IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of
Information Technology & Computer Engineering

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

ASSESSMENT OF DEEP LEARNING TECHNIQUES IN SENTIMENT EVALUATION FROM TWITTER STATISTICS

Dr.Mohammad Sanaullah Qaseem[1], Mohd Khaleel Ahmed[2]
Professor[1], Asst.Prof[2]
Department of cse
NAWAB SHAH ALAM KHAN COLLEGE OF ENGINEERING & TECHNOLOGY
NEW MALAKPET, HYDERABAD-500024

Abstract: This take a look at provides a assessment of different deep gaining knowledge of methods used for sentiment evaluation in Twitter statistics. on this domain, deep learning (DL) strategies, which make a contribution on the equal time to the solution of a extensive variety of issues, won popularity amongst researchers. specifically, two classes of neural networks are utilized, convolution neural networks (CNN), which are particularly performant in the location of photograph processing and recurrent neural networks (RNN) which might be implemented with success in natural language processing (NLP) duties. on this paintings we compare and compare ensembles and combinations of CNN and a category of RNN the lengthy shortterm memory (LSTM) networks. moreover, we compare one of a kind phrase embedding systems including the Word2Vec and the worldwide vectors for phrase representation (GloVe) fashions. For the evaluation of those strategies we used information furnished by way of the global workshop on semantic assessment (SemEval), that's one of the most famous international workshops at the location. Diverse tests and combos are applied and best scoring values for every version are as compared in terms of their overall performance. This take a look at contributes to the sphere of sentiment analysis with the aid of analyzing the performances, blessings and barriers of the above methods with an evaluation method underneath a unmarried testing framework with the identical dataset and computing surroundings.

key phrases: sentiment evaluation, deep gaining knowledge of, convolution neural networks, LSTM, word embedding models, Twitter statistics.

INTRODUCTION:

In latest years, thanks to the boom within the use of social media, sentiment evaluation gained recognition among a wide range of human beings with different hobbies and motivations. As customers everywhere in the international have the possibility to specific their opinion approximately unique topics associated with politics, schooling, travel, subculture, commercial merchandise, or topics of well known interest, extracting knowledge from those records have become a topic of excellent significance and significance. Besides facts concerning users' visited sites, buying choices etc., knowing their emotions as they're expressed by way of their messages in diverse systems, turned out to be an essential detail for the estimation of human being's opinion about a specific problem. a very common technique is to categorise the polarity of a text in phrases of user's pride, dissatisfaction or neutrality. The polarity can vary in terms of labeling or wide variety of tiers from effective to poor but in widespread it denotes the feelings of a textual content varying from a glad to an unhappy mode. The tactics used for sentiments analysis are numerous and are primarily based on one-of-a-kind strategies of herbal language processing and system learning strategies for extracting ok functions and classifying text in suitable polarity labels. in view that some years, with the popularity that deep learning techniques have received, various deep neural networks were applied on the field with achievement. Particularly, the convolution neural networks and LSTM networks proved to be performant for sentiment analysis obligations. Various researches showed their effectiveness alone or in mixture among them. in the subject of natural language processing, most of the strategies which might be used for extracting features from words, Word2Vec and the global vectors for phrase representation (GloVe) are the most popular ones. The accuracy completed with the above strategies is high however still no longer excellent, for this reason making sentiment analysis an ongoing and open research issue. because of this researchers try and broaden new methods or enhance the present ones. As the present techniques have a massive variety in terms of network configuration, tuning, and

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

many others., a studies upon the evaluation of the already used techniques remains essential that allows you to have a clear an specific idea about their limits and the challenges on sentiment evaluation. This paper contributes to this area through comparing the maximum popular deep mastering techniques and configurations based totally on an accepted dataset approximately sentiment evaluation that's constructed from Twitter information below a single testing framework. The paper is split as follows:

