Academic Journal of Nawroz University (AJNU), Vol.11, No.1, 2022

https://doi.org/10.25007/ajnu.v11n1a1250

A New Model for Emotions Analysis in Social Network Text Using Ensemble Learning and Deep learning

[1] Umran Abdullah Haje, [2] Mohammed Hussein Abdalla, [3] Reben Mohammed Saleem Kurda, [4] Zhwan Mohammed Khalid

[1,2,4] Department of Computer Science, College of Basic Education, University of Raparin, Ranya, Sulaymanaiha, Iraq
[3] Department of Information System Engineering Techniques, Erbil Technical Engineering College, Erbil Polytechnic University, Erbil, Iraq

BSTRACT

Recently, emotion analysis has become widely used. Therefore, increasing the accuracy of existing methods has become a challenge for researchers. The proposed method in this paper is a hybrid model to improve the accuracy of emotion investigation; Which uses a blend of convolutional neural network and ensemble learning. In the proposed method, after receiving the dataset, the data is pre-processed and converted into process able samples. Then the new dataset is divided into two categories of training and analyzing. The proposed model is a structure for machine learning in the form of ensemble learning. It contains blocks consisting of a combination of convolutional networks and basic classification algorithms. In each convolutional network, the base classification algorithms replace the fully connected layer. Evaluate the proposed method, in IMDB, PL04 and SemEval dataset with accuracy, precision, recall and F1 criteria, shows that, on average, for all three datasets, the precision of polarity detection is 90%, the recall of polarity detection is 93%, the F1 of polarity detection is 91% and finally the accuracy of polarity detection is 92%.

KEYWORDS: Emotion Analysis, Social Network, Ensemble Learning. Deep learning, CNN.

1. Introduction

One of the important challenges in emotion analysis is choosing the right algorithm for detecting polarity [1]. According to research, supervised machine learning methods with less complexity and also high performance and low time, in comparison to other methods to build the model, are the best choice for classification [2]. Researchers have used various algorithms to classify or detect emotions [3-5]. Meanwhile, the use of deep learning techniques has shown very high accuracy compared to other classification algorithms [6, 7]. Among the various methods of deep learning, convolutional neural networks are the most broadly used for such problems [6]. However, various methods including recursive neural networks, short-term memory networks, and gateway recursive networks have been used for such problems. [1].

One of the most important disadvantages of convolutional networks is the presence of fully connected layers [8]. This section contains the largest number of learning parameters. These layers are responsible for learning the properties extracted by the convolutional layers. The last layer of fully connected layers is also known as the category score calculator [9]. This layer calculates the outputs for each category by a probabilistic function such as Softmax [10]. The network error rate is then calculated by a cost function [11]. To reduce the amount of network error, the error is propagated into the network and is changed by the after-propagation function of the weights of each neuron in such a way that layers are completely connected and the number of kernels in the convolutional layers to reduce the error [12].

All-connected layers on CNN are computationally heavy and time consuming [13]. On the other hand, they are less accurate than classifiers such as support vector machines. Support vector machine is one of the most widespread monitored classification methods that has a strong foundation and according to research reports, is probably the more accurate method in text classification that is common in emotion classification [14]. On the other hand, according to the explanations provided in the ensemble learning section, because there is no single specific training algorithm that can be the best and most accurate for all applications [15]. It seems; It can use a combination of convolutional networks and initial classification algorithms with a group learning approach. Increase the accuracy of emotion analysis. Accordingly, through a divisive classifier recognizing emotions in social network text, which uses attached networks and the basic classification logarithm as a group.

2. Convolutional Neural Network

CNN can be defined as a learning tool for the purpose of data processing. This model uses network patterns including photos encouraged by the group of intended to learn the spatial hierarchy of structures from low-slung to high. Convolution networks have other applications in addition to image processing. One of these examples is text processing and natural language processing. In this article, it will be use in word processing.

Fig 1 shows a general structure for a convolutional network.

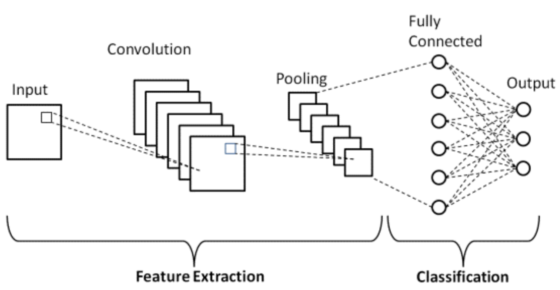


Fig. 1. General structure for a convolutional network

3. Ensemble learning

The purpose of mission categorization ensemble learning was proposed [16]. The main object of the ensemble learning is to teach several base learners then assemble their expectation into a particular output, that should give better implementation than any other members with uncorrelated error on the target data sets [17]. The projected methods have to deal with crucial issues of ensemble learning, while they are sourced from different perspectives [15]. The outputs gained from multiple base learners how to be assembled. How the base learners could be trained. And that's is the substantial factor to find the degree of the achievement of ensemble learning (ensemble diversity) [18].

