Global Journal of Engineering and Technology Advances

eISSN: 2582-5003

Cross Ref DOI: 10.30574/gjeta

Journal homepage: https://gjeta.com/

(Research Article)

Check for updates

Sentiment analysis for airline twitter based of machine learning and deep learning

Shajan Mohammed Alsowaidi *

Department of Computer Science, College of Education, Mustansiriyah University, Baghdad, Iraq.

Global Journal of Engineering and Technology Advances, 2024, 20(02), 125–134

Publication history: Received on 16 July 2024; revised on 17 August 2024; accepted on 20 August 2024

Article DOI: https://doi.org/10.30574/gjeta.2024.20.2.0133

Abstract

The study of public opinion can be beneficial for obtaining certain knowledge. Thus, the sentiment analysis of the social networks, for example, the Twitter or Facebook has grown into an effective way of understanding users' opinion and has many uses. Nevertheless, the efficiency and accuracy of sentiment analysis seem to be hampered by the problems rising from natural language processing processes that are inherent with the text. Currently the airline sector is considered the significant field of the market. To sustain that sector and constantly update it, mind mining becomes inevitable. This paper, proposed a model for sentiment analysis based on extracting two different features. Term frequency-inverse document frequency and Word2vec. These feature introduced separately to different classifiers to classify the sentences as positive, negative, or neutral. Twitter US Airline dataset used to evaluate the performance of the proposed model. Bi-directional Long short Term memory outperformed others methods with recall, Precision, and F-Score reached to 0.97, 0.98,and 0.97 respectively when using Word2vec feature.

Keywords: Sentiment analysis; Word2Vec; BI-LSTM; CNN; LSTM

1. Introduction

Sentiment analysis (SA) can be defined as methods, techniques, and tools about detecting and extracting subjective. that is, opinion and attitudes from information which is derived from language [1]. Conventionally sentiment analysis is all about opinion decision maker's attitude, or as known, polarity, whether a decision maker has positive, neutral or negative attitude towards an object or not [2]. The object of whereas, a sentiment analysis has more often been a product or a service whose performance analysis has been displayed on the Internet. This may explain why what we describe as sentiment analysis and opinion mining are often considered equivalent although, we believe that the latter is more. compelling, to regard sentiments as opinion-like affect-laden states [3].

The issue identified as automatic sentiment analysis (SA) has become rather popular. Even though SA is a significant area and is already being actively applied in various fields and industries, some factors undoubtedly remain from this statement and it can be noted that SA has many NLP issues, and it is not an easy task. roughly, it is found that many developed analysis types still suffer from theoretical and technical problems regarding their polarity sensitiveness [4]. The investigation of the correlation between those problems and the sentiment structure also the effect on the result correctness. makes it possible to confirm that accuracy is a significant issue in the most recent documents dealing with the sentiment analysis problem and demonstrates that some difficulties, like how to deal with negation or achieving domain independence, do influence accuracy [5].

It can be noted that social media are associated with SA and are valuable sources of information. Currently social networks are being created in large quantities. They allow creating much more compound and mutually connected information [6]. Thus, Thai et al argued while deciding on what to do with formal data it is unhelpful to concentrate only on their structure and correlation but When handling presentation of data, inference, analysis, visualization, search and

* Corresponding author: Shajan M alsowaidi

navigation, and decision-making in complex networks, attempt to embrace a lifelong learning strategy[7]. Many works concern the creation of robust models to address strong growth in big data and the extension of sentiment analysis to various fields, including financial and marketing prediction, medical applications , and others . However, they scarcely consider the comparative assessment of various deep-learning approaches to exhibit empirical outcomes of their efficacies [8].

2. Related Works

Another categorical question with respect to the conduct of sentiment analysis is the categorisation of sentiment polarity. The challenge, therefore, is to classify the given piece of the written text into one of the sentiments' polarity type or non-sentiment: positive or negative, or neutral. According to the division of the lengths of the texts, there are three degrees of sentiment polarity categorization which include the document level, sentence level and the entity, and aspect level. The document level also relates to whether a document in general is positive or negative while the sentence level focuses to each individual sentence if it is positive or negative; The entity and aspect level then zeroes down to the details of what people actually like or dislike from their opinions [9].

