

Comparison of Neural Network Models for Nostalgic Sentiment Analysis of YouTube Comments

Seda Postalcioglu  Senem Aktas 

Bolu Abant İzzet Baysal University, Department of Computer Engineering, Bolu, Turkey

ABSTRACT

For this study Sentiment Analysis (SA) is applied for the music comments using different Neural Network (NN) Models. SA is commonly used for Natural Language Processing (NLP). With the help of NLP, the evaluations / tips about the future can be obtained by analyzing the correspondences and comments. The aim of the study is to draw conclusions from the comments made under the songs whether they are nostalgic. Data is captured using the YouTube Data API. Data extraction is done by entering the link of the song whose comments will be taken. CSV files are obtained and then labeled as nostalgic and non-nostalgic. Different neural network models as MLPNN (Multi-Layer Perceptron Neural Network), CNN (Convolutional Neural Network), RNN-LSTM (Recurrent Neural Network-Long Short-Term Memory) are applied for sentiment analysis. Their performances are analyzed. MLPNN, CNN, RNN-LSTM performance results are 78%,88%,88%, respectively.

Keywords:

CNN, MLPNN, NLP, RNN, YouTube comments.

Article History:

Received: 2020/05/15

Accepted: 2020/09/01

Online: 2020/09/30

Correspondence to: Seda Postalcioglu,

E-mail: postalcioglu_s@ibu.edu.tr;

Phone: +90 3742541000

INTRODUCTION

With the rapid development of social networking sites, people need to explain their opinions and thoughts. People share not only information with websites but also they begin to express their thoughts and feelings more easily [1]. They expressed these posts by tweeting them on Twitter, posting on Instagram or commenting on social media. Based on this case, the feelings and thoughts of the interpreters can be analyzed [1]. Reference 11 can be shown as an example.

It is possible to analyze such comments or posts according to some criteria such as what team hits, political opinion, positive/negative/neutral content of comments. This kind of analyzes may affect future production progress by enabling the public to learn the opinions, parties, and thoughts before the political election or to make a customer analysis for the products produced. It has a wide usage area. These analyses are carried out for many different purposes. With these studies, it is tried to predict the future. These analyses are referred to as "Sentiment Analysis" in terminology [1]. Sentiment analysis is one of the techniques commonly used in Natural Language Processing (NLP). NLP helps us to understand the content of a text [2]. There are different

approaches based on natural language processing and machine learning techniques to get the features for the sentiment analysis [3].

In this study sentiment analysis is studied for the music comments. Music is universal, it has no language, and it has always been the best way to express emotions, at least that is how many people think. When you listen to a song, you can remember your memories of your childhood/youth or find something from yourself.

One of the most used music video platforms today is YouTube. It has a very large archive and is free. We can listen to any song from the 1960s or a song from today. YouTube also provides a section that allows you to post comments under these music videos. Here, feelings and opinions about the music listened to are shared.

The aim of the study is to draw conclusions from the comments made under the songs whether they are nostalgic. Comments captured using the YouTube Data API are tagged according to whether they are nostalgic in the CSV file. Choices are made from comments were taken before tagging, comments that will not be useful for analysis are deleted from the CSV file. "Natural lan-

guage processing" (NLP) operations are applied to the data. These data are divided into two as education data and test data. After training the models with training data, performance evaluation is made with test data. The difference of this study from the other articles is compared with the performance of NN using deep learning technique.

In this study, nostalgic sentiment analysis is done using different neural network models as MLPNN, CNN, RNN-LSTM. These models are performed on the database which is constructed by us. Performance results are compared.

MATERIAL AND METHODS

It is very important to create a suitable database for the study. Nostalgia literally means longing for the past. In the labeling of the comments, it is labeled as nostalgic in the comments that contain information about the real meaning of the word nostalgia as well as "remembering the past, childhood-youth years, memories of those years".

To find nostalgic comments, it is necessary to look at the comments under the old songs. The relevant website is given in reference 4 has been used to easily access such songs. CSV files are obtained and then labeled as nostalgic and non-nostalgic. The data is partially ready.

