M82B at CheckThat! 2021: Multiclass fake news detection using BiLSTM

Sohel Siddique Ashik [1], Abdur Rahman Apu [1], Nusrat Jahan Marjana[1], Md. Sanzidul Islam [1] and Md. Arid Hasan[1]

[1] *Daffodil International University, Dhaka, Bangladesh*

Abstract

The rapid advancement of web technologies enabled to spread of information online faster. This enabled the spreading of both informative and uninformative information among the people. Fake news represents misinformation which is published in the newspapers, blogs, newsfeed, social media. Fake news is expanding very quickly via social media, generated by humans or machines as well as originating the unrest situation in the society, country. Meanwhile, fake news detection using machine learning is becoming a prominent area in the research to identify the credibility of the news instantaneously. To present this work, we used Bidirectional Long Short-Term Memory (BiLSTM) to predict the news is either fake or true. We participated in the fake news classification shared task of CheckThat! 2021 workshop. We obtained the dataset from this event to train and evaluate our model. Finally, we were able to achieve 36% model accuracy and 29.0 F1-macro score with our training data.

Keywords

BiLSTM, Deep Learning, Fake news, Multiclass Classification.

1. Introduction

In the modern era, the uses of social media as news collecting, sharing are not only increasing but also generating various news, contents on the internet every day. News helps people to know, gain knowledge about the current situation of the world, country. There are various media available for collecting news, but social media as the source of collecting news is becoming a very popular source than traditional media like television, radio, newspaper etc. for various reasons. The reasons are:

- Its easiest accessibility,
- It requires lower cost as well as less time to access news,
- It provides sharing with friend's, comment features on social media [1].

Moreover, most of the news are circulating on the social media is fake news [2]. Social media like Facebook, Twitter, Instagram opens many facilities with some challenges such as these platforms are the source of spreading fake news and lots of social bots which helps to spread fake news [3] [4]. Fake news is disinformation or rumor presented to the people as true news for influencing people's opinion and thrust them into political action and can often be a profitable business for online publishers. Fake news can hoax people by looking like faithful websites or using duplicate names, web addresses to honorable news organizations. These websites can be found in America, China, Russia and other countries [5]. Human is responsible for creating fake news for self-interest, creating inconstancy in the country, especially in the political area.

According to the report in 2017, social bots are available in the Twitter is 23 million, in the Facebook is 140 million and in the Instagram is 27 million [4]. Fake news worked as a factor to expand political polarization. The issue of fake news has gained prominence in the USA presidential election 2016 [2].

[1]CLEF 2021 – Conference and Labs of the Evaluation Forum, September 21–24, 2021, Bucharest, Romania

EMAIL: sohel15-12898@diu.edu.bd (A. 1); abdur15-13025@diu.edu.bd (A. 2); nusrat15-12999@diu.edu.bd (A. 3); sanzid.cse@diu.edu.bd (A. 4); arid.cse0325.c@diu.edu.bd (A. 5);
ORCID: 0000-0001-6901-2737 (A. 1); 0000-0003-2415-4357 (A. 2); 0000-0001-7068-301X (A. 3); 0000-0002-9563-2419 (A. 4); 0000-0001-7916-614X (A. 5);

Social media is becoming the home of sharing fake news [3]. Almost 90% of adults use the internet and approximately 16% use Twitter in the USA [17]. Dwoskin [18] stated that fake news had an important role to attract and influence people's thoughts, view in the USA presidential election 2016. Thus, it became a great threat for the world. It became necessary to create tools to identify the fake news.

This paper introduced a Recurrent Neural Network (RNN) model – BiLSTM model for detecting fake news. The dataset for training and evaluating our model is achieved from CheckThat! 2021 workshop [19-21] where we participated. The aim of this research is to experiment on how perfectly the BiLSTM model works for detecting fake news.

The remaining parts of this paper are arranged as follows. In the second part, we will present previous work on fake news detection. We will present our proposed approach and model, in section 3. In section 4, we will show our experimental result. We will conclude our research with future research work direction, in section 5.