Phase 2 provides the associated work in this discipline. Section 3 demonstrates the method and the extraordinary neural network configurations which can be applied. Segment four suggests the effects, compares the extraordinary strategies among them and discusses the findings. in the end, section five concludes the paper. II. Background With the enlargement and recognition of social media and numerous systems allowing humans to explicit their opinion upon different topics, sentiment evaluation and opinion mining became a topic that attracted the eye of researchers global. In a work published in 2008, the authors described the various methods that have been used until that day. Within the ultimate years deep neural networks proved to be specially performant in sentiment evaluation tasks. Among them, convolution neural networks and recurrent neural networks were broadly implemented because CNN respond thoroughly to the dimensionality discount hassle and a category of RNN the LSTM networks take care of with achievement temporal or sequential records. Within the pioneer works provided within the authors demonstrated that CNN architectures may be utilized with achievement for sentence class. furthermore it was tested that CNN perform slightly higher than traditional strategies the performance of RNN turned into proven as they outperformed the state of the art strategies and in an implementation of CNN and LSTM networks was offered, displaying the big blessings of the use of together these neural networks. In parallel, the GRU networks, added in 2014 can be used efficiently with similar results within the place of LSTM. In a survey of deep learning strategies in sentiment analysis , it may be seen that the phrase embedding is completed specially with two techniques, Word2Vec or GloVe .these days, Twitter is one of the maximum influencing social media systems which serves as an records sharing medium in nations everywhere in the global. Therefore, extracting public opinion from tweets approximately various topics, measuring the influence of different activities or classifying sentiments have become a subject of great hobby. The early works for sentiment analysis were the use of one of a kind strategies for extracting capabilities based totally specially on bi-grams, unigrams, POS unique polarity functions and have been utilizing device getting to know classifiers like the Bayesian networks or help vector machines. in the remaining years, in specific places all over the international, numerous technology competitions were prepared with the intention to appeal to the hobby of researchers. Amongst them, the international workshop on semantic assessment is organizing competitions on this field for the last thirteen years. Today, deep getting to know strategies are dominant and the associated research try to advantage high ratings in the competition the use of particularly extraordinary combinations of neural networks and various configurations of word embedding functions. Regarding the sentiment analysis in Twitter information, some of studies had been distinguished in terms of overall performance. Within the authors proposed two different CNN configurations the usage of one of a kind word embeddings, Word2Vec and GloVe, respectively, where their effects are combined in a random woodland classifier. In some other look at the authors are using embeddings trained on lexical, element-of-speech and sentiment embeddings which might be initializing the enter of a deep CNN structure. in the authors proposed two configurations primarily based on bidirectional LSTM networks. The word embedding is accomplished with GloVe. Some other have a look at proposed a mixture of CNN and LSTM networks. The authors experimented with three special phrase embedding models, the Word2Vec, GloVe and FastText and they reported that GloVe had a negative performance as compared to the opposite models. Sooner or later, a variant of CNN's the RCNN's were used efficiently in . Regardless of the performances of the above research, when looking to do a comparison between them it became especially hard to assess the position of a dataset, a network configuration or a selected setup and tuning. the inducement of this study got here from this trouble, aiming to create a unmarried framework in an effort to compare these strategies and clarify the advantages and barriers of each particular configuration. technique. in this segment we present the dataset, the phrase embedding models with their configurations, and the one of a kind deep neural network configurations which are used in this have a look at. Inside the following setups GRU networks and RCNN's aren't covered due to the fact they give similar outcomes with LSTM networks and CNN's. A. Dataset and Preprossesing A corpus of various datasets become applied based totally on three datasets utilized in SemEval competitions. greater specially, the SemEval2014 Task9-SubTask B full statistics, the SemEval2016 complete information Task4 and the SemEval2017 improvement facts had been used forming a complete of round 32.000 tweets. They

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

encompass a body of 662.000 words with a vocabulary of around 10.000 words the next step changed into to method the tweets in an effort to boom the device's overall performance in the course of schooling. because of this, an additional preprocessing task turned into finished aiming to eliminate and modify a few characters. This task included the conversion of all letters to lowercase, the removal of a few unique characters and emoticons or the tagging of urls.