These prepared data are exposed to some pre-processing stages before modeling such as removing punctuation, numbers, stopwords, extra spaces. After a pre-processing dataset is used in neural network models. 80% of the dataset obtained is used as training data to train the model, and the remaining 20% is used as test data to test the model. Fig. 1 shows the dataset for the study. Dataset size is 1500.

Sentiment Analysis needs a pre-processor to transform a word into a vector [2]. This process is called "Word Embedding"[2]. Extracting features from words for NLP, there are two popular methods; word2vec and GloVe (The Glo-

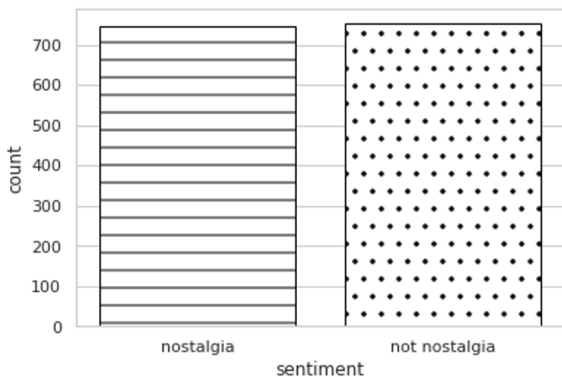


Figure 1. Dataset for the study

bal Vectors for Word Representation) [3]. Word embedding maps the words in vocabulary into real-valued vectors [5].

The word vectors shows semantics and their dimension is usually low, this transformation obtains to calculate the connection among words and dimensionality reduction for efficient representation [6,5]. The NN takes words from a vocabulary as input and then embeds them as vectors into a lower-dimensional space, which is denoted to as Embedding Layer [5]. Fig. 2 shows the word embedding model using a NN [5].

The word embedding model used in this study. The GloVe is utilized. GloVe stands for global vectors for word representation. Fig. 3 shows the system architecture. Pre-processing and feature extraction are important for the study.

Word Embedding is applied as the Glove technique. For pre-processing stages, tokenization, stop word, noise removal is used. Tokenization breaks a stream of text into words, phrases, symbols, or other meaningful elements named tokens. This process provides the analysis of the words in a sentence [7]. The most common technique to deal with "a", "about", "above", "across", "after", "afterward", "again" which is not contain important impact for the sentence is Stop Words. This method removes them from the text [7]. For this reason, stop words are applied for the study. Text datasets cover redundant characters such as punctuation and special fonts. Critical punctuation and special fonts are significant for people understanding of text [7]. It can be causes a problem for classification algorithms. So for this study noise removal is applied. An example of the dataset before pre-processing is given in Fig. 4. An example of the dataset after pre-processing is given in Fig. 5.

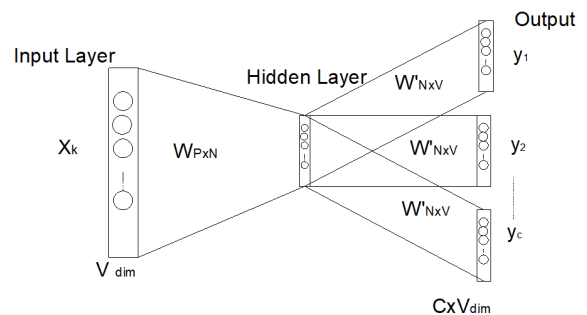


Figure 2. Word Embedding Model [5]