Mulki et al. 2018 introduced a multiclassification system to identify emotions in tweets in three languages: Arabic, English, and Spanish. The system was used for subtask E-c, one of the subtasks of SemEval-2018 Task1, and it could identify one or more of the following 11 emotion labels: love, trust, optimism, joy, pessimism, sadness, fear, surprise, anger, anticipation, and disgust. They broke down the Multilabel Classification (MLC) problem into smaller issues that could be solved by conventional single-label classifiers using the Binary Relevance (PR) technique. In order to prepare all datasets, all URLs, usernames, dates, numbers, hashtag symbols, and punctuation marks were first deleted. Next, stop words were eliminated, emojis were replaced with tags to indicate their moods, and the words were stemmed and lemmatized[10]

In 2019, A. Farisi and colleagues offer a solution to help travelers comprehend hotel evaluations. They do this by employing the Multinomial Naïve Bayes Classifier technique to categorize positive and negative opinion reviews, and then preprocessing, feature extraction, and feature selection to achieve the best possible outcome. The dataset utilized comes from Data Finitis's Business database, which has 5,000 English hotel reviews in CSV format. Following preprocessing steps to remove desired characters, a bag of words is used to extract text features. Features are selected based on frequency, and those with the lowest frequency word are deleted. Another feature involves removing features with the smallest difference between positive and negative probability values. The experimental findings with 10-fold feature selection and preprocessing[11].

In 2020 Alsalman was presented a Arabic Tweet SA. The model consists of two main steps: preprocessing and classification. The preprocessing step include normalization, tokenization, and stemming. The normalization aims to standardized the word and all the letter be in the same representation. The tokenization step include separating the sentences into tokens based on the space between them. Stemmer based on remove all the prefix and suffix from the token the proposed method used Khoja stemmer. The second step based on using the discriminative multinomial naïve Bayes (DMNB) to classify the Tweet as positive, negative, or neutral. The classifier accuracy reached to 87.5% [12].

In 2020 Abdelrahman was presented a model for twitter airline sentiment analysis. The proposed model consists of three main steps: preprocessing, feature extraction, and classification. The preprocessing step include four cleaning stage. In first stage all stop words was removed from the text. This words may cause redundancy and make analysis more complex. The second stage include removing punctuation from the text such as "@" "*" and so on. The third stage include converting all letter to lowercase that ensure the stand razing of all words. The last step involve using the stemming to return all words to it's original form. The second step consists of using Bag of word to represent each term in the dataset. The last step include using different machine learning approach to classify the Tweet. The classifiers include SVM,RF, LR,NB, DT, and XGB. The SVM outperformed other methods with accuracy reached to 83.31% [13].

In 2021 Fatma Jemai et al proposed a method for SA based on using machine learning. The model consists of two main steps: preprocessing and classification. The processing step include five main stages: tokenization, stop word removing, deleting the URL, converting to lowercase, and Lemmatization. While the classification step include using of Naive Bayes, Multinomial NB , Bernoulli NB , Logistic Regression, and LinearSVC classifier. NB outperformed other classifiers with accuracy reached to 99.73%..Figure -2 shows the approaches of SA [14].