4. RELATED WORKS

Emotion analysis can be defined as a field of study which focuses on the analysis of views, feelings, assessments, behaviors, tendencies, and emotions represented in a language [19]. Emotion analysis is a tool for understanding the emotions of a text [20]. In fact, the purpose of emotion analysis is to understand if the information in the text is positive or negative [21].

In [7] Researchers use deep learning techniques to analyze emotions to evaluate sustainable transportation. To develop such a model, they semi-automatically generated a set of ideas about transportation, which were then used to fine-tune a large pre-trained language model based on recent deep learning techniques. Experimental results show the robustness of this approach, which can be used for automated processing of large amounts of feedback on transportation. They believe that this approach can help complement official statistics data and traditional studies on transport sustainability.

The paper [30] Offers a framework for analyzing emotions in Thai with the Thai-SenticNet5 collection. This framework employs a variety of attributes, namely embedding the word, part of speech, emotional attributes, and all combinations of these attributes. In addition, in this study integrated convolutional neural network (CNN) and two-way short-term memory (BLSTM) algorithms in different ways and compared it with several other fused combinations. Three Thai datasets were used: ThaiTales, ThaiEconTwitter, and Wisesight. Experimental results show that combining all three features and combining deep learning algorithms can improve overall performance. The best in-depth blend learning was BLSTM-CNN, which achieved F1 scores of 0.7436, 0.7707 and 0.5521 in the ThaiTales, ThaiEconTwitter and Wisesight datasets, respectively. Based on the experimental results, this conclude that the combination of features and hybrid deep learning algorithms can improve the overall performance.

5. Proposed Method

Our proposed method is an applied algorithm for analyzing textual emotions, which includes three main phases: preprocessing, feature extraction, and processing. In the preprocessing phase, tweets are converted to a process able form and sent to the next step. In the next stage, the features of each tweet are taken out using convolutional networks and deep learning techniques. Then a processing data set is formed. Then at the end (processing) using ensemble learning and majority voting strategy, the polarity of the tweets is detected. Figure-2 Block diagram shows the proposed method.

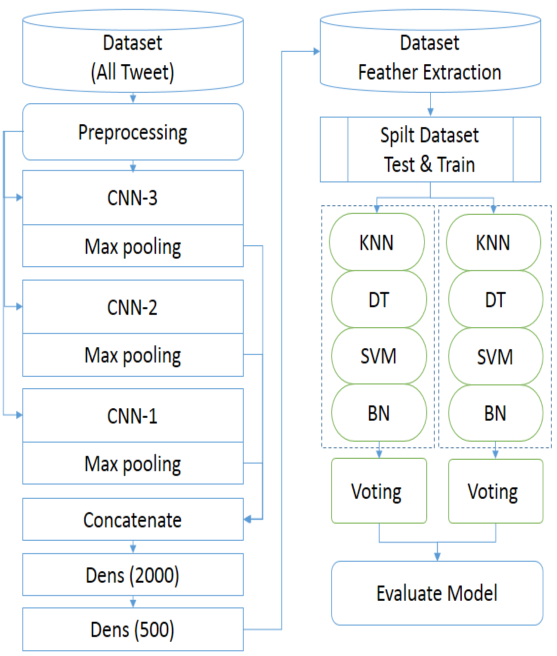


Fig. 2. Block diagram of the proposed method

5.1 Preprocessing

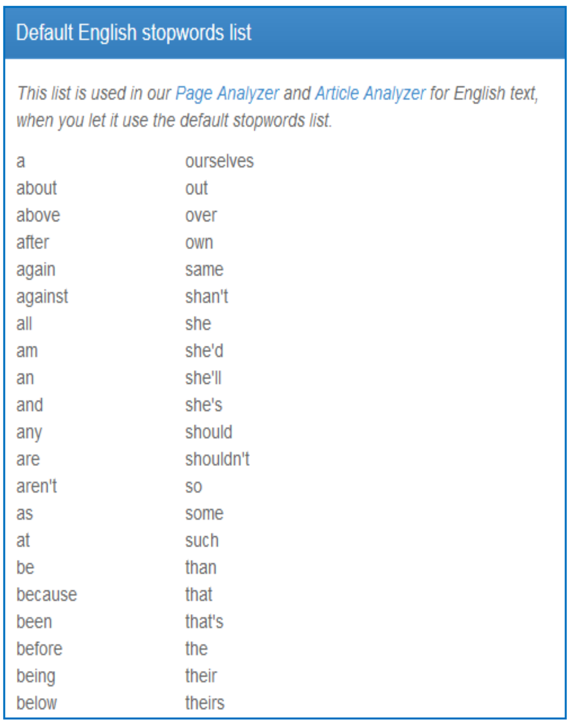
At this stage, the database is converted to a process able form by pre-processing operations. The most important preprocessing steps in the proposed method are to separate the words of the tweets, convert uppercase letters to lowercase letters, return the words to their original roots using the Wordent word chain, and finally tag the samples. Algorithm 1 shows the preprocessing stage actions.