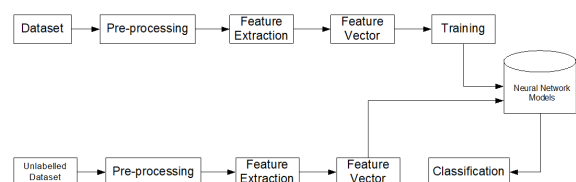


Figure 3. System Architecture

Index	sentiment	comment
0	not nostalg...	I have listened to and enjoyed this song for many ...
1	nostalgia	My Dad and Mon used to play this all the time and ...
2	not nostalg...	He was a singer with a golden voice that I love to...
3	nostalgia	The mist beautiful voice ever I listened to hin wh...
4	nostalgia	I have most of Mr. Reeves songs. Always love his ...
5	not nostalg...	30 day leave from 1st tour in Viet Nam to convince...
6	nostalgia	listening to his songs reminds me of my mum who pl...
7	nostalgia	Every time I heard this song as a child, I used to...
8	nostalgia	My dad loved listening to Jim Reeves, when I was a...
9	nostalgia	i HAVE ALSO LISTENED TO Jim Reeves since childhood...
10	not nostalg...	Wherever you are you always in my heart
11	not nostalg...	Elvis will always be number one no one can come cl...
12	not nostalg...	ill bet if they begin to play this song on the rad...

Figure 4. Dataset before pre-processing

Index	Type	Size	Value
0	str	1	listened enjoyed song many years years old widower driving distant cit...
1	str	1	dad mon used play tine great reeves fans grew music love lost nana num...
2	str	1	singer golden voice love hear time great fan age years days still alth...
3	str	1	mist beautiful voice ever listened kid still love singing never forgot...
4	str	1	mr reeves songs always love smooth voice comforting sounds like peopl...
5	str	1	day leave st tour viet convince mary lou marry came back nd tour years...
6	str	1	listening songs reminds mum played everyday
7	str	1	every time heard song child used cry reminded funerals got fear death ...
8	str	1	dad loved listening reeves teenager back tine changes things
9	str	1	listened reeves since childhood nostalgic feeling comes tunes love fee...
10	str	1	wherever always heart
11	str	1	elvis always number close taking place king always king
12	str	1	bet begin play song radio new generation could still hit goes anytime ...
13	str	1	conn came center wear high school taxic nrlfriend loved forbidden kn...

Figure 5. Dataset after pre-processing

The system has two stages as training and testing. For training stage, to cope with over fitting problem, L2 regularization is preferred. The aim of L1 and L2 regularizations are avoiding over-fitting problem. L2 is the most common regularization in deep learning [12]. After training, the test dataset is applied to the system. The performance results of the networks are calculated using Equation (1). TP is the true positive, TN is the true negative, FP is false positive and FN is the false-negative predictions.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (1)$$

RESULTS AND DISCUSSION

Nowadays, there are different neural networks structures. Each of them has different architectural and working mechanisms. Specially used neural networks have been developed for various machine learning processes such as image processing, natural language processing etc. In this study, three different neural networks are used. These are MLPNN, CNN, RNN-LSTM.

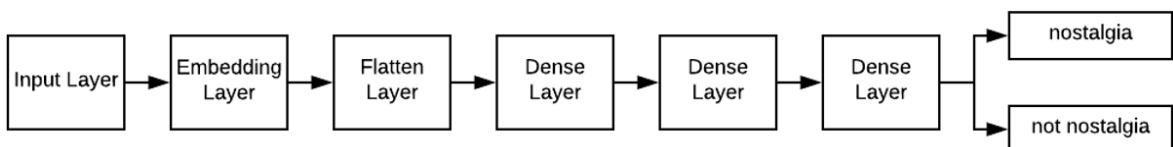


Figure 7. Architecture of the MLPNN Model

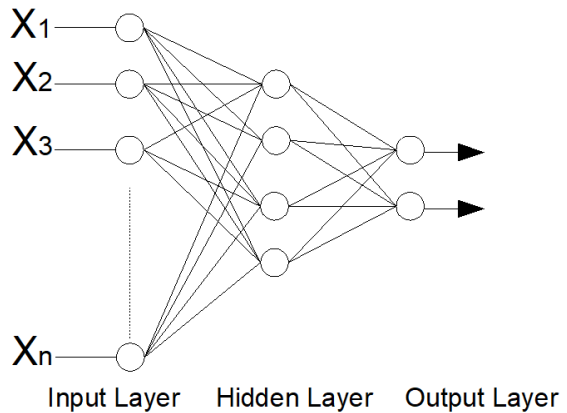


Figure 6. MLPNN Model

MLPNN is a feed-forward neural network. The structure has units which has one-way connections to other units. The units can be arranged in layers so that connections go from one layer to a later layer [13]. It has one or more layers called a hidden layer. MLPNN works especially effectively in classification, regression prediction problems. It can be used for image, text and many other different datasets. Linearly not separable problems are solved by MLPNN [8]. MLPNN Model is shown in Fig. 6.