B. word Embedding the phrase embedding models used in this look at were the Word2Vec, and Glove. The Word2Vec model changed into applied to create 25-dimensional word vectors based at the dataset defined before. The configuration of Word2Vec became finished by way of using the CBOW model. Additionally, words that seemed less than five times have been discarded. Eventually, the most pass length among phrases changed into set to 5. GloVe changed into applied with its pertained phrase vectors. they're also 25-dimensional vectors and had been constructed from 2 billion tweets, which constitutes a notably larger schooling dataset than the dataset extracted from SemVal information. All vectors have been normalized using the following equation

$$v_i' = \frac{v_i - v_{min}}{v_{max} - v_{min}} \qquad (1)$$

Where is the normalized i-value of a 25-dimensional vector minimum value and the maximum value of the vector? 1) Sentence vectors the sentence vectors are created after concatenating the word vectors of a tweet in order to form a unique vector. After experiencing with various lengths we created sentences with a length of 40 words. As tweets vary in length, in case that a tweet has more words, the extra words were removed. When they were less than 40 the words of the tweet were repeated until the desired size was achieved. An alternative method is to use zero padding in order to fill the missing words in a sentence. In the approach followed in this work, zero padding was used only in case of words that were not present in the vocabulary.

2) Sentence Regions A supplementary approach in word embedding is to divide the word vectors of a sentence in regions, in an effort to preserve information in a sentence and long-distance dependency across sentences during the prediction process . The division is done with the punctuation marks existing on a sentence. In the current configuration each region is composed of 10 words and a sentence has eight regions. In case of missing words or regions, zero padding is applied. Figure 1 presents the structure of regions in a sentence.

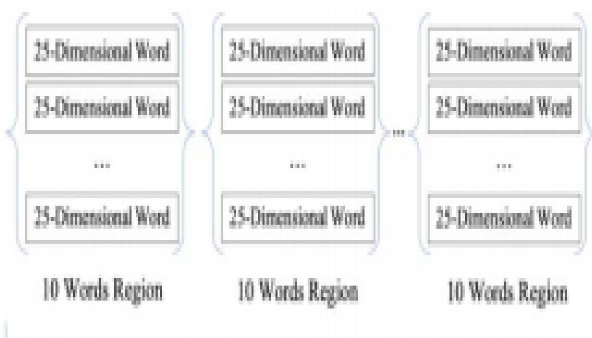


Fig. 1. Regional structure of a sentence. Every sentence has eight regions and every region has 10 25-dimensional words. In case of missing words or regions zero padding is applied in order to fill the missing regions.

In the end the dataset is converted two times forming two distinct datasets, one with non-regional and another with regional based sentences. In the first case the input size is 1000 (a sentence has 40 words where each of them has a size of 25) and in the second is 2000 (a sentence is divided into eight regions where each of them has 10 words of size 25) C. Neural Networks The neural network configurations that are proposed for the evaluation of twitter data

International Journal of
Information Technology & Computer Engineering

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

are based on CNN and LSTM networks. Additionally, in one case a SVM classifier is used. All the networks were tested with both non-regional and regional datasets. In total, eight network configurations are proposed. As mentioned above, RCNN and GRU networks are not utilized because in our experiments they had very similar performance with CNN and LSTM networks correspondingly. All networks were trained with 300 epochs and used sigmoid activation function. 1) Single CNN network In this network a single 1-dimensional CNN layer is used. Figure 2 presents this configuration where the sentence vector is convolved with 12 kernels with size 1×3 (from our tests it performed better when compared with other kernel configurations). The max pooling layer has a size of 1×3 . The CNN parameters will be the same for the following CNN configurations. Finally, a 3-dimensional output predicts the polarity in terms of positive, negative or neutral answer.

