A Deep Neural Network Based Hybrid Approach for Twitter Sentiment Analysis

M.A.L Manthrirathna
Department of Computing and Information Systems
Sabaragamuwa University of Sri Lanka
Belihuloya, Sri Lanka
manthrirathna.mal@gmail.com

W.M.H.G.T.C.K. Weerakoon
Department of Physical Sciences and Technology
Sabaragamuwa University of Sri Lanka
Belihuloya, Sri Lanka
thariweera@gmail.com

R.M.K.T. Rathnayaka
Department of Physical Sciences and Technology
Sabaragamuwa University of Sri Lanka
Belihuloya, Sri Lanka
kapila.tr@gmail.com

***Abstract* — Twitter sentiment analysis results are used in multiple fields. Hybrid approaches which combine lexicons and machine learning models are proved to achieve best accuracies in sentiment analysis. Among all the machine learning techniques, Deep Neural Networks are widely used to conduct sentiment classification and are proved to perform better than other machine learning models. This research aims to compare the performance of a few Deep Neural Networks in Twitter sentiment analysis done in the hybrid approach. A set of 22,932 Tweets extracted from Twitter Standard API using the name of a popular mobile phone brand were preprocessed and classified according to SentiwordNet lexicon. In the initial step, the classified tweets were converted to numerical form and an embedding layer was created using the GloVe word embedding. Next, a simple Neural Network, Convolutional Neural Network and Long Short Term Memory Neural Network were build using the embedding layer and their performance was measured in terms of test accuracy. The results show that Long Short Term Memory Neural Network outperforms the other two models with a test accuracy of 84.94%. This result can be justified based on the nature of Twitter data. Long Short Term Memory Neural Networks are a type of Recurrent Neural Networks. RNN perform well with sequential data and Twitter data consists of sequences of words. As the conclusion, this study suggests that Long Short Term Memory Neural Networks are more suitable for Twitter sentiment analysis conducted in hybrid approach than simple Neural Networks and Convolutional Neural Networks.**

Keywords-Twitter, Sentiment Analysis, Deep Neural Networks

I. INTRODUCTION

Sentiment analysis is also known as emotion mining, sentiment classification or opinion mining. It is a sub field of Natural Language Processing research area. The purpose of sentiment analysis is to identify the emotion or the opinion with the intensity of negativity or positivity of text data according to [1],[2]. Also, sentiment analysis can identify if a text is subjective or objective [3].

Social media sentiment analysis aims to analyze social media user generated text content using sentiment analysis tools and classify data as positive, negative or neutral. This is a field of study that has been studied in many research due to its large number of applications.

With 330 million users around the world, Twitter is identified as the worlds' second mostly used social media network. Twitter users share text content related to numerous topics in the form of small text messages (tweets) creating zeta bytes of data each month. Most importantly, Twitter is used as a trusted source of user generated content by many researchers. Sentiment classification outputs of Twitter data can be utilized for so many applications. Few such applications are prediction of stock market prices [4], analysis of product recommendations [5], prediction of election results [6] and health planning [7].

The methodologies and approaches used for social media sentiment analysis can be divided under three main classes: lexicon based, machine learning based and hybrid approach. Lexicon based methods use a lexicon of classified words. Text classification is done by assigning a sentiment score to each word in analyzing text according to the lexicon. The final class is selected by averaging the sentiment score of all words. Machine learning based methods train a machine learning model to learn the patterns from training text data and then classify unseen test data using the learned patterns. Hybrid methods are a combination of machine learning and lexicon based approaches. It is proved that hybrid methodologies produce best results than the other two methods [8].

There are a large number of machine learning models commonly used in hybrid approaches for sentiment analysis. Support Vector Machine, Logistic Regression, Naïve Bayes, Decision Trees and Random Forest and Neural Networks are some of them. Among them, Neural Networks are proved to produce the best accuracies in sentiment classification tasks. According to [9] Neural Networks are good adaptive learners and are good at learning patterns, recognizing sequences and generalization.