The Embedding layer was used as the first layer in all our neural networks. This layer turns positive integers (indexes) into fixed-sized dense vectors to reduce the costs of one-hot encoded vectors, which are high-dimensional and sparse. The flatten layer was used to flatten the input, converting the layer to the 1D array so it could feed the layer. Then three dense layers were used to connect each node of the current layer to the next node of the next layer. The single-unit last dense layer gives the classification result. MLPNN model architecture is given in Fig. 7.

Table 1 shows the MLPNN properties and performance results. ReLu is used as an activation function. Adam is used as an optimizer. Mean Squared Error is used. The batch size technique is used for modeling. The batch size is a hyper parameter of gradient descent and it controls the number of training samples [9]. Fig. 8 shows the MLPNN results for Sentiment Analysis (SA) as to accuracy and loss, respectively. Accuracy is calculated as 78%.

CNN, which is widely used especially in computer vision tasks, gives very successful results. Although they have been specially developed and used for image data, they are

also used for sentiment analysis and classification processes obtained using text data. CNNs are constructed with convolutional layers [14]. Reference 15 states that convolutional networks are suitable for recognizing or rejecting shapes [15].

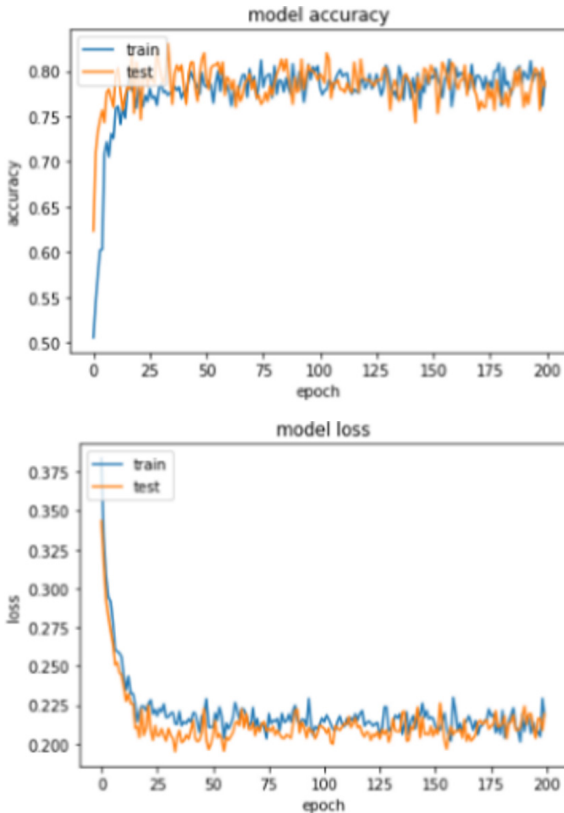


Figure 8. MLPNN results for Sentiment Analysis

As can be seen in Fig. 9, CNN architecture has Conv1D layers and Global Max Pooling layer different from MLPNN architecture. Conv1D was applied because the texts are one-dimensional data strings. This layer creates a convolution kernel.