Algorithm 1: Pre-processing operations in the proposed method

- *Tweets tokenizing*
- *Remove stop words*
- *Convert uppercase letters to lowercase letters*
- *Stemming using the Wordent word chain*
- *Labeling of samples*

5.2 Tokenizing

All methods of emotion analysis require identifying the boundaries of words to distinguish them. Boundary of words is done by examining the separating marks. These symbols are: empty space, jump mark, new line mark, ",", ",", ">", "<", "]", "[", "-",



"_" and "/". More sophisticated methods are usually used to improve performance [31].

5.3 Remove stop words

Stop Words are repetitive but unimportant words [32]. If the abundance of words is a proof of their importance [33], it is necessary; Examine the insignificance of a word. Because unimportant words are limited (up to a few hundred words), they are not difficult to find and a list is enough to identify them. Table 1 shows a list of these words. Table. List of redundant words in English

5.4 CONVERT UPPERCASE LETTERS TO LOWERCASE LETTERS

Given that text processing operations are highly dependent on uppercase and lowercase letters, it has been observed that authors do not follow a fixed pattern in writing; Therefore, in order to process texts accurately, it is necessary; All words should be analyzed in the same format after separation. Accordingly, in the proposed method, this process will be performed for all the letters in the dataset.

5.5 STEMMING

The process of reducing words to their origins is called steaming. So "computer" and "compute" and "computing" are reduced to "compute" which is the root [34]. The meaning of the root in this section is not exactly the root of the words in linguistics [35]. Rather, the root is a representative for words that are semantically and syntactically in the same domain. Not all word processing / emotion analysis systems use the same type of root finder [36]. In this research, Wordnet vocabulary and Wordnetleetizer library in Python software will be used to find the roots of words.

5.6 LABELING OF SAMPLES

The next step is to define data classes for each tweet. To model the problem with the proposed cross-method, a target column is created for each data class; This column contains the values 0 and 1. This is called one-hot encoding. Table 3-2 shows an example of defining data classes in the proposed method. If the dataset has 2 data classes (positive polarity and negative polarity); The number of columns will be 2.

Table 2 one-hot-encoding method

Class/ Tweet	*Negative*	*Positive*
Tweet 1	*1*	*0*
Tweet 2	*1*	*0*
Tweet 3	*0*	*1*
Tweet 4	*1*	*0*

5.7 Feature Extraction

The first step in attribute extraction is to embed word vectors. The Word2Vec model will be used for this purpose. Based on the Word2Vec model, tweets are mapped to the corresponding n-dimension vectors. To use the word embedding vector, its weights are copied to the embedding layer. After embedding in this layer, the input enters the network (with any type of structure). This operation is also called word embedding. The basic concept of word embedding is that all the words employed in a language is modeled with a set of decimal numbers (in the form of a vector). Word embedding are n-dimensional vectors that attempt to grasp the meaning of words and their content with their numerical values. Each set of numbers is considered a valid "word vector" that is not necessarily useful to us; it is a set of word vectors useful for our intended applications that understand the connotation of words, the connection between them, and the content of unlike words as is naturally the case. Which has been used then, earned.

In Word2vec method, supported by neural network, a tiny and fixed vector is considered to display all words and texts. This vector is calculated with appropriate numbers in the training phase of the model or training for each word. In this vector, each column does not display a specific word or attribute and only a number is shown. If you consider this vector to be 400, will have a 400-dimensional space in which each word will have an exceptional illustration. For the purpose of the accuracy acceleration of this method, the basic data set needed to teach the model must include the billions of words used within several million texts. After creating vectors related to each word, to display any text or news, the vector of each word used in it can be found and the average of the numbers in each column can be obtained, which will result in a vector for each text or document. After embedding the word vector, convolutional networks are called to extract textual features. The proposed model is a non-sequential method for the feature extraction phase; That is, in that deep learning models (networks) work in parallel. At this stage, three independent networks (A, B and C) are designed to work in parallel. In this structure, after calling the networks, the layers related to each network are also called and Keras is imported. Reshape, flatten, dropout and concatenate are also used to join parallel networks. In the proposed method, convolutional networks, although operating independently of each other; But they have a similar structure. In these networks, the number of filters is 64 and the size of filters [3, 4, 5] is considered. In Figure 3, a 3 × 3 random filter is placed on a log; In this figure, the parts on which the 3 × 3 filter is placed; Multiplied by the filter values and forms a computational output.

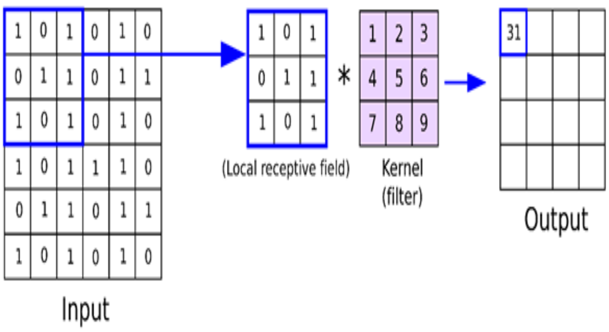


Fig. 3. An example of applying a convolution filter

to a word embedding matrix

The next action to be taken on tweets; is padding; The padding operation means adding additional layers with a value of zero around the attribute vector. This step is performed to normalize the input matrices. If 3 texts are available as input so that each contains 7, 4 and 5 words, respectively. Input matrices will not be the same size. To equalize the size of the input matrices, the largest matrix is considered as the axis and the other matrices are measured with it. In the example mentioned, the largest size for a text is 7; For a matrix of 5 words to be the same size; It is necessary after converting the words to embedded vectors (5 vectors), 2 zero vectors are added to the end of the word vector. In this case, the amount of padding will be equal to 2. Figure 4 below shows this process.