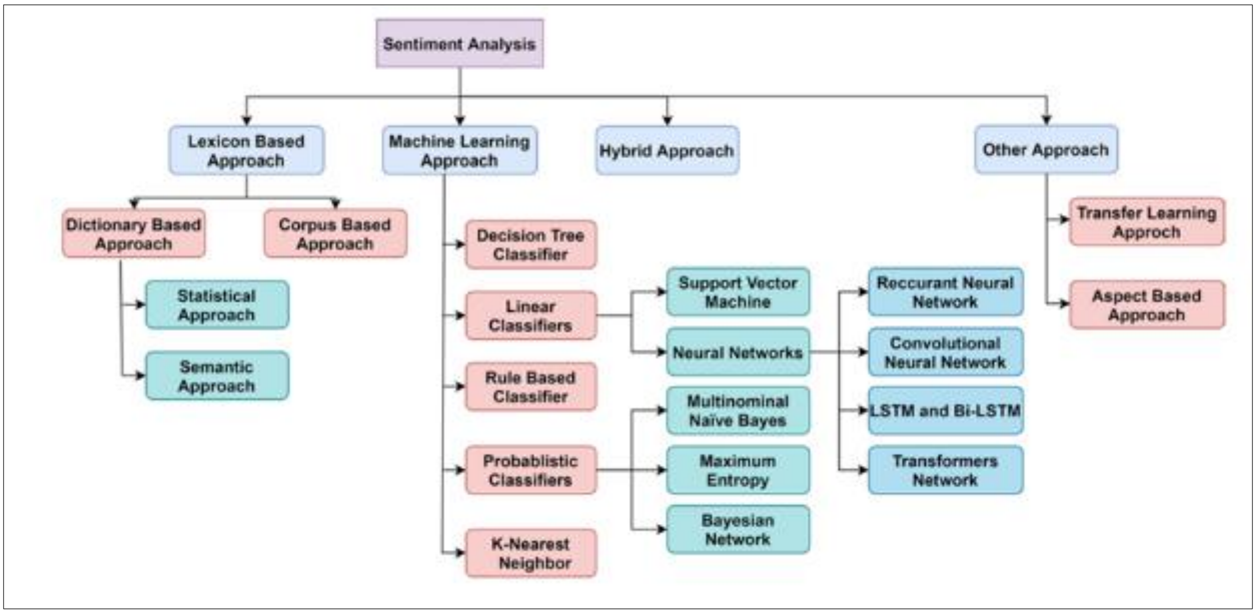

Figure 1 Approaches of Sentiment Analysis(15)

3. Sentiment Analysis Levels

There are four levels for SA. Document Level, Sentence Level, Phrase Level, and Aspect Level As shown in Figure(2).

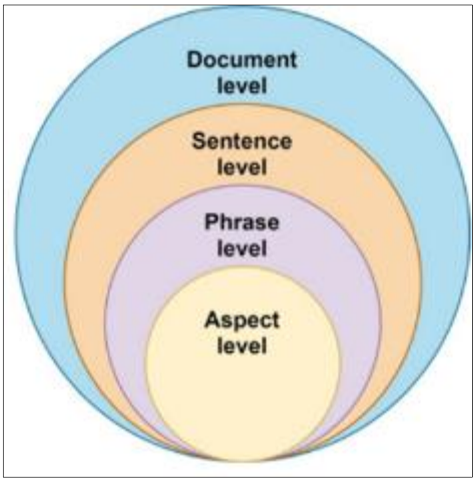

Figure 2 Sentiment Analysis levels (15)

3.1. Document Level

This level concern with analysis of all the document as one polarity. This approach usually used when analysis of pages or chapters as negative, positive or neutral. This type of analysis is least used than other types of analysis. In this approach both supervised and unsupervised method can be used for the classification purposes. Cross-domain and cross-language SA are the two utmost important concerns in document-level SA [16].

3.2. Sentence Level

At this level of analysis, each sentence is assigned a score and given its own polarity index. This is particularly useful w hen documents are associated with multiple combinations of sentiment. There are also two levels: document and sent ence, so the polarity of each sentence is calculated separately. With additional training data and processing capability, this employs the same method as the document level. To assess the overall sentiment of the paper, the polarity of each sentence can also be utilized alone or added together. Sentiment analysis at the document level isn't always enough for specific uses. [17].

3.3. Phrase level

The sentiment analysis can be done at phrase level when the opinion words are mined. This level may contain single or many aspects, which help in reviewing multiple lines at the same time. Word is the smallest building brick in language or the fundamental entity of language; word polarity is directly connected with subjectivity of a particular sentence or document. If a sentence includes an adjective, there is likely to be a greater than random probability that the sentence is a subjective one. Also, The terms chosen for this expression reveal the demographic indicators of individuals, including gender and age, as well as their aspirations, social status, and personality, as well as other psychological and social factors. Therefore, this terminology is considered to be the starting point for text sentiment analysis [18].

3.4. Aspect level

Using pre-training knowledge, we combine syntactic and semantic features for aspect-level sentiment polarity, and we present a syntactic and semantic improved multi-layer graph attention network for feature extraction in syntactic and semantic perspective. aspect-level sentiment analysis is carried out for the text. It would be useful to point out that each sentence can have several aspects; thus, the Aspect level sentiment analysis. Focusing on every aspect applied in the construction of the sentence and endows polarity to every aspect before arriving at the overall score. Again, sentiment has been calculated for the whole sentence [19].