CNN (Convolutional Neural Network) contains three layers as the Convolutional, Pooling and Full-Connected Layers [6]. Convolution layer extracts the features of the input by using a slide filter and applies the scalar product [6]. Feature map is the output of convolution [6]. Matrices obtained by convolution operators are transformed with an activation function. ReLu (Rectified Linear Unit) is used as an activation function. ReLu is given in Equation (2) and its effect is given in Fig. 10 [6]. For the output layer, the sigmo-

id function is used as an activation function. The sigmoid function is given in Equation (3).

$$ReLU(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad (2)$$

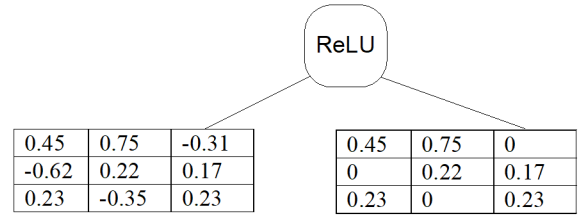


Figure 10. ReLU

$$f(x) = 1/(1 + e^{-x}) \quad (3)$$

The second one is the Pooling layer. It reduces the size of the feature map. Pooling works by sliding a window over the input. Finally, the Fully Connected layer filters high-level representation of the input and converts them into votes. Fully connected is a way to connect layers in two layers together [6].

Fig. 11 shows the CNN results for SA as accuracy and loss, respectively. Table 1 gives the CNN model properties. Accuracy is calculated as 88%.

LSTM algorithm is very popular for NLP [2]. It is one of the RNN models that can learn sequential data [2]. For training stage, RNN is hard to report the vanishing gradient problem. By using its internal memory, RNN sends the output it produces back to the network to obtain the next new output. Therefore, all the inputs in RNN are dependent. LSTM-RNN is advanced to solve the vanishing gradient problem and capture long-term dependency. So, LSTM-RNN is applied to language modeling [10]. The simple design of RNN is shown in Fig. 12 [10]. We consider $X=(x_1,x_2,\dots,x_T)$ as the input word sequence and $Y=(y_1,y_2,\dots,y_T)$ as the hidden sequence. RNN's hidden layer is the recurrent layer. Recurrent Layer is used to capture the contextual information of sentence [10].

LSTM-RNN adds output gate, forget gate, input gate and memory cell and an important information can be store over long time duration [10]. For this model, the tanh activation function is used. The activation function is given in Equation (4).