2. Related work

Fake news is becoming a very popular area in research. In this section, we will discuss some related work which is already done on fake news detection. In this paper [6], Zhang et al. proposed FakeDetector Which is an automatic fake news credibility inference model and a new diffusive unit model, namely GDU.They achieved an excellent Bi-Class Inference Result and Multi-Class Inference Result. The accuracy of Bi-Class Inference Result is 0.63 and Multi-Class Inference Result is 0.28. Recently, CNN models have been applied in identifying Fake news widely. Umer et al. used [7] neural network and CNN-LSTM model Where PCA and chi-square extract quality features and the accuracy score are 97.8%. The Authors in [8] have described a fake news detection model that uses n-gram analysis and machine learning techniques. In this paper they studied and compare two different features selection methods which are Term Frequency (TF) and Term Frequency-Inverse Document Frequency (TF-IDF) And six different machine learning algorithms (classification techniques) namely Stochastic Gradient Descent (SGD), Support Vector Machines (SVM), Linear Support Vector Machines (LSVM), K-Nearest Neighbor (KNN) and Decision Trees (DT). By using (TF-IDF) and LSVM, the accuracy is 92% which is the best accuracy above them. Granik et al. proposed [9] a simple way or approach for detecting fake news using Naïve Bayes Classifier and achieved classification accuracy 74% on a data set of Facebook news posts. Mainly this paper shows that simple AI Algorithms also provide such good accuracy.The Authors in [10] presented a hybrid model namely Bidirectional LSTM which is based on POS (part of speech) and Convolutional Neural Network (CNN).The accuracy score of this paper is 42.2% on Liar-Liar dataset.

The aim of this paper is to analyze the most well-known super-wise strategy for detecting fake news.The Authors in [11] want to find out the best classifier model using three feature extraction techniques such as Term Frequency–Inverse Document Frequency (TF–IDF), Count-Vectorizer (CV) and Hashing-Vectorizer (HV).They also have described a novel multi-level voting fake news detection model that uses the best three machine learning(ML) models are combined from each feature extraction technique. This paper uses 12 classifiers using three dataset, classifiers are Multinomial Naïve Bayes (MultinomialNB), Passive Aggressive (PA), Stochastic Gradient Descent (SGD), Logistic Regression (LR), Support Vector classifier (SVC), NuSVC, LinearSVC, Multi-Layer Perceptron (MLP), Decision Tree (DT), AdaBoost, Gradient Boosting and Voting Classifier etc were analyzed on the basis of their performance measures.The Accuracy of Passive Aggressive model is 0.8%, Logistic Regression model is 1.3%, LinearSVC model is 0.4% using TF-IDF, CV and HV.The Main purpose of this paper [12] is to show how to use semantic features to enhance fake news detection and Improves accuracy by adding semantic features significantly. In this paper, they described a semantic fake news detection method which is extracted directly from a text and formed around relational features like sentiment, entities or facts.

Recently deep learning techniques have been applied in identifying fake news widely. Karimiha et al. Proposed a Multi-source Multi-class Fake news Detection framework (MMFD). The accuracy score of this paper [13] is 38.81 on the LIAR dataset. The Authors in [14] presented a Dense neural network (DNN) model which accurately predicts the position between a given pair of headline and article body. On the test dataset the achieved accuracy is 94.21% and outperform existing model architectures by 2.5%. Singhal et al. introduced SpotFake- a multi-modal framework for fake news detection which exploits the textual and visual features of an article.To learn text features they use models like BERT which is a language mode and image features are learned from VGG-19 on ImageNet dataset [15].

3. Methodology

3.1. Dataset

The dataset of fake news is obtained from CLEF Checkthat! 2021 workshop. Train dataset contains 900 records and the test set consists of 364 records. Both datasets have four attributes – public_id, text, title and rating which are text type data. Text column defines full news, title is the headline of that news and rating column states the identity (true, false, partially false, other) of that news. The statistics of the target attribute is shown in table 3.1.

Table 3.1: The overview of the training and test data set.

Classes	Train size	Test size
False	465	113
True	142	69
Other	76	41
Partially False	217	141

3.2. Data-preprocessing

Text type data cannot be used to implement machine learning or deep learning models. It can work with numerical type data. We applied data preprocessing techniques on text and rating dimension.

Dimension reduced: Unnecessary attributes are responsible for expanding the execution of time. Before inserting data in the model, public_id, title attributes are removed from the dataset. Text column is input data and the target column is rating.

Noise removal: All data has to be clear and noiseless. Otherwise, unexpected trouble might be happened. Words have to be eliminated from the input sequence which are useless to tokenize and vectorize. The uppercase characters are converted into the lowercase characters in order to normalize the tokenization process.