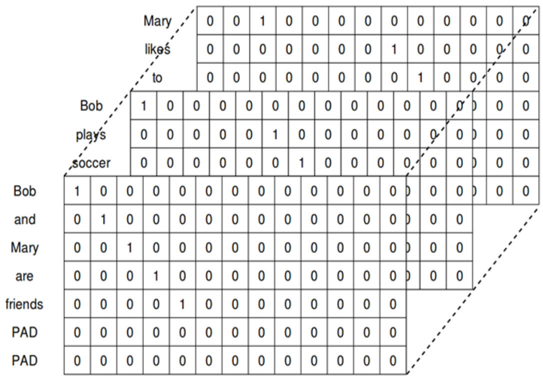


Fig. 4. Shows zero padding in the proposed method

Then Max poling layer is applied on convolutional networks. Pooling is done to reduce the number of parameters and prevent "over-fitting". At this stage, a 2 × 2 filter and maximizing operation have been used. Pooling can also be done with other operators such as averaging, but maximizing has had the highest success rate. Figure 5 shows the pooling operation with a 2. 2 filter. Based on this example, a 4. 4 matrices taken as input is divided into 4 2 × 2 sections and the maximum value of each section is placed at the output.

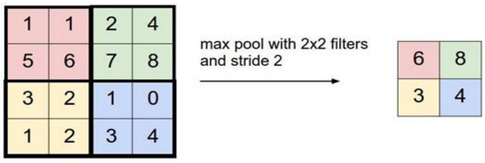


Fig. 5. pooling operations in the proposed method

In the next step, the network output (Max poling) must be connected (Concatenate); The proposed function for this purpose is Flatten. Then, using the Dense function, the number of features extracted that are used in the classification step; With two consecutive operations, it is reduced first to 2000 and finally to 500 (Dens).

After the proposed model structure was formed; The machine learning step is done. In this step, Adam [37] model and accuracy evaluation index are used to update the weights and optimize the error. Adam optimization is a procedure that functions to update network weights frequently relies on training data instead of the outdated random gradient descent method. One is obtained by calculating the evolutionary moment. This algorithm is used for deep learning [38]. Adam combines the benefits of two other random descending slope extensions with the AdaGrad Algorithm, which maintains the speed of learning in each parameter, improving performance on scattered slope problems [39]. Root Mean Square Propagation (RMSProp) also sustains the learning rate of each parameter, which is weighted according to the mean of new magnitudes. Offline and non-fixed problems, this algorithm does well [40].

Adaptive torque estimation (Adam) is another method that calculates the learning rate according to the data. In addition to storing the exponential mean exponential squared of the previous gradients, ie v_t, like the AdaDelta and RMSprop methods, the Adam method also preserves the exponential average exponential damping of the gradients m_t, like the acceleration method:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \quad (1)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \quad (2)$$

m_t and v_t are the estimations of the first torque (mean) and the second torque (decentralized variance) of the gradients, respectively, which is why the method is named. Because the initial state of the m_t and v_t vectors is zero, the method makers observed that the results tend to be zero, especially in the first steps and particularly when the damping rate is small (or in other words β_1 and β_2 are close to 1). They solved this problem with modified first and second torque estimates:

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \quad (3)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^2} \quad (4)$$

Then they used these two formulas to calculate the changes in the parameters, which gives the formula for the changes for this method:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t}+\epsilon}\hat{m}_t \quad (5)$$

The developers have proposed a method of default values of 0.9 for β_1, 0.999 for β_2 and 10^{-8} for ϵ.They have shown empirically that the Adam method works well in practice and is superior to other adaptive learning methods.

5.8 Processing and Emotion Analysis

At this stage, the process of emotion analysis is achieved using ensemble learning technique. For this purpose, after separating the data set into two groups of training and testing, the machine learning process is done with ensemble learning and the polarity of the tweets is detected. The important steps of this phase are described below. In ensemble learning, there are a number of independent classifiers that perform the machine learning step independently of each other and using the training data set. The training dataset is sent to each of the classifiers. Classifiers perform the machine learning phase based on their structure. The number of classifiers in this article will be n. Figure 6 shows the process of mapping a dataset into several categories. The proposed classifiers for this step are SVM, KNN, BN and DT.