As stated by [10],[11],[12] Deep Neural Networks are a variety of Artificial Neural Networks that have multiple hidden layers. In the natural language processing field, deep learning based methods are used for recognition of named entities [13], POS tagging and to create word-vector representation [14],[15], [16], etc. In this research, the performance of a few most commonly used Deep Neural Networks for the hybrid approach for Twitter sentiment analysis is reviewed.

II. OBJECTIVES

Deep Neural Networks play a significant role in sentiment analysis. There are a large number of Deep Neural Networks used for sentiment analysis including Convolutional Neural Networks, Recurrent Neural Networks, Long Short-term Memory Network, Long Short-Term Memory network with attention mechanism, Gated Recurrent Unit, Deep Recurrent Belief Networks, Memory Networks and Transformer Neural Networks. The objective of this research is to compare the performance of Convolutional Neural Networks, Long Short Term Memory Networks and Simple Neural Networks for Twitter sentiment classification in terms of test accuracy to comment on best suitable Deep Learning Model for Twitter sentiment analysis conducted in a hybrid approach.

III. METHODOLOGY

Twitter provides a Standard API for academic purposes. To gather data for this research, we entered the name of a popular mobile phone brand as the search keyword into the Twitter Standard API and collected 22,932 tweets. This number was reduced to 13,521 after preprocessing. It was observed that 41.04% number of Tweets were removed when preprocessing. The reason for this is the large number of tweets posted as retweets as part of marketing campaigns. During preprocessing duplicate tweets, emoji, emotions, numbers, special characters and stop words were removed. Case folding, lamentation and word segmentation were also performed on the tweets. Then the preprocessed data set was classified as positive, negative and neutral using the SentiwordNet lexicon. Results of lexicon-based classification are shown in Table 1. The data set was then divided as the training set (80% of the data) and testing set (20% of the data). Both data sets were then converted to numerical form using Keras python library. GloVe embedding was used to create an embedding layer.

To build the simple Neural Network, firstly the created embedding layer was added, and flattening was done to be able to directly add a densely connected layer. Then, a dense layer with sigmoid activation function was added to the model. Then Adam optimizer and binary cross entropy was used to compile the model. After compiling the model 80% of the training data set was used to train the model and the rest 20% was kept aside to calculate the training accuracy of the model. Finally, the test accuracy of the created model was measured using the test data set kept aside.

When building the Convolutional Neural Network, firstly the created embedding layer was added and then a one-dimensional convolutional layer with 128 kernels was added (kernel size=5). Then, feature size was reduced by adding a global max pooling layer. Finally, a dense layer with sigmoid activation was added. After that, the model was compiled and tested in the same way as in simple Neural Network.

The last model, Long Short-Term Memory Network (LSTM) was created by adding a LSTM layer with 128 neurons after adding the embedding layer. Then a global max pooling layer and a dense layer was added in the same way as when building the Convolutional Neural Network. Model compiling and testing was also done in the same way as before.

Table 2 contains the training and testing accuracies of the three models.

IV. RESULTS AND DISCUSSION

Table 1. Outcomes of lexicon based classification

Positive	Negative	Neutral
4051	2995	6475

Table 2. Performance of DNN models

	CNN	LSTM	Simple NN
Training Accuracy	90.01%	95.40%	85.52%
Test Accuracy	80.59%	84.94%	74.68%

The results show that there is a significant gap in training and testing accuracies in each model. This indicates that the models are overfitted.

Simple Neural Network produced the lowest test accuracy and Convolution Neural Network produced the second-best accuracy. Long Short-Term Memory Neural Network produced the highest test accuracy.

V. CONCLUSION

Twitter sentiment analysis outcomes are utilized in a variety of fields. Hybrid approaches are proved to produce the best results for Twitter sentiment analysis. There are a large number of machine learning models that can be used in hybrid approaches and among them all, Neural Networks are proved to produce the best accuracies.

Deep Neural Networks created by adding multiple layers to Neural Networks have an important role in sentiment analysis. The objective of this study was to conduct a comparison between the performance of simple Neural Networks, Convolutional Neural Networks and Long Short Term Memory Neural Networks for Twitter Sentiment Analysis in hybrid approach and identify which Neural Network is best suitable.