4. Application of sentiment Analysis

It is commonly acknowledged that analysis of sentiment has many applications in the commercial, government, and biomedical fields, among other areas. Businesses in the domains of e-commerce and business intelligence can leverage consumer feedback analysis to enhance customer service, develop superior products, or refine marketing tactics to draw in new clients. Sentiment analysis is a useful tool for determining what people think about events or goods. The outcomes of SA contribute to a deeper understanding of the preferences or viewpoints of consumers on market developments. In this regard, Jain and Dandannavar presented a quick, adaptable, and scalable sentiment analysis framework for Twitter data using Apache Spark and a few machine learning techniques [20].

The sheer amount of requests, the range of topics covered, the variety of departments within a business, and the immediacy of each request make customer service management extremely challenging. Natural language understanding (NLU) is a sentiment analysis technique that looks for meaning, sentiment, tone, and other characteristics in everyday human language to interpret consumer inquiries as a human would. You may automatically handle emails, phone conversations, online chats, and customer care requests based on sentiment to prioritize essential issues [21].

SA can be used also for enhance customer services. A positive customer experience increases the likelihood that the client will return. A prosperous company understands that quality control over delivery methods is just as vital as quality control over product quality. We can obtain priceless and unbiased information on consumer attitude through brand monitoring. Nevertheless, this approach can also be used to customer service correspondence and surveys. By asking a straightforward question, such as "Will you recommend this brand, product, or service to your friend or family?" NPS (Net Promoter Score) surveys enable you to get feedback for your company. A single score on the number scale is the result. Businesses evaluate customers as promoters, critics, and passives using these sentiment scores[22].

5. Methodology of the proposed model

The proposed system for SA consists of three main steps: preprocessing, feature extraction, and classification s each of them include many stages. Figure-3 shows the main steps of the proposed model.