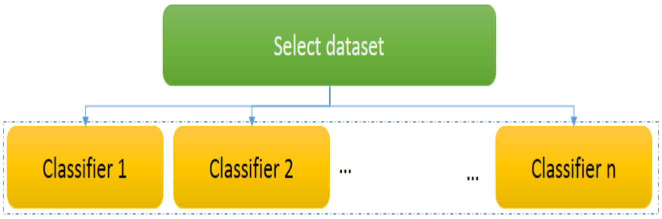


Fig. 6. Mapping the data set into several categories in the proposed method

In the proposed method, the polarity of tweets will be determined based on the majority vote. Figure 7 shows an example for a majority vote. In this strategy, each of the learner models (classifiers) independently detects the polarity of the tweets. The model votes are counted to determine the polarity of the tweet, and the polarity that has the most votes; Tweets are assigned to it.

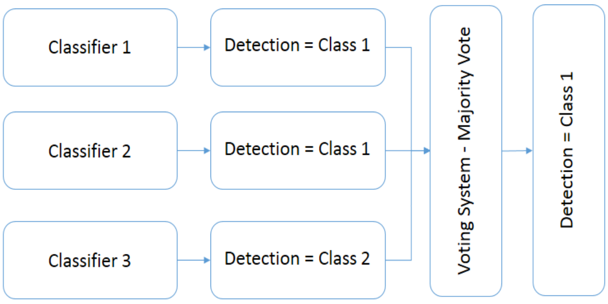


Fig. 7. Majority voting strategy

6. Dataset

The datasets used in this paper are IMDB, PL04 and SemEval, which are public and available datasets. Each of the mentioned data sets has two classes, positive and negative, so they are called 2- classes. Table 3 shows the specifications of this set.

Table 3: dataset specifications

Average words	Negative emotions	Positive emotions	Number of classes	Number of samples	Dataset
255	**25000**	**25000**	**2**	**50000**	*IMDB*
723	**1000**	**1000**	**2**	**2000**	*PL04*
22	**932**	**2509**	**2**	**3441**	*SEmEVAl*

7. Evaluation criteria

In this study, accuracy, precision, recall and F1 have been used to assess the planned method. Table 4 introduces these indicators and the relationships for them.

Table 4 Evaluation criteria

Formula	Description
$Acc = \frac{TP + TN}{TP + FP + FN + TN}$	Accuracy: The most common performance criterion for classification algorithms. It can be defined in terms of the number of correct predictions made relative to all the predictions made.
$Pre = \frac{TP}{TP + FP}$	Precision: The main focus of this criterion is on the correctness of positive diagnoses by the algorithm. In fact specifies; The algorithm detects what percentage of positive classes.
$Rec = \frac{TP}{TP + FN}$	Recall: This criterion can be defined in terms of the number of positives returned by the model.
$F1 = \frac{2 * Pre * Rec}{Pre + Rec}$	F1: This criterion is the harmonic mean of precision and recall. The F1 criterion is the weighted average of precision and recall.

In the above relations, TP shows the quantity of models whose class is positive and the proposed method correctly identifies their class. FP displays the number of examples whose class is negative; however, the proposed method incorrectly identifies their class. Positive, FN indicates the number of samples whose actual class was positive, nonetheless the proposed method incorrectly identified their class as negative, and TN indicates the number of samples whose class is negative, and the proposed method is correct the class recognized them negatively.

8. Results

In this section, the efficiency of the model in analyzing textual emotions with different evaluation criteria is examined. The results of the experiments are shown below.

8.1 precision

The results in the precision criterion show that the proposed method detects positive and negative emotions with a precision of 0.91 in the IMDB dataset, with a precision of 0.92 in the PL dataset and with an accuracy of 0.88 in the SemEval dataset. Table 5 displays the concluding points of this study. The conclusion shows that as the length of messages increases, the precision of emotion recognition also increases.

Table 5 Evaluation of the proposed method in the precision criterion

SemEval	*PL*	*IMDB*	*Class*
0.91	**0.94**	***0.89***	**Positive**
0.85	**0.9**	**0.93**	**Negative**
0.88	**0.92**	**0.91**	**Total**

The results in this study show that the highest precision was for PL dataset and the lowest precision was for SemEval dataset. Figure 8 Displays the results.

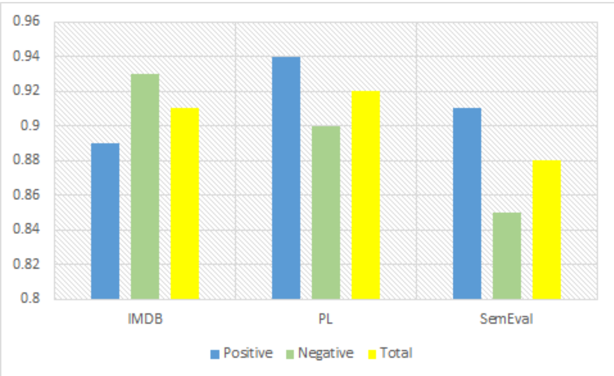


Fig. 8.Evaluation of the proposed method in the accuracy criterion

8.2 Recall

The results in the recall criterion show that the proposed method detects positive and negative emotions with a recall of 0.92 in the IMDB dataset, with a recall of 0.98 in the PL dataset and with a recall of 0.90 in the SemEval dataset. Table 6 demonstrates this study's result. The results show that as the length of messages increases, the recall of emotion recognition also increases.

