social Media Data Using Intelligent Techniques,” International Journal of Engineering Research and Technology, 2019.

[2] S. Joshi and D. Deshpande, “Twitter Sentiment Analysis System,” Int. J. Comput. Appl., vol. 180, no. 47, pp. 35–39, 2018.

[3] A. Kumar and G. Garg, “Systematic literature review on context-based sentiment analysis in social multimedia,” Multimed. Tools App., 2019.

[4] M. Amajd, Z. Kaimuldenov, and I. Voronkov, “Text Classification with Deep Neural Networks,” International Conference on Actual Problems of System and Software Engineering, pp. 364-370,2017.

[5] V. VardanReddy, M. Maila, S. Sri Raghava, Y. Avvaru, and S. V. Koteswarao, “A Survey on Analysis of Twitter Opinion Mining Using Sentiment Analysis,” Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., pp. 537–542, 2019.

[6] L. Zhang, S. Wang, and B. Liu, “Deep learning for sentiment analysis: A survey,” Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 8, no. 4, Jul. 2018.

[7] B. Lopez and X. Sumba, “IMDb Sentiment Analysis,” pp. 2–6, 2019.

[8] R. Adyan Marendra, and H. S. Goo, "Twitter sentiment analysis using deep learning methods," 2017 IEEE 7th International Annual Engineering Seminar (InAES), pp. 1-4, 2017.

[9] V. N. Patodkar and I. R. Shaikh, “Twitter as a Corpus for Sentimental Analysis using Naïve Bayes,” vol. 6, no. 7, pp. 375–379, 2017.

[10] K. K. Uma and D. K. Meenakshisundaram, “A Novel Optimizer Technique for Sentiment Polarity in Social Web Environment,” International Journal of Applied Engineering Research, pp. 118–125, 2019.

[11] B. Duncan and Y. Zhang, “Neural networks for sentiment analysis on Twitter,” Proc. 2015 IEEE 14th Int. Conf. Cogn. Informatics Cogn. Comput. ICCI*CC 2015, pp. 275–278, 2015.

[12] A. Tripathy, A. Agrawal, and S. K. Rath, “Classification of Sentimental Reviews Using Machine Learning Techniques,” Procedia Comput. Sci., vol. 57, pp. 821–829, 2015.

[13] M. Z. Asghar, S. Ahmad, M. Qasim, S. R. Zahra, and F. M. Kundi, “SentiHealth: creating health-related sentiment lexicon using hybrid approach,” Springerplus, vol. 5, no. 1, 2016.

[14] R. A. Ramadhani, F. Indriani, and D. T. Nugrahadi, “Comparison of Naive Bayes smoothing methods for Twitter sentiment analysis,” 2016 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2016, pp. 287–292, 2017.

[15] S. Mehla, and M. M “Sentiment Analysis of Movie Reviews using Machine Learning Classifiers,” Int. J. Comput. Appl., vol. 182, no. 50, pp. 25–28, 2019.

[16] H. Han, X. Bai, and P. Li, “Augmented sentiment representation by learning context information,” Neural Comput. Appl., vol. 4, 2018.

[17] B. Ay Karakuş, M. Talo, İ. R. Hallaç, and G. Aydin, “Evaluating deep learning models for sentiment classification,” Concurr. Comput, vol. 30, no. 21, pp. 1–14, 2018.

[18] D. M. Reddy, D. N. V. S. Reddy, and D. N. V. S. Reddy, “Twitter Sentiment Analysis using Distributed Word and Sentence Representation,” arXiv preprint arXiv:1904.12580, 2019.

[19] N. Mohamed Ali, M. M. A. El Hamid, and A. Youssif, “Sentiment Analysis for Movies Reviews Dataset Using Deep Learning Models,” Int. J. Data Min. Knowl. Manag. Process, vol. 09, no. 03, pp. 19–27, 2019.

[20] X. Li, P. Wu, W Wang “Incorporating stock prices and news sentiments for stock market

[21] prediction,"Inf Process Manag 57(5):102212. https ://doi.org/10.1016/j.ipm.2020.10221, 2020.
[22] U. D. Gandhi, P. Malarvizhi Kumar, G. Chandra Babu, and G. Karthick, "Sentiment Analysis on Twitter Data by Using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM)," Wirel. Pers. Commun., no. 0123456789, 2021.
[23] M. Mahyarani, A. Adiwijaya, S. Al Faraby, & M. Dwifebri, "Implementation of Sentiment Analysis Movie Review based on IMDB with Naive Bayes Using Information Gain on Feature Selection," In 2021 3rd International Conference on Electronics Representation and Algorithm (ICERA) (pp. 99-103). IEEE, 2021
[24] M. Pota, M. Ventura, R. Catelli, M. Esposito, "An Effective BERT-Based Pipeline for Twitter Sentiment Analysis: A Case Study in Italian," Sensors, 21, 133, 2021.
[25] B. AlBadani, R. Shi, J. Dong, "A Novel Machine Learning Approach for Sentiment Analysis on Twitter Incorporating the Universal Language Model Fine-Tuning and SVM," Appl. Syst. Innov. , 5, 13, 2022.
[26] T. Swathi, N. Kasiviswanath, & A. Rao, "An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis," Appl Intell, 2022.
[27] Kaggle website, https://www.kaggle.com/ [Online], accessed on March, 2022.
[28] IMDB review, http://reviews.imdb.com/Review/ [Online], accessed on March, 2022.
[29] Natural Language Toolkit, https://www.nltk.org/ [Online], accessed on March, 2022
[30] A., Alqudsi, N. Omar, & Ibrahim, R. W. (2016). Rule Based and Expectation Maximization algorithm for Arabic-English Hybrid Machine Translation. International Journal of Artificial Intelligence (IJ-AI), 5.
[31] A., Alqudsi, N. Omar, & Shaker, K. (2014). Arabic machine translation: a survey. Artificial Intelligence Review, 42(4), 549-572.
[32] Alqudsi, A., Omar, N., & Shaker, K. (2019, May). A hybrid rules and statistical method for Arabic to English machine translation. In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-7). IEEE.