Punctuation removal: Punctuation is important in our natural language conversion to add clarity of the sentence. Punctuation like comma, does not provide any special value to understand the meaning of the sentence for data analysis. The punctuation removing process is shown in the figure 3.2.1.

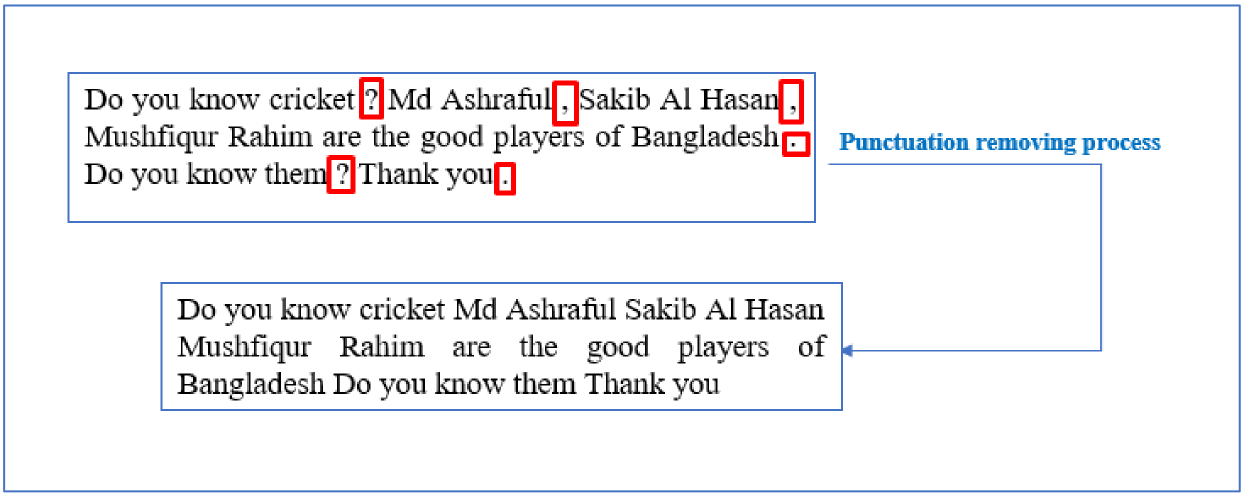

Figure 3.2.1: Punctuation removing example

Stopwards removal: Stopwards are less important in a sentence because they are common in natural language and do not provide any special meaning. In data analysis, stopwards can take more execution time. So, it is necessary to filter out stopwards from the sentence. We used the NLTK library for removing stopwords from the sentence. The example of stopwords removal is shown in the figure 3.2.2.

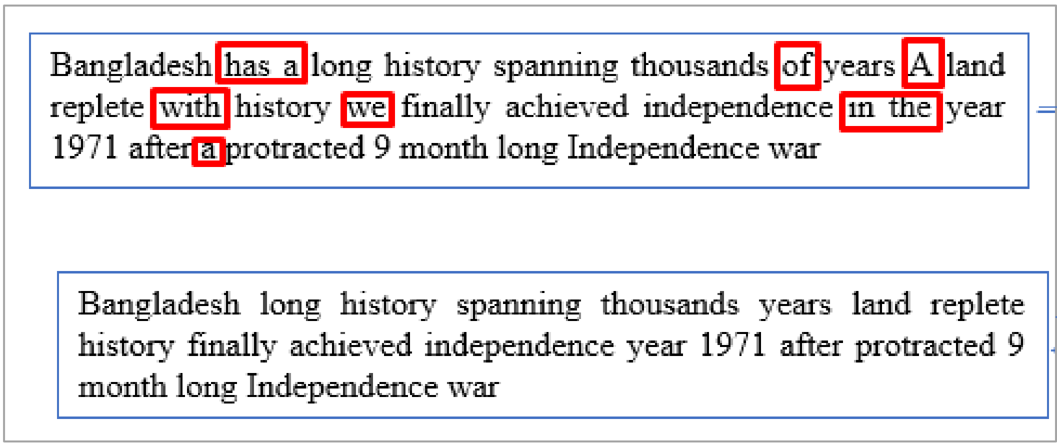

Figure 3.2.2: An example of removing stopwards

Stemming removal: Stemming is the process of removing suffixes and prefixes form a word. By applying the stemming technique, we can convert a word to its common base structure.