The results show that between the three models compared, Long Short-Term Memory Neural Network (LSTM) produce the best accuracy than the other two neural network models. This observation can be justified based on the nature of the data. LSTM is a type of Recurrent Neural Network (RNN). RNN is proved to produce better results with sequence data and Twitter data consist of a sequence of words.

Convolutional Neural Network (CNN) has produced an accuracy that is only slightly less than LSTM (80.59%). This proves that CNNs are also an optimal choice for Twitter sentiment analysis. CNNs are proved to work well with data that has a spatial relationship. When considering Twitter data, there is an order relationship between words that appear in the text. This explains the good performance of CNN for Twitter sentiment analysis.

In conclusion, this study suggests that Long Short Term Memory Neural Networks are better suitable for Twitter sentiment analysis than simple Neural Networks and Convolution Neural Networks.

REFERENCES

[1] V. Brock and H. U. Khan, "Big data analytics: does organizational factor matters impact technology acceptance?," *J. Big Data*, 2017, doi: 10.1186/s40537-017-0081-8.

[2] C. Bhadane, H. Dalal, and H. Doshi, "Sentiment analysis: Measuring opinions," in *Procedia Computer Science*, 2015, doi: 10.1016/j.procs.2015.03.159.

[3] W. Wolny, "Emotion analysis of twitter data that use emoticons and emoji ideograms," *25th Int. Conf. Inf. Syst. Dev. ISD 2016*, no. August 2016, pp. 476–483, 2016.

[4] K. Guo, Y. Sun, and X. Qian, "Can investor sentiment be used to predict the stock price? Dynamic analysis based on China stock market," *Phys. A Stat. Mech. its Appl.*, 2017, doi: 10.1016/j.physa.2016.11.114.

[5] Q. Dong and O. Cooper, "An orders-of-magnitude AHP supply chain risk assessment framework," *Int. J. Prod. Econ.*, 2016, doi: 10.1016/j.ijpe.2016.08.021.

[6] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with Twitter: What 140 characters reveal about political sentiment," in *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010.

[7] J. C. Eichstaedt *et al.*, "Psychological Language on Twitter Predicts County-Level Heart Disease Mortality," *Psychol. Sci.*, 2015, doi: 10.1177/0956797614557867.

[8] H. Thakkar and D. Patel, "Approaches for Sentiment Analysis on Twitter: A State-of-Art study," 2015.

[9] A. Sharma and S. Dey, "A document-level sentiment analysis approach using artificial neural network and sentiment lexicons," *ACM SIGAPP Appl. Comput. Rev.*, vol. 12, no. 4, pp. 67–75, 2012, doi: 10.1145/2432546.2432552.

[10] H. Schulz and S. Behnke, "Deep Learning: Layer-Wise Learning of Feature Hierarchies," *KI - Kunstl. Intelligenz*, vol. 26, no. 4, pp. 357–

ISSN 2756-9160 / November 2020.
International Conference on Advances in Computing and Technology (ICACT–2020) Proceedings

363, 2012, doi: 10.1007/s13218-012-0198-z.

[11] R. Pascanu, C. Gulcehre, K. Cho, and Y. Bengio, “How to construct deep recurrent neural networks,” *2nd Int. Conf. Learn. Represent. ICLR 2014 - Conf. Track Proc.*, pp. 1–13, 2014.

[12] J. Schmidhuber, “Deep Learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[13] T. Demeester, T. Rocktäschel, and S. Riedel, “Lifted rule injection for relation embeddings,” *EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 1389–1399, 2016, doi: 10.18653/v1/d16-1146.

[14] J. Pennington, R. Socher, and C. D. Manning, “GloVe: Global vectors for word representation,” *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, no. June 2018, pp. 1532–1543, 2014, doi: 10.3115/v1/d14-1162.

[15] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural language processing (almost) from scratch,” *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, 2011.

[16] T. Weerasooriya, N. Perera, and S. R. Liyanage. A method to extract essential keywords from a tweet using NLP tools. In 2016 Sixteenth International Conference on Advances in ICT for Emerging Regions (ICTer), pp. 29-34, 2016.