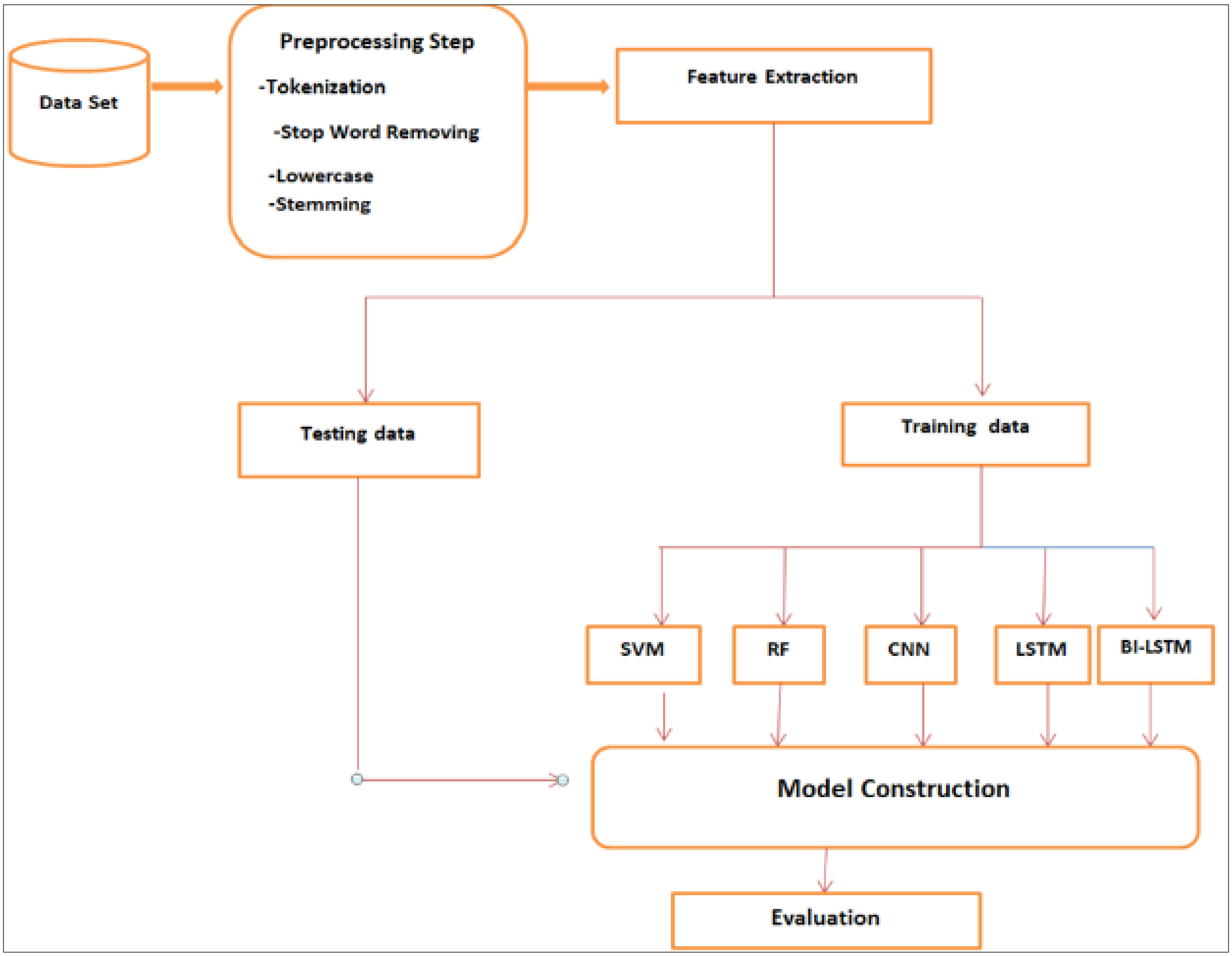

Figure 3 The structure of The proposed Model

5.1. Preprocessing

Preprocessing is an essential step in SA that helps to analysis data in the next steps. Preprocessing involves four main stages as follows [23].

5.1.1. Tokenization

Tokenization, in the field of Natural Language Processing (NLP) as well as machine learning is the action of splitting a stream of text into individual units, called tokens. These tokens can certainly be as small as characters if needed, or as large as words. The first and obvious reason is that such a process is indispensable due to the fact that it aids the machine in understanding the human language by slicing the language into possibly manageable chunks that may be easier for the machine to handle.

5.1.2. Stop word Removing

The words that are removed prior to actually processing natural language are termed as stop words. These are the basic of English and any other language such as article, prepositions, pronouns, conjunctions among others and does not really assist in conveying a lot of information in the text.

5.1.3. Lowercase

Is the process of converting all words in the dataset to a lowercase this process helps to standardized words and to help easy manipulation of in the next step of SA model.

5.1.4. Stemming

Stemming is the removal of affixes from a word either derived or inflected in a word and particularly in the written form. Stem does not necessarily be the morphological root of the word; It is adequate that related words correspond to the same stem even if this stem itself is not properly a root.

5.2. Feature Extraction

Feature extraction is defined as the conversion of raw data into numerical features that can easily be handled while the info from the data set is retained. This is considered to be producing a better result compared to if the machine learning were to be applied directly on the data. Two different types of feature were used in the proposed model Term Frequency-Inverse Document Frequency and word embedded.

5.2.1. Term Frequency-Inverse Document Frequency(TF-IDF)

It can be defined as the determination of how suitable a word from a series or corpus is to a text. The meaning increases proportionally to the number of occurrences of the given word in the text while the influence of the word frequency within the corpus of texts is compensated equation 3 shows how can calculate the TF-IDF[24].

$$TF = \frac{Number\ of\ term\ in\ document}{total\ number\ of\ terms\ in\ document} \text{.....................}(1)$$

$$IDF = \log\left(\frac{number\ of\ documents\ in\ the\ dataset}{total\ number\ of\ document\ in\ the\ dataset\ that\ contain\ the\ term}\right) \text{.................} (2)$$

$$TF - IDF = TF * IDF \text{ ....}(3)$$

5.2.2. Word embedding

The technique that is used for the conversion of words to vectors of real numbers is called Word Embedding that is a language modeling technique. It represents the words or phrases in vector space which is of several dimensions. The word embedding can be drawn in different methods including: Neural networks, Co- occurrence matrix, Probability models and and so on Word2Vec comes with various models for generating word embedding. These models are the simplest form of neural networks having one input layer, one hidden layer and one output layer. Word2Vec is among the most popular strategies in NLP which introduces ways to express words, which are positioned within a continuous vector space. Word2Vec is an experiment of regressing words to high level of dimensionality in the syntactic relationship between words by Google researchers. The same should be said with meaning, as the main principle of Word2Vec indicates that words with close meaning should be represented with close vectors[25].