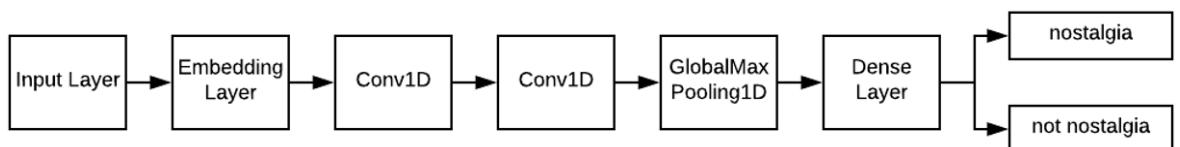


Figure 9. Architecture of the CNN Model

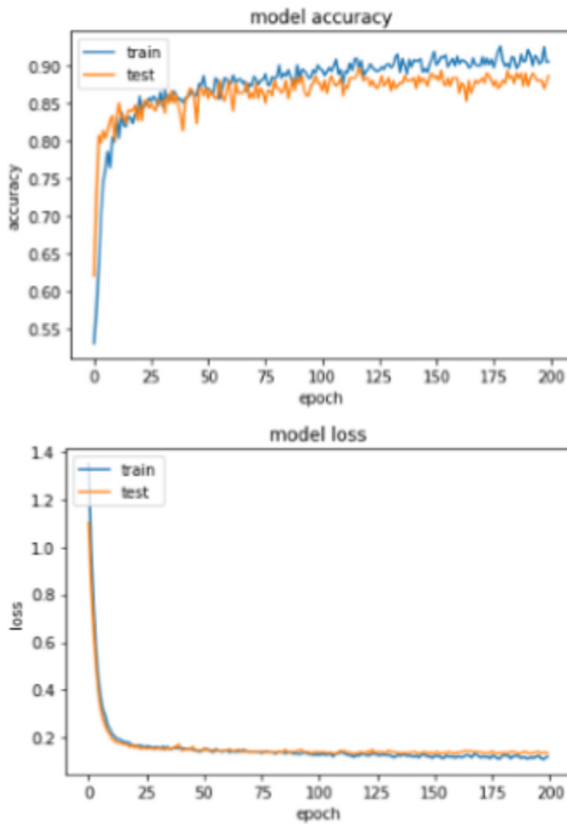


Figure 11. CNN results for Sentiment Analysis

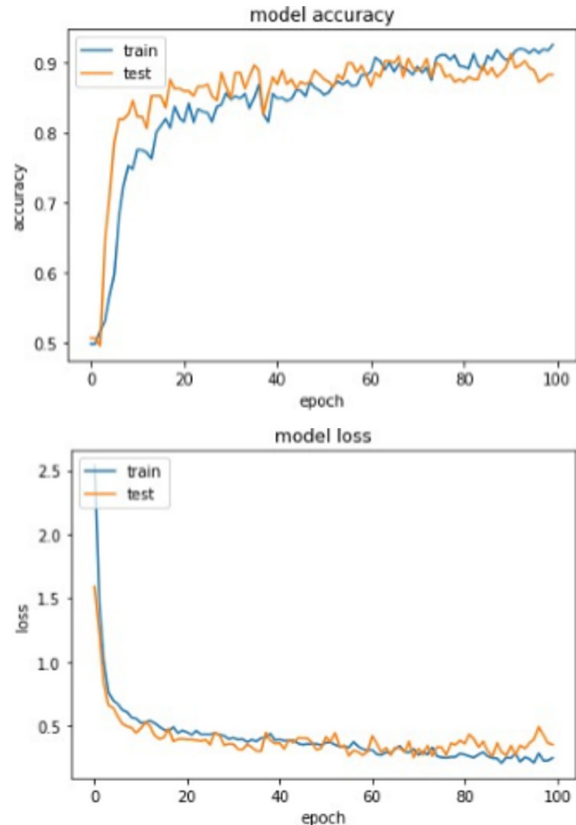


Figure 13. RNN-LSTM results for Sentiment Analysis

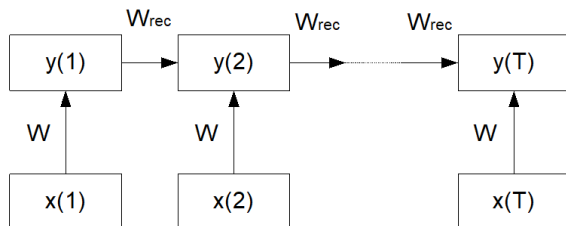


Figure 12. The basic architecture of RNN

$$\tanh(x) = 2 / (1 + e^{-2x}) - 1 \quad (4)$$

Fig. 13 shows the RNN-LSTM results for SA as accuracy and loss, respectively. Table 1 gives the RNN-LSTM model properties. Accuracy is calculated as 88%.

The architecture of RNN-LSTM model is shown in Fig. 14. It consists of three layers. Embedding layer is used as the input layer and dense layer is used as the output layer. The power of the RNN-LSTM neural network has been demonstrated using a not so deep architecture.

Table 1 gives the models properties. Adam is an efficient stochastic optimization method that only requires first-order gradients with little memory requirement [16]. As stated in reference 16, Adam gives good results in our practice compared to other stochastic optimization methods. Adam can be used to update network weights. Update expressions are shown in Equation (5) [17].

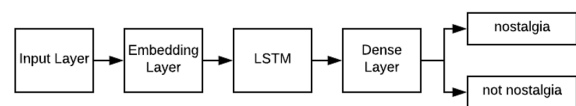


Figure 14. Architecture of the RNN-LSTM Model

$$\begin{aligned} v_t &= \beta_1 * v_{t-1} - (1 - \beta_1) * g_t \\ s_t &= \beta_2 * s_{t-1} - (1 - \beta_2) * g_t^2 \\ \Delta w_t &= -\eta \frac{v_t}{\sqrt{s_t + \epsilon}} * g_t \\ w_{t+1} &= w_t + \Delta w_t \end{aligned} \quad (5)$$

η is the initial learning rate, g_t is the gradient at time t along w^j , v_t is the exponential average of gradient along w_j , s_t is the exponential average of squares of gradients along w_j , β_1 and β_2 are hyperparameters [17].