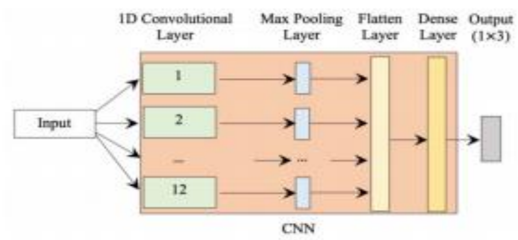


Fig. 2. CNN configuration with one layer and a 3-dimensional output for positive, neutral and negative polarity prediction.

2) Single LSTM network In this configuration a single LSTM layer is used with a dropout of 20%. The output is again 1×3 in order to predict the polarity (positive, neutral or negative). 3) Individual CNN and LSTM networks the aim of this configuration is to take the outputs of individual CNN and LSTM networks and evaluate together their results. A soft voting based on the outputs of the networks decides about the prediction answer. Figure 3 shows the structure of this configuration where the CNN and the LSTM networks have the same settings as in the two previous configurations (for CNN 12 kernels with size 1×3 and a max pooling layer with a size of 1×3).

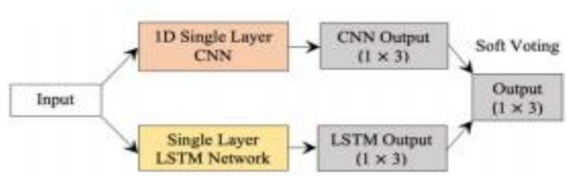
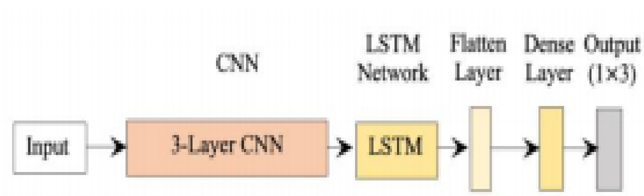


Fig. 3. Individual CNN and LSTM networks. The final prediction answer is given after soft voting calculated from the network outputs.

4) Single 3-Layer CNN and LSTM Networks This setup utilizes a 3-layer 1-dimensional CNN and a single layer LSTM network. Figure 4 displays this configuration where the input is directed to a 3-layer CNN. The input has a size of 1000 if it is based on words (nonregional) or a size of 2000 if it is based on regions (regional).



ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

Fig. 4. Combination of a 3-Layer CNN and a LSTM network

5) Multiple CNN's and LSTM Networks In the current setup the input is divided into its basic elements, words for non-regional inputs and regions for regional inputs. Those elements serves as an input to individual CNN's. Then the output of every CNN is directed as an input to a single LSTM network. Figure 5 presents the network structure. In total according to the type of the input we have 40 or eight corresponding CNN's (40 words or eight regions). Every CNN network utilizes as previously 12 kernels.

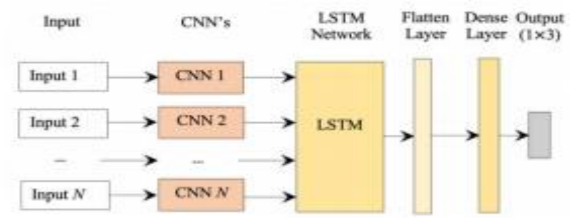


Fig. 5. Combination of CNN's and LSTM networks for an input that is divided into N inputs. N is equal to 40 (words) if the input is non-regional or 8 (regions) if the input is regional.

6) Single 3-Layer CNN and bidirectional LSTM Network This setup includes a configuration same as (5) with the difference that this time a bidirectional LSTM network is used. The aim of this setup is to test the effectiveness of bidirectional LSTM networks compared to simple LSTM networks. 7) Multiple CNN's and bidirectional LSTM Network Again this setup includes a configuration identical to (6) with the difference that this time a bidirectional LSTM network is used.