5.3. Classification

Depending on the way the text is classified sentiment classification models are of several types such as binary, and multi-class models. The binary classification is where the text is classified either as positive or negative while the multi-classification classifies the text into three sentiments; the positive, the negative and the neutral. The proposed system used multi-class for the classification purpose.

5.3.1. Support vector machine

Support Vector Machine (SVM) is a flexible supervised machine learning algorithm for linear as well as non-linear classification, regression as well as outlier detection. The applications of SVMs can be classification of text documents, remote sensing image classification, email spam filtering, handwritten digit recognition, tumor analysis, facial features recognition and detection of unusual patterns. SVMs are versatile and fast for many uses because the model can handle large attributes and can also deal with non-linear connections. The support vector machine algorithm has a goal to define a hyper plain in an N-dimensional space where N is a number of feature to categorize the points. To split the two classes of the points many possibilities of hyperplanes might be selected. Thus, let our task be to find a plane where the maximum interclass distance is achieved, that is, the separation between the class locations. In order to increase the certainty of data point classification in the future, some reinforcement is therefore made by maximizing the margin distance. [26].

5.3.2. Random Forest

Random forest is an advanced machine learning method that falls under the area of supervised learning approaches. The approach is suitable for solving classification and regression problems in machine learning. It is based on the concept of ensemble learning, which involves solving complex problems with the help of multiple classifiers and improving the performance of the model. Random forest can be defined as a classifier that consists of many decision trees built from different copies of the original dataset, and the scores are averaged to improve the predictability of that dataset. Unlike a single decision tree, random forest studies the final prediction results of all decision trees and determines its result based on the majority of these results. More trees in the forest help to improve the accuracy of the results, and the availability of a large number of trees avoids overfitting [27].

5.3.3. Convolution Neural Network

A Convolutional Neural Network (ConvNet/CNN) is a deep learning algorithm that can accept an input image and recognized and given a specific weight to various aspects/objects in the image and distinguish between them. The pre-processing which is needed for ConvNet is much lesser than what is needed for other classification methods. In primitive methods… the filters are created by hand but for approval, ConvNets have the capabilities of training these filters/characteristics. ConvNet structure was proposed based on the arrangement of the Visual Cortex and is based on the connectivity pattern of neurons in the human brain. Only a small portion of the visual field—referred to as the Receptive Field—is where neurons are responsive to inputs. To encompass the entire visual field, another set of these fields overlaps [28].

5.3.4. Long Short Term Memory

Long Short Term Memory is a modification of Recurrent Neural Network and devised by Hochreiter & Schmidhuber. A typical RNN has one hidden node which is recurrent through time which makes it hard for the network to capture long dependency arithmetic’s. LSTMs model address this problem in a unique way by providing a memory cell which is simply a box that is used to contain information for sometime. LSTM architectures have the ability to capture long-term dependencies in sequential data, thus they are mostly useful in applications such as translation, speech recognition and time series analysis among others. Specifically, the proposed Long Short-Term Memory (LSTM) is suitable for issues involved with big data sequences. They are capable of detecting short and long term relation in temporal sequences. LSTM has a set of recurrently interconnected subnets called memory blocks and three multiplicative elements, the forget gate, the input gate, and the output gate. The switch of LSTM’s gates helps it to get rid of the vanishing gradient and in this way it forms a sort of long term memory scheme. Thus, it can identify what temporal information needs to be sent through the network or what temporal information can be deleted [29].

5.3.5. Bidirectional Long Short Term Memory

Bi-LSTM (Bidirectional Long Short-Term Memory) is one of the types of the recurrent neural network (RNN) and it deals with the sequential data in forward and backward manner. The learner uses both forward and backward tracking as they happen within the LSTM, therefore, the model is in a position to look at past as well as the future of the input sequences. This architecture can be viewed as having two different LSTM one gets the sequence of tokens as they are while the other gets the tokens in reverse order. It is observed that both of these LSTM network returns a probability vector as output and the final output is the fusion of both such probabilities [30].

6. Experiential Results

In this section the proposed model was evaluated to measure the accuracy of each method.