Table 6 Evaluation of the proposed method in the recall criterion

SemEval	*PL*	*IMDB*	*Class*
0.93	**0.98**	***0.91***	**Positive**
0.87	**0.98**	**0.92**	**Negative**
0.90	**0.98**	**0.92**	**Total**

The results in this study show that the highest recall was for PL dataset and the lowest recall was for SemEval dataset. Figure 9 Displays the results.

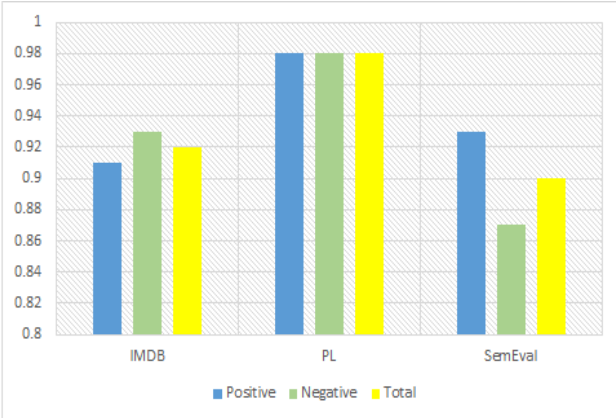


Fig. 9. Evaluation of the proposed method in the recall criterion

8.3 F1

The results in the F1 criterion show that the proposed method detects positive and negative emotions with a F1 of 0.91 in the IMDB dataset, with a F1 of 0.95 in the PL dataset and with a F1 of 0.89 in the SemEval dataset. Table 7 shows the results of this study. The results show that as the length of messages increases, the F1 of emotion recognition also increases.

Table 7 Evaluation of the proposed method in the F1 criterion

SemEval	*PL*	*IMDB*	*Class*
0.91	**0.96**	***0.9***	**Positive**
0.85	**0.94**	**0.93**	**Negative**
0.89	**0.95**	**0.91**	**Total**

The results in this study show that the highest F1 was for PL dataset and the lowest F1 was for SemEval dataset. Figure 10 Displays the results.

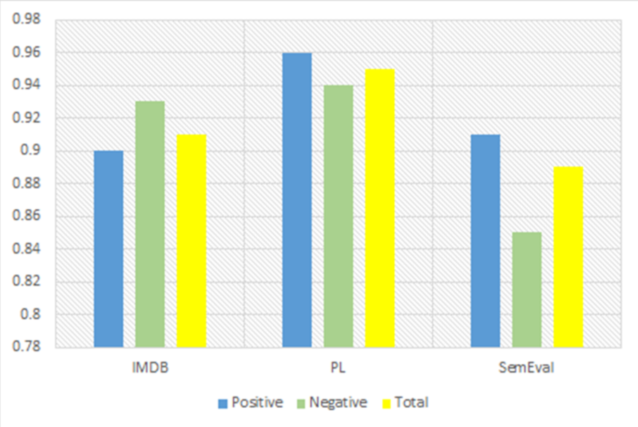


Fig. 10. Evaluation of the proposed method in the F1 criterion

8.4 accuracy

The results in the accuracy criterion show that the proposed method detects positive and negative emotions with an accuracy of 0.92 in the IMDB dataset, with an accuracy of 0.95 in the PL dataset and with an accuracy of 0.88 in the SemEval dataset. Table 8 displays the results of this study. The results show that as the length of messages increases, the accuracy of emotion recognition also increases.

Table 8 Evaluation of the proposed method in the accuracy criterion

SemEval	*PL*	*IMDB*	*Class*
0.91	0.98	*0.90*	Positive
0.85	0.92	0.94	Negative
0.88	0.95	0.92	Total

The results in this study show that the highest accuracy was for PL dataset and the lowest accuracy was for SemEval dataset. Figure 11 Displays the results.

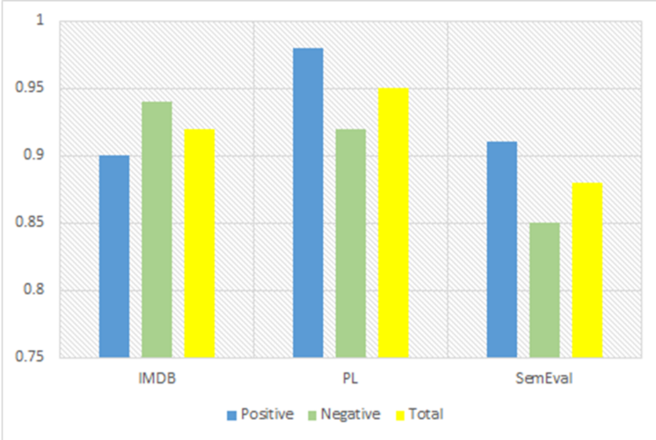


Fig. 11. Evaluation of the proposed method in the accuracy criterion

Examination of F1 emotion analysis by the proposed method and DLEL [41] shows that the proposed method has improved the detection accuracy by 0.1 in all three datasets of IMDB, PL and SemEval. But the point to consider in both the proposed method and its basis is that with increasing the length of the messages, the accuracy of detection has also increased. Figure 12 shows the results of this study.