As Loss function, MSE was used for MLPNN and CNN. MSE has been described as an excellent metric in the context of optimization, it is widely preferred as it saves time and effort [18]. The best result was obtained by using binary cross-entropy loss function in RNN-LSTM. Other features of the models are defined on the table.

Fig. 15 shows the confusion matrices of MLPNN, CNN, RNN-LSTM networks are given respectively. According to these matrices, Precision, Recall and F1 score values can be

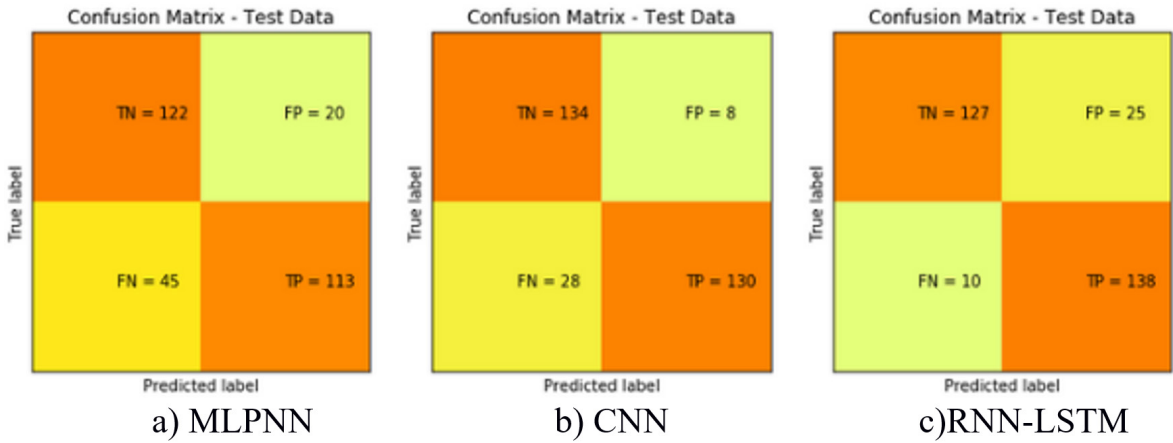


Figure 15. Confusion Matrices

obtained to better measure and evaluate the performance of the models. The Equation 6 is given below [19]. The F1 scores of the models are calculated as 0.78, 0.88, and 0.89 for MLPNN, CNN, RNN-LSTM, respectively.

$$F_1 = \left(\frac{2}{recall^{-1} + precision^{-1}} \right) = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (6)$$

Table 1. Model properties.

Model	Optimizer	Loss function	Batch size	Epoch	Accuracy (%)
MLPNN	Adam	MSE	64	200	78
CNN	Adam	MSE	32	200	88
RNN-LSTM	Adam	Binary Crossentropy	128	100	88

Fig. 16 shows the comparison of MLPNN, CNN, and RNN performance results. CNN and RNN-LSTM give better results. RNN-LSTM model completed the training for 100 epochs. It means RNN-LSTM is faster than CNN for this study.