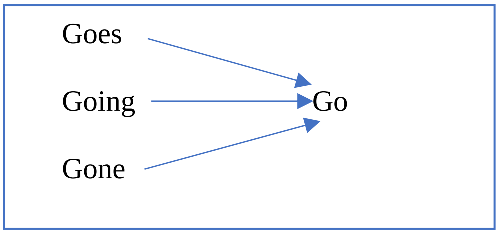

Figure 3.2.3: An example of stemming removal

Encoding technique: Encoding technique is an important data preprocessing technique which is used to convert text type label into numerical form for machine friendly because text is not understand able by machine. We used the Label Encoding technique into the label column which is shown in the table 3.2.4.

Table 3.2.4: Label Encoding technique on rating column

Before	After
False	0
True	1
Other	2
Partially False	3

Word tokenization technique: Word tokenization is a required technique for natural language processing tasks which is used to split words from a sentence for classifying, counting words and other analysis. We used NLTK library for word tokenization.

3.3. Proposed methodology

This paper introduced a BiLSTM model for analysis whether the news is fake or authentic. The process of predicting whether the news is true, false, partially false or other type by using this model is shown in figure-3.3.1.

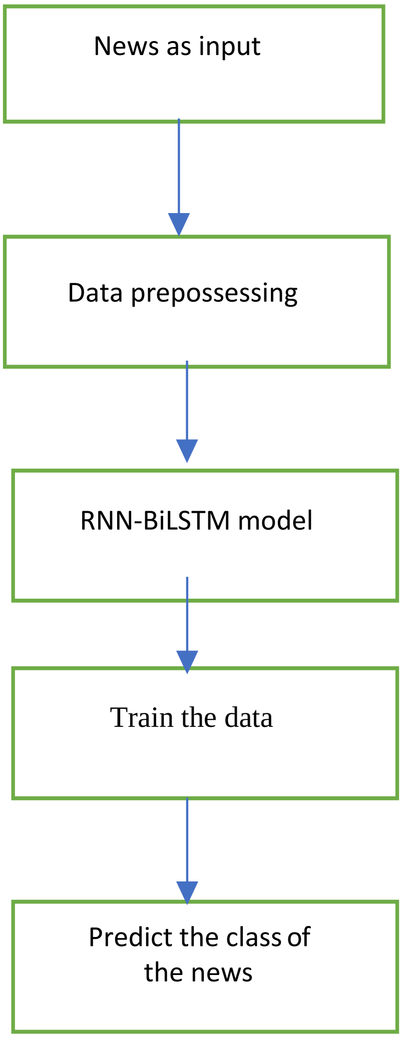

Figure 3.3.1: steps of predicting the class of the news

3.4. BiLSTM

Bi-directional long short-term memory (BiLSTM) performs better on sequence classification problems than Long Short-Term Memory (LSTM). BiLSTM comprises two LSTMS-one LSTM takes input in the forward direction and another takes in the backward direction for training [12]. By using hidden states, it can flow information in both directions. At each time step, the output of two LSTMs is joined. This BiLSTM model helps to remove barriers of conventional RNN. BiLSTM provides good accuracy so that the context can understand in a very better way.