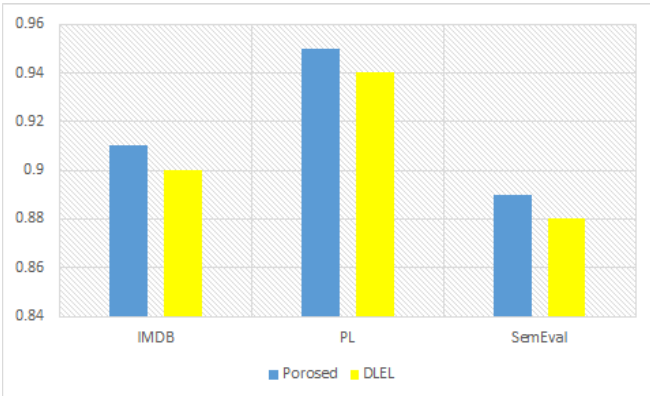


Fig. 12. Compare results with similar work

9. Conclusion

In conclusion as social media has become widespread among the individuals, the growth of information (especially textual) has become a process and this has faced challenges in sorting and extracting knowledge. User data in cyberspace and social media has valuable information about the environment and analyzing the emotions in them can help managers achieve organizational or social goals. Therefore, the analysis of emotions from texts has been considered by researchers in recent years. Emotion analysis means discovering and recognizing people's positive or negative feelings about an issue or product in the texts. Emotion analysis is the process of analyzing opinions, feelings, evaluations, behaviors, and tendencies and emotions written in a written language. In this article, using deep learning and ensemble learning, a new approach to emotion analysis was proposed. In the proposed method, after receiving the data set, the data is pre-processed and converted into workable samples. Then the new data set is categorized into two groups of training and test and the training data is sent to the proposed model. The proposed model is a machine learning model in the formula of ensemble learning in which there are blocks consisting of a combination of convolutional networks and basic classifiers. In each convolutional network, the base classifier replaces the fully connected layer. To evaluate the proposed method, used IMDB, PL and SemEval dataset and accuracy, precision, recall and F1 criteria. The conclusion of the study clarifies that the proposed model has always performed better than the similar model. Also, with increasing the length of messages, the accuracy of detection has increased.

10. References

1. R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," Information Processing & Management, vol. 58, no. 1, p. 102435, 2021.
2. D. A. Pereira, "A survey of sentiment analysis in the Portuguese language," Artificial Intelligence Review, vol. 54, no. 2, pp. 1087-1115, 2021.
3. K. Sailunaz and R. Alhajj, "Emotion and sentiment analysis from Twitter text," Journal of Computational Science, vol. 36, p. 101003, 2019.
4. A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency, "Tensor fusion network for multimodal sentiment analysis," arXiv preprint arXiv:1707.07250, 2017.
5. J. K. Rout, K.-K. R. Choo, A. K. Dash, S. Bakshi, S. K. Jena, and K. L. Williams, "A model for sentiment and emotion analysis of unstructured social media text," Electronic Commerce Research, vol. 18, no. 1, pp. 181-199, 2018.
6. C. Dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, 2014, pp. 69-78.
7. A. Serna, A. Soroa, and R. Agerri, "Applying Deep Learning Techniques for Sentiment Analysis to Assess Sustainable Transport," Sustainability, vol. 13, no. 4, p. 2397, 2021.
8. A. Dhillon and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," Progress in Artificial Intelligence, vol. 9, no. 2, pp. 85-112, 2020.
9. N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in 2017 International Conference on Communication and Signal Processing (ICCSP), 2017, pp. 0588-0592: IEEE.
10. G. Yao, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition," Pattern Recognition Letters, vol. 118, pp. 14-22, 2019.
11. S. Sony, K. Dunphy, A. Sadhu, and M. Capretz, "A systematic review of convolutional neural network-based structural condition assessment techniques," Engineering Structures, vol. 226, p. 111347, 2021.
12. Hussein, M. and Özyurt, F., 2021. A New Technique for Sentiment Analysis System Based on Deep Learning Using Chi-Square Feature Selection Methods. Balkan Journal of Electrical and Computer Engineering, 9(4), pp.320-326.
13. N. F. Waziralilah, A. Abu, M. Lim, L. K. Quen, and A. Elfakharany, "A review on convolutional neural network in bearing fault diagnosis," in MATEC Web of Conferences, 2019, vol. 255, p. 06002: EDP Sciences.
14. J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," Neurocomputing, vol. 408, pp. 189-215, 2020.
15. X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," Frontiers of Computer Science, vol. 14, no. 2, pp. 241-258, 2020.
16. X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu, "An adaptive ensemble machine learning model for intrusion detection," IEEE Access, vol. 7, pp. 82512-82521, 2019.
17. C. Zhang and Y. Ma, Ensemble machine learning: methods and applications. Springer, 2012.
18. H. M. Gomes, J. P. Barddal, F. Enembreck, and A. Bifet, "A survey on ensemble learning for data stream classification," ACM Computing Surveys (CSUR), vol. 50, no. 2, p. 23, 2017.
19. J. Serrano-Guerrero, J. A. Olivas, F. P. Romero, and E. Herrera-Viedma, "Sentiment analysis: A review and

comparative analysis of web services," Information Sciences, vol. 311, pp. 18-38, 2015.