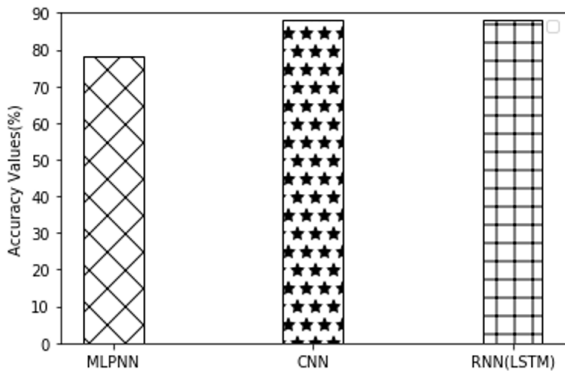


Figure 16. Comparison of MLPNN, CNN and RNN performance results

CONCLUSIONS

SA is one of the techniques used in NLP. It analyzes the text or comment and tries to find the main point. By using such techniques, planning for the future can be made by predicting the feelings and thoughts of people in all areas such as product development and marketing. In this study, comment analysis is performed with SA. Classification is done after the analysis using different NN models as deep learning which is popular today. A performance comparison is made. As a result of the performance, CNN and RNN-LSTM gives good results for the study. In future studies, analysis can be expanded by creating deeper networks and using various word embeddings. The created networks can be used easily in the analysis of application data such as Twitter, not just YouTube data.

References

1. Kumar A, Rani R. Sentiment analysis using neural network, 2nd International Conference on Next Generation Computing Technologies (NGCT), Dehradun, pp. 262-267, 2016.
2. Ayutthaya T S N, Pasupa K. Thai Sentiment Analysis via Bidirectional LSTM-CNN Model with Embedding Vectors and Sentic Features. International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP), Pattaya, Thailand, pp. 1-6, 2018.
3. Goularas D, Kamis S. Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data. International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), Istanbul, Turkey, pp. 12-17, 2019.
4. <http://www.officialcharts.com/charts/singles-chart/>, Acces date:02/01/2020

5. Bhoir S, Ghorpade T, Mane V. Comparative analysis of different word embedding models. International Conference on Advances in Computing, Communication and Control (ICAC3), Mumbai, pp. 1-4, 2017.
6. Huynh T, Le A. Integrating Grammatical Features into CNN Model for Emotion Classification. 5th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, pp. 243-249, 2018.
7. Kamran K, Meimandi J, Heidarysafa K, Mendu M, Barnes S, Brown L, Id D, Barnes L. Text Classification Algorithms: A Survey. Information (Switzerland), 10 (2019) 1-68.
8. Wan F. Sentiment Analysis of Weibo Comments Based on Deep Neural Network. International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, pp. 626-630, 2019.
9. Brownlee J. <https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/> Accessed on: 20.11.2019
10. Lu Y, Shi Y, Jia G, and Yang J. A new method for semantic consistency verification of aviation radiotelephony communication based on LSTM-RNN. IEEE International Conference on Digital Signal Processing (DSP), Beijing, pp. 422-426, 2016.
11. Timoney J, Raj A, Davis B. Nostalgic Sentiment Analysis of YouTube Comments for Chart Hits of the 20th Century. 26th AIAI Irish Conference on Artificial Intelligence and Cognitive Science, AICS, pp:1-10, 2018.
12. Xiong H, Liu Q, Song S, Cai Y. Region-based convolutional neural network using group sparse regularization for image sentiment classification. Journal on Image and Video Proc. 30 (2019), pp. 2-9.
13. Ripley B. Feed-forward Neural Networks. Pattern Recognition and Neural Networks Cambridge: Cambridge University Press, pp. 143-180, 1996. doi:10.1017/CBO9780511812651.006
14. Huang X. Convolutional Neural Networks In Convolution, https://www.researchgate.net/publication/328189519_Convolutional_Neural_Networks_In_Convolution. Accessed date:18.06.2020.
15. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), pp.2278-2324, 1998. doi:10.1109/5.726791
16. Kingma D, Ba J. Adam: A method for stochastic optimization. ICLR Conference, pp:1-15,2015.
17. Kathuria A, Intro to optimization in deep learning: Momentum, RMSProp and Adam, <https://blog.paperspace.com/intro-to-optimization-momentum-rmsprop-adam/> (Visited:01.09.2020)
18. Wang Z, Bovik A. Mean squared error: love it or leave it? – A new look at signal fidelity measures. Signal Processing Magazine, IEEE. 26 (2009) 98 – 117. 10.1109/MSP.2008.930649.
19. Powers D, Ailab. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. 2 (2011),pp. 2229-3981. 10.9735/2229-3981.