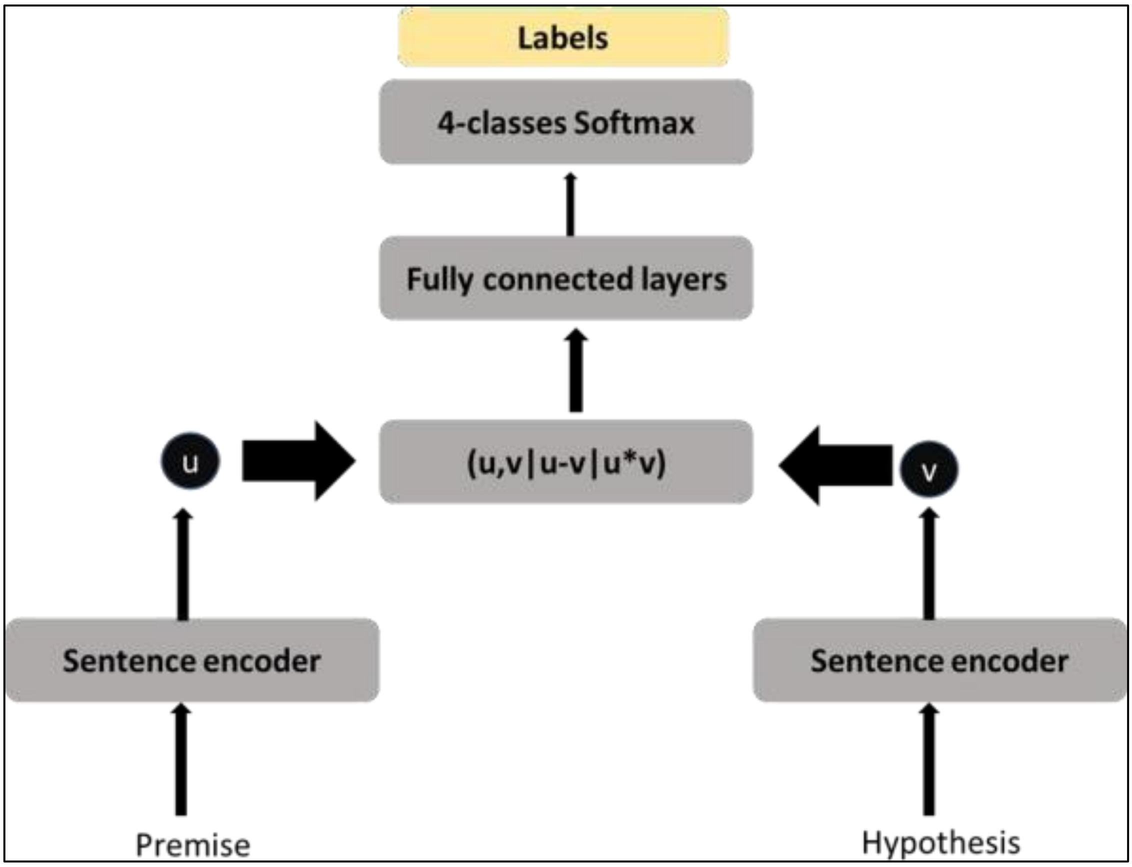

Figure 3.4.1: Typical classification model

We used Keras library for building our BiLSTM model. 100d Glove embedding is used for inputting in the model. The model which is used in this paper is a sequential model. Embedding, dropout layers, fully connected layer with 256 neurons and relu activation are used. In the data set, we have multiclass. That's why, soft-max activation is applied in the output layer. In this model, all data from the training dataset is used for training and all data from the test set is for testing. The model is trained using training data with 20 epochs and 128 batch size. This model gave 36% accuracy and F1-macro score is 29.

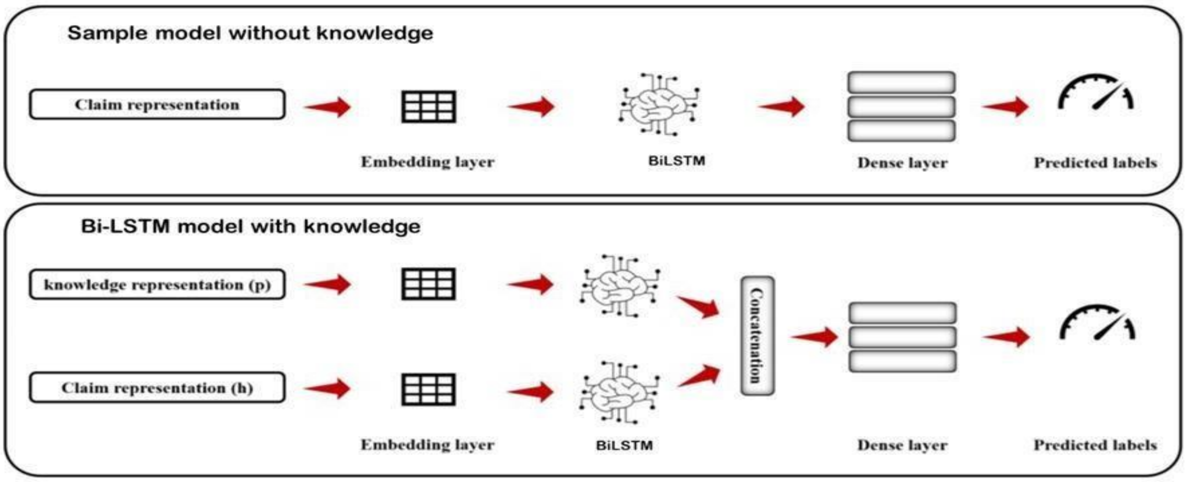

Figure 3.4.2: Sample model vs BiLSTM model

Table 3.4.3: Hyper-parameters of the model

Hyper-parameter	Selection
Embedding dimension	100
Hidden layer	2
Number of classes	4
Dropout	0.3
Maximum sequence length	3200
Lstm node	32
Batch size	128
Dense	256
Activation function	Softmax
Loss function	Sparse_categorical_crossentropy
Optimizer	Adam