IV. RESULTS

This section presents the performance results of the previous network configurations in terms of Accuracy, Precision Recall, and F-measure (F1) as described in the following equations:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F\_Measure = \frac{2*Recall*Precision}{Recall+PRecision} \quad (5)$$

In the above equations are the true positive, are the true negative, false positive and false negative predictions. Table I and Table II presents the performance results of the proposed combinations using CNN and LSTM networks with Wor2Vec and GloVE word embedding systems correspondingly. First, we can observe that utilizing the GloVe system increased the performance of almost all configurations (5%-7%). The reason behind it lies to the fact that with Word2Vec the vectorization of words has been made with a relatively small training dataset, around 32.000 tweets compared to the pretrained word vectors made with GloVe that used a significantly larger training dataset. The second observation is that using multiple CNN with LSTM networks instead of simple configurations increases the performance of the system, independently of the word embedding system (3%-6%). We can observe that the configurations have almost always the best performance when compared with the other configurations. A third observation is that separating the text input into regions in most cases doesn't really improve the performance of a configuration (1%-2%). Concerning the use of SVM classifier instead of a soft-voting procedure it can be seen that

International Journal of
Information Technology & Computer Engineering

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

it gives a slightly worse performance. A last observation is that the use of bidirectional LSTM networks instead of simple LSTM networks doesn't present any real advantage, which can be eventually explained by the nature of the data (the structure of words in a sentence). Table III compares the best results of this study with the results of other works that used similar neural networks. We can observe that the current study has similar but slight inferior performance with the literature studies (6% difference). This is expected and it is due to the different datasets and specialized methods that are used in the other studies for shaping the dataset or tuning the network. Moreover, the scope of this study was not focused on achieving the best performance in comparison with other studies, but rather to evaluate and compare different deep neural networks and word embedding systems on a single framework. At this point, it is worth to mention that the best performance in the literature in terms of accuracy (~65%) it is still not satisfactory, thus revealing that on sentiment analysis deep learning methods are still far from guaranteeing a performance comparable to other fields where the same networks are used with higher success rate (e.g. deep learning networks for object recognition in images).

TABLE I. Sentiment prediction of different combinations of CNN and LSTM networks with Word2Vec word embedding system with no-regional and regional settings from a set of around 32.000 tweets

Embedding Word System: Word2Vec						
Network Model	**Type**	**Recall**	**Prec.**	**F1**	**Acc.**
1. Single CNN network	N-R[a]	0.33	0.35	0.33	0.49
	R	0.32	0.34	0.33	0.51
2. Single LSTM network	N-R	0.43	0.51	0.39	0.51
	R	0.44	0.49	0.39	0.50
3. Individual CNN and LSTM Networks	N-R	0.43	0.47	0.37	0.50
	R	0.46	0.52	0.42	0.52
4. Individual CNN and LSTM Networks with SVM classifier	N-R	0.45	0.46	0.43	0.49
	R	0.42	0.54	0.38	0.51
5. Single 3-Layer CNN and LSTM Networks	N-R	0.41	0.52	0.40	0.46
	R	0.40	0.46	0.35	0.48
6. Multiple CNN's and LSTM Networks	N-R	0.43	0.47	0.37	0.50
	R	0.46	0.52	0.43	0.52
7. Single 3-Layer CNN and bi-LSTM Networks	N-R	0.42	0.45	0.39	0.48
	R	0.42	0.47	0.36	0.48
8. Multiple CNN's and bi-LSTM Networks	N-R	0.43	0.50	0.38	0.51
	R	0.46	0.51	0.44	**0.52**

TABLE II. Sentiment prediction of different combinations of CNN and LSTM networks with GloVe word embedding system with no-regional and regional settings from a set of around 32.000 tweets