6.1. Dataset Collection

The dataset used in the analysis is the Twitter US Airline Sentiment which was obtained in csv format from Kaggle. Originally, it was obtained from Crowdflower’s Data for Everyone library. In February 2015, 2,800 tweets about each of the major US airlines were collected and scraped from Twitter. Each tweet was then labelled as either positive, neutral, or negative where contributors had to state why they labelled the tweet as negative and state a confidence level as to why they had labelled it thus. E200 has 15 columns and 14,640 rows . The included features are: (tweet id, name, sentiment, tweet text, sentiment confidence score, negative reason, date of tweet, negative reason confidence, airline, sentiment gold, retweet count, tweet coordinates, time of tweet, tweet location, and user time zone).

6.2. Evaluation Metrics

Three different evaluation where used to compute the performance of the proposed model [31].

- Recall: is the fraction of relevant instances that were retrieved. As in equation (4).

$$R = \frac{TP}{TP+FN} \text{ ……………………… (4)}$$

- Precision: is the fraction of relevant instances among the retrieved instances. As in equation(5)

$$P = \frac{TP}{TP+FP} \text{ ………………(5)}$$

- F-Score:It is the harmonic mean of recall and precision or refers to the degree of accuracy of the precision of the searches made from the collected documents. Its range is [0,1]. This is normally an indication that informs us how accurate or specific (It gets and classifies the right number of instances) and reliable or stable (does not misclassify or forgets a large number of instances) our classifier is. Equation (6) shows how F-score computed.

$$F - score = \frac{2*P*R}{R+P} \text{……………..(6)}$$

6.3. Performance of The proposed Model

In this section the performance of the proposed model was evaluated based on precision, recall and F-score. Table(1) show the results of the proposed model. The model based on using word2vec features

Table 1 The Results of the Proposed Model using Word2vec

Method	Precision	Recall	F-Score
SVM	0.88	0.90	0.889
RF	0.86	0.87	0.864
CNN	0.93	0.95	0.939
LSTM	0.94	0.97	0.954
BI-LSTM	0.97	0.98	0.974

It's clear from table-1 that BLLSTM outperformed other due to the fact that the data passes in two way forward and backward. This technique helps to remember the data in good way and not forget what they learned in the previous step.

Table-2 shows the results of the proposed model using TF-IDF features.

Table 2 The Results of the Proposed Model using TF-IDF

Method	Precision	Recall	F-Score
SVM	0.85	0.87	0.859
RF	0.83	0.84	0.83
CNN	0.90	0.92	0.92
LSTM	0.91	0.93	0.91
BI-LSTM	0.92	0.94	0.92

As it's clear from table-2 also BI-LSTM outperformed other methods, but the results of using word2vec features is better than using TF-IDF features. This due to the fact that Word2vec helps to use all terms in the model and not ignore any one, while TF-IDF based on computing the frequency of the each token in the dataset that can lead to gives high importance to some token and less importance to others which reduce model performance. The dataset was divided into training and testing, Where the 70% of the data for training and 30% for testing.

7. Conclusions

The fundamentals of deep learning models and associated methods that are used for sentiment analysis of airline data were covered in this paper. In order to prepare the input data for feeding into deep learning models and machine learning, was used embedding of word and TF-IDF. and performed sentiment analysis by analyzing the architecture of all models used in the system and combining them with TF-IDF and Word Embedding. and was analyzed the architecture of the (DNN, CNN, RNN) and combined them with TF-IDF and word embedding for sentiment analysis. The experimental results shows that BI-LSTM outperformed others methods in both cases when using TF-IDF and word2vec. Conclusion

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. Computer Science Review, 27, 16-32.

[2] Keramatfar, A., & Amirkhani, H. (2019). Bibliometrics of sentiment analysis literature. Journal of Information Science, 45(1), 3-15.

[3] Manosso, F. C., & Domareski Ruiz, T. C. (2021). Using sentiment analysis in tourism research: A systematic, bibliometric, and integrative review. Journal of Tourism, Heritage & Services Marketing (JTHSM), 7(2), 17-27

[4] Unnikrishnan, A., Kannan, B., & Shameem C, M. (2021, December). A Literature Review of Sentiment Evolution. In Proceedings of the International Conference on Systems, Energy & Environment (ICSEE).

[5] Sánchez-Núñez, P., De Las Heras-Pedrosa, C., & Peláez, J. I. (2020). Opinion mining and sentiment analysis in marketing communications: a Science mapping analysis in Web of Science (1998–2018). Social Sciences, 9(3), 23.