4. Results & discussions

The output of the proposed model provides the identity of the given news. The news is true, false, partially false or other. To find the identity of the news, True is considered as 1, false as 0, partially false as 3 and other as 2 because the RNN model cannot work with text. We used 20 epochs to train our model and got 36% accuracy and 29% f1 -macro average score. The following table 4.1 shows about the classification report and table 4.2 shows the accuracy and f1 macro avg score of the model.

Table 4.1: Classification report

Class	Precision	Recall	F1-macro
False	0.62	0.83	0.71
True	0.18	0.24	0.20
Other	0.00	0.00	0.00
Partially False	0.18	0.41	0.25

Table 4.2: Accuracy report

Methodology	Accuracy	F1-macro avg
BiLSTM	36%	29%

The accuracy we achieved from our model is not unsatisfactory. The dataset which we used in our work is imbalanced which is responsible for our result. The following line graphs are shown that the graphical representation of the accuracy vs evaluation accuracy (Figure 4.3) and loss vs evaluation loss (Figure 4.4) which is achieved by our proposed model. It is clear from those graphs that our model is learning from the previous.

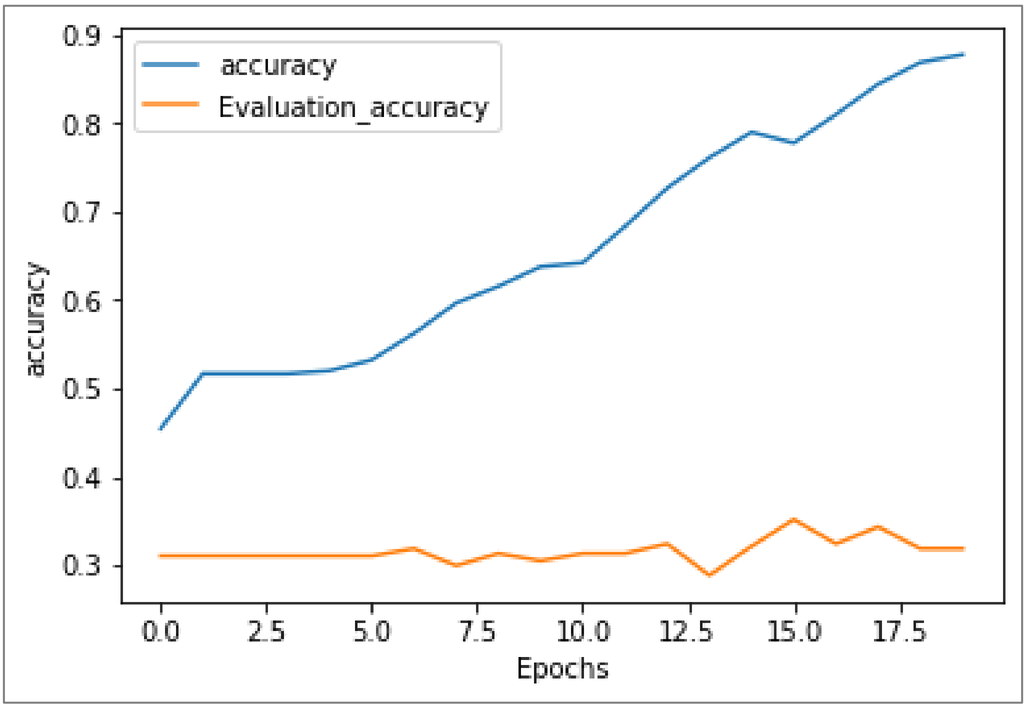

Figure 4.3: Accuracy vs evaluation accuracy

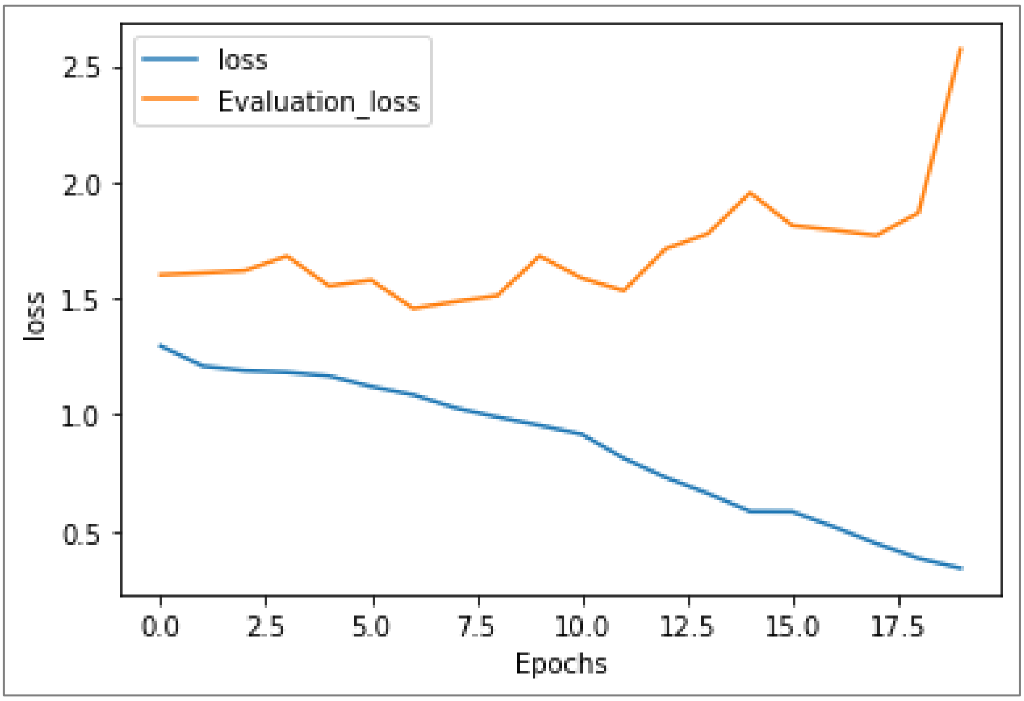

Figure 4.4: Loss vs evaluation loss

5. Conclusions

In this paper, we proposed a BiLSTM model for detecting fake news. We used some filter approaches to eliminate the irrelevant attribute from our dataset to improve the execution of the training time. Feature extraction techniques- Countvectorizer, word tokenizer is used on the features for converting into machine readable form. Features, stopwards, Stemming, etc. are used as data preprocessing tasks for eliminating less important words for training to save the execution time of our proposed model. This model gives overall accuracy of 36% with 29 F1-macro score which is not unsatisfactory. The imbalanced dataset is responsible for this