International Journal of
Information Technology & Computer Engineering

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

Network Model	Type	Recall	Prec.	F1	Acc.
1. Single CNN network	N-R[a]	0.44	0.41	0.4	0.54
	R	0.35	0.31	0.31	0.48
2. Single LSTM network	N-R	0.5	0.58	0.48	0.55
	R	0.51	0.55	0.51	0.55
3. Individual CNN and LSTM Networks	N-R	0.53	0.6	0.53	0.58
	R	0.55	0.6	0.55	0.56
4. Individual CNN and LSTM Networks with SVM classifier	N-R	0.52	0.55	0.53	0.56
	R	0.49	0.6	0.5	0.56
5. Single 3-Layer CNN and LSTM Networks	N-R	0.5	0.5	0.5	0.52
	R	0.43	0.61	0.39	0.53
6. Multiple CNN's and LSTM Network	N-R	0.53	0.60	0.53	0.58
	R	0.55	0.6	0.56	**0.59**
7. Single 3-Layer CNN and bi-LSTM Network	N-R	0.52	0.59	0.53	0.57
	R	0.50	0.57	0.50	0.55
8. Multiple CNN's and bi-LSTM Network	N-R	0.54	0.60	0.55	**0.59**
	R	0.55	0.6	0.56	**0.59**

TABLE III. Comparison of the state of the art methods with the best results of the current study.

Study	Network System	Word Embedding	Dataset (labeled Tweets)	Accuracy
Baziotis et al. [22]	bi-LSTM	GloVe	~50.000	**0.65**
Cliche [23]	CNN+LSTM	GloVe FastText Word2Vec	~50.000	**0.65**
Deriu et al. [20]	CNN	GloVe Word2Vec	~300.000	**0.65**
Rouvier and Favre [21]	CNN	Lexical, POS, Sentiment	~20.000	0.61
Wange et al. [27]	CNN+LSTM	Regional Word2Vec	~8.500	1.341[a]
Current study	CNN+LSTM	Regional, GloVe	~31.000	0.59

V. CONCLUSION:

In this paper different configurations of deep learning methods based on CNN and LSTM networks are tested for sentiment analysis in Twitter data. This evaluation gave slight inferior but similar results with the state of the art methods, thus allowing to extract credible conclusions about the different setups. The relatively low performance of these systems showed the limitations of CNN and LSTM networks on the field. Concerning their configuration, it was observed that when CNN and LSTM networks are combined together they perform better than when used alone. This is due to the effective dimensionality reduction process of CNN's and the preservation of word dependencies when using LSTM networks. Moreover, using multiple CNN and LSTM networks increases the performance of the system. The difference in accuracy performance between different datasets demonstrates that, as expected, having an appropriate dataset is the key element for increasing the performance of such systems. Consequently, it looks like spending more time and effort in order to create good training sets presents more advantages rather than experimenting with different combinations or settings for CNN and LSTM networks configurations. To summarize, the contribution of this paper is that it allowed to evaluate different deep neural network configurations and experimented with two different word embedding systems under a single dataset and evaluation framework allowing to shed more light on their advantages and limitations.

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

REFERENCES

[1] L. L. Bo Pang, «Opinion Mining and Sentiment Analysis Bo», Found. Trends® Inf. Retr., vol. 1, no. 2, pp. 91–231, 2008.

[2] K. Fukushima, «Neocognition: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position», Biol. Cybern., vol. 202, no. 36, pp. 193–202, 1980.

[3] Y. Lecun, P. Ha, L. Bottou, Y. Bengio, S. Drive, eta R. B. Nj, «Object Recognition with Gradient-Based Learning», in Shape, contour and grouping in computer vision, Heidelberg: Springer, 1999, pp. 319–345.

[4] D. E. Ruineihart, G. E. Hinton, eta R. J. Williams, «Learning internal representations by error propagation», 1985.

[5] S. Hochreiter eta J. Urgen Schmidhuber, «Ltsm», Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[6] Y. Kim, «Convolutional Neural Networks for Sentence Classification», arXiv Prepr. arXiv1408.5882., 2014.

[7] C. N. dos Santos eta 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.

[8] P. Ray eta A. Chakrabarti, «A Mixed approach of Deep Learning method and Rule-Based method to improve Aspect Level Sentiment Analysis», Appl. Comput. Informatics, 2019.

[9] S. Lai, L. Xu, K. Liu, eta J. Zhao, «Recurrent Convolutional Neural Networks for Text Classification», Twenty-ninth AAAI Conf. Artif. Intell., pp. 2267–2273, 2015.