20. M. Z. Asghar, A. Khan, S. Ahmad, and F. M. Kundi, "A review of feature extraction in sentiment analysis," Journal of Basic and Applied Scientific Research, vol. 4, no. 3, pp. 181-186, 2014.
21. H. Peng, E. Cambria, and A. Hussain, "A review of sentiment analysis research in Chinese language," Cognitive Computation, vol. 9, no. 4, pp. 423-435, 2017.
22. A. Krouska, C. Troussas, and M. Virvou, "Deep learning for twitter sentiment analysis: the effect of pre-trained word embedding," in Machine Learning Paradigms: Springer, 2020, pp. 111-124.
23. M. Umer, I. Ashraf, A. Mehmood, S. Kumari, S. Ullah, and G. Sang Choi, "Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model," Computational Intelligence, vol. 37, no. 1, pp. 409-434, 2021.
24. H. Juwiantho, E. I. Setiawan, J. Santoso, and M. H. Purnomo, "Sentiment Analysis Twitter Bahasa Indonesia Berbasis Word2Vec Menggunakan Deep Convolutional Neural Network," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 7, no. 1, pp. 181-188, 2020.
25. W. L. Lim, C. C. Ho, and C.-Y. Ting, "Tweet sentiment analysis using deep learning with nearby locations as features," in Computational Science and Technology: Springer, 2020, pp. 291-299.
26. I. Kandasamy, W. Vasantha, J. M. Obbineni, and F. Smarandache, "Sentiment analysis of tweets using refined neutrosophic sets," Computers in Industry, vol. 115, p. 103180, 2020.
27. R. P. Mehta, M. A. Sanghvi, D. K. Shah, and A. Singh, "Sentiment analysis of tweets using supervised learning algorithms," in First International Conference on Sustainable Technologies for Computational Intelligence, 2020, pp. 323-338: Springer.
28. N. Chouchani and M. Abed, "Enhance sentiment analysis on social networks with social influence analytics," Journal of Ambient Intelligence and Humanized Computing, vol. 11, no. 1, pp. 139-149, 2020.
29. A. H. Ombabi, W. Ouarda, and A. M. Alimi, "Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks," Social Network Analysis and Mining, vol. 10, no. 1, pp. 1-13, 2020.
30. K. Pasupa and T. S. N. Ayutthaya, "Hybrid deep learning models for thai sentiment analysis," Cognitive Computation, pp. 1-27, 2021.
31. A. Hassan and A. Mahmood, "Deep learning approach for sentiment analysis of short texts," in 2017 3rd international conference on control, automation and robotics (ICCAR), 2017, pp. 705-710: IEEE.
32. A. Agrawal and A. An, "Kea: Sentiment analysis of phrases within short texts," in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 380-384.
33. C. Bouras and V. Tsogkas, "Noun retrieval effect on text summarization and delivery of personalized news articles to the user's desktop," Data & Knowledge Engineering, vol. 69, no. 7, pp. 664-677, 2010.
34. A. K. Dipongkor, M. A. Nashiry, K. A. Abdullah, and R. S. Ritu, "A Study on Bengali Stemming and Parts-of-Speech Tagging," in Emerging Technologies in Data Mining and Information Security: Springer, 2021, pp. 35-44.
35. Z. Nanli, Z. Ping, L. Weiguo, and C. Meng, "Sentiment analysis: A literature review," in 2012 International Symposium on Management of Technology (ISMOT), 2012, pp. 572-576: IEEE.
36. I. K. Raharjana, D. Siahaan, and C. Fatichah, "User Stories and Natural Language Processing: A Systematic Literature Review," IEEE Access, vol. 9, pp. 53811-53826, 2021.
37. K. K. Chandriah and R. V. Naraganahalli, "RNN/LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting," Multimedia Tools and Applications, pp. 1-15, 2021.
38. Abdalla, M. H., & Karabatak, M. (2020, June). To review and compare evolutionary algorithms in optimization of distributed database query. In 2020 8th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-5). IEEE. [39] M. Yaqub et al., "State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images," Brain Sciences, vol. 10, no. 7, p. 427, 2020.
39. D. O. Melinte and L. Vladareanu, "Facial expressions recognition for human–robot interaction using deep convolutional neural networks with rectified adam optimizer," Sensors, vol. 20, no. 8, p. 2393, 2020.
40. O. Araque, I. Corcuera-Platas, J. F. Sanchez-Rada, and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," Expert Systems with Applications, vol. 77, pp. 236-246, 2017.