[6] Shafqat, W., & Byun, Y. C. (2019). A recommendation mechanism for under-emphasized tourist spots using topic modeling and sentiment analysis. Sustainability, 12(1), 320.

[7] Bordoloi, M., & Biswas, S. K. (2020). Graph based sentiment analysis using keyword rank based polarity assignment. Multimedia Tools and Applications, 79(47).

[8] Rachmawati, R. (2021). Pre-research Study based on Bibliometrics, Deep Learning Research for Aspect-Based Sentiment Analysis. Indonesian Journal of Librarianship, 113-128.

[9] Rajput, G. K., Kumar, A., & Kundu, S. (2020, December). A comparative study on sentiment analysis approaches and methods. In 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 427-431). IEEE.

[10] Mulki, H., Ali, C. B., Haddad, H., & Babaoğlu, I. (2018, June). Tw-star at semeval-2018 task 1: Preprocessing impact on multi-label emotion classification. In Proceedings of The 12th International Workshop on Semantic Evaluation (pp. 167-171).

[11] Farisi, A. A., Sibaroni, Y., & Al Faraby, S. (2019, March). Sentiment analysis on hotel reviews using Multinomial Naïve Bayes classifier. In Journal of Physics: Conference Series (Vol. 1192, No. 1, p. 012024). IOP Publishing.

[12] AlSalman, H. (2020, March). An improved approach for sentiment analysis of arabic tweets in twitter social media. In 2020 3rd International conference on computer applications & information security (ICCAIS) (pp. 1-4). IEEE.

[13] Jemai, F., Hayouni, M., & Baccar, S. (2021, June). Sentiment analysis using machine learning algorithms. In 2021 International Wireless Communications and Mobile Computing (IWCMC) (pp. 775-779). IEEE.

[14] Jemai, F., Hayouni, M., & Baccar, S. (2021, June). Sentiment analysis using machine learning algorithms. In 2021 International Wireless Communications and Mobile Computing (IWCMC) (pp. 775-779). IEEE.

[15] Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731-5780.

[16] Bhatia, P., Ji, Y., & Eisenstein, J. (2015). Better document-level sentiment analysis from rst discourse parsing. arXiv preprint arXiv:1509.01599.

[17] Yang, B., & Cardie, C. (2014, June). Context-aware learning for sentence-level sentiment analysis with posterior regularization. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 325-335).

[18] Fredriksen-Goldsen, K. I., & Kim, H. J. (2017). The science of conducting research with LGBT older adults-an introduction to aging with pride: National health, aging, and sexuality/gender study (NHAS). The Gerontologist, 57(suppl_1), S1-S14.

[19] Gao, Z., Feng, A., Song, X., & Wu, X. (2019). Target-dependent sentiment classification with BERT, IEEE Access 7 (2019) 154290–154299.

[20] Gursoy, U. T., Bulut, D., & Yigit, C. (2017). Social media mining and sentiment analysis for brand management. Global Journal of Emerging Trends in e-Business, Marketing and Consumer Psychology, 3(1), 497-551.

[21] Song, M., Park, H., & Shin, K. S. (2019). Attention-based long short-term memory network using sentiment lexicon embedding for aspect-level sentiment analysis in Korean. Information Processing & Management, 56(3), 637-653.

[22] Cirqueira, D., Pinheiro, M. F., Jacob, A., Lobato, F., & Santana, Á. (2018, December). A literature review in preprocessing for sentiment analysis for Brazilian Portuguese social media. In 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI) (pp. 746-749). IEEE.

[23] Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. Procedia computer science, 17, 26-32.

[24] Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. Journal of documentation, 60(5), 503-520.

[25] Rong, X. (2014). word2vec parameter learning explained. arXiv preprint arXiv:1411.2738.

[26] Wang, L. (Ed.). (2005). Support vector machines: theory and applications (Vol. 177). Springer Science & Business Media.

[27] Biau, G. (2012). Analysis of a random forests model. The Journal of Machine Learning Research, 13(1), 1063-1095.

[28] O'shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

[29] Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. Artificial Intelligence Review, 53(8), 5929-5955.

[30] Su, Y., & Kuo, C. C. J. (2019). On extended long short-term memory and dependent bidirectional recurrent neural network. Neurocomputing, 356, 151-161.

[31] Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process, 5(2), 1.