[10]D. Tang, B. Qin, eta T. Liu, «Document Modeling with Gated Recurrent Neural Network for Sentiment Classification», in Proceedings of the 2015 conference on empirical methods in natural language processing, 2015, pp. 1422–1432.

[11]K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, eta Y. Bengio, «Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation», arXiv Prepr. arXiv1406.1078, 2014.

[12]J. Chung, C. Gulcehre, K. Cho, eta Y. Bengio, «Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling», arXiv Prepr. arXiv1412.3555, 2014.

[13]L. Zhang, S. Wang, eta B. Liu, «Deep Learning for Sentiment Analysis: A Survey», Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 8, no. 4, pp. e1253, 2018.

[14]T. Mikolov, K. Chen, G. Corrado, eta J. Dean, «Efficient Estimation of Word Representations in Vector Space», arXiv Prepr. arXiv1301.3781, 2013.

[15]J. Pennington, R. Socher, eta C. Manning, «Glove: Global Vectors for Word Representation», Proc. 2014 Conf. Empir. Methods Nat. Lang. Process., pp. 1532–1543, 2014. [16]H. Kwak, C. Lee, H. Park, eta S. Moon, , in Proceedings of the 19th international conference on World wide web, 2010, pp. 591–600.

[17]A. Go, R. Bhayani, eta L. Huang, «Proceedings - Twitter Sentiment Classification using Distant Supervision (2009).pdf», CS224N Proj. Report, Stanford, vol. 1, no. 12, 2009.

[18]A. Srivastava, V. Singh, eta G. S. Drall, «Sentiment Analysis of Twitter Data», in Proceedings of the Workshop on Language in Social Media, 2011, pp. 30–38.

International Journal of
Information Technology & Computer Engineering

ISSN NO: 2347 – 3657
Volume 8, Issue 4, 2020

[19]E. Kouloumpis, T. Wilson, eta M. Johanna, «Twitter Sentiment Analysis: The Good the Bad and the OMG!», in Fifth International AAAI conference on weblogs and social media, 2011. [20]J. Deriu, M. Gonzenbach, F. Uzdilli, A. Lucchi, V. De Luca, eta M. Jaggi, «SwissCheese at SemEval-2016 Task 4: Sentiment Classification Using an Ensemble of Convolutional Neural Networks with Distant Supervision», Proc. 10th Int. Work. Semant. Eval., pp. 1124–1128, 2016. [21]M. Rouvier eta B. Favre, «SENSEI-LIF at SemEval-2016 Task 4 : Polarity embedding fusion for robust sentiment analysis», Proc. 10th Int. Work. Semant. Eval., pp. 207–213, 2016. [22]C. Baziotis, N. Pelekis, eta C. Doulkeridis, «DataStories at SemEval2017 Task 4: Deep LSTM with Attention for Message-level and Topicbased Sentiment Analysis», Proc. 11th Int. Work. Semant. Eval., pp. 747–754, 2017.

[23]M. Cliche, «BB twtr at SemEval-2017 Task 4: Twitter Sentiment Analysis with CNNs and LSTMs», arXiv Prepr. arXiv1704.06125.

[24]P. Bojanowski, E. Grave, A. Joulin, eta T. Mikolov, «Enriching Word Vectors with Subword Information», Trans. Assoc. Comput. Linguist., Vol 5, pp. 135–146, 2016.

[25]T. Lei, H. Joshi, R. Barzilay, T. Jaakkola, K. Tymoshenko, A. Moschitti, eta L. Marquez, «Semi-

supervised Question Retrieval with Gated Convolutions», arXiv Perry. arXiv1512.05726, 2015.

[26]Y. Yin, S. Yangqiu, eta M. Zhang, «NNEMBs at SemEval-2017 Task 4: Neural Twitter Sentiment Classification: a Simple Ensemble Method with Different Embeddings», Proc. 11th Int. Work. Semant. Eval., pp