result. There are some limitations and need some improvement. At first, our dataset is imbalanced which impacts on our result. We can combine BiLSTM and CNN to achieve better accuracy as well as the performance. We can increase our model performance by using the optimal hyperparameters which requires a lot of time and high-quality hardware components. In future, we will work with a large balanced dataset for validating our proposed model and try to overcome our limitations.

6. Acknowledgements

This is unbelievable support we've got from some of the faculty members and seniors of DIU NLP and ML Research Lab to continue our whole research flow from the beginning. We acknowledge Dr. Firoj Alam for guiding us and informing the secretes of research workshops. Also, we're thankful to Daffodil International University for the workplace support and the academic collaboration in some cases. Dr. Touhid Bhuiyan and Dr. Sheak Rashed Haider Noori also supported us with guidance, motivation, and advocating in institutional supports. Lastly, for sure we're thankful to our God always for every fruitful work with our given knowledge.

7. References

[1] Shu, K., Sliva, A., Wang, S., Tang, J. and Liu, H., 2017. "Fake news detection on social media: A data mining perspective." *ACM SIGKDD explorations newsletter*, *19*(1), pp.22-36.

[2] Umer, M., Imtiaz, Z., Ullah, S., Mehmood, A., Choi, G.S. and On, B.W., 2020. "Fake news stance detection using deep learning architecture (cnn-lstm)." *IEEE Access*, *8*, pp.156695-156706.

[3] Granik, M. and Mesyura, V., 2017, May. "Fake news detection using naive Bayes classifier." In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)* (pp. 900-903). IEEE.

[4] Bots and propagation of fake news. URL: https://www.cits.ucsb.edu/fake-news/spread [Online]. Accessed on: 25 July, 2021.

[5] Fake news websites list. URI: https://en.wikipedia.org/wiki/Fake_news_website [Online]. Accessed on: 25 July 2021.

[6] Zhang, J., Cui, L., Fu, Y. and Gouza, F.B., 2018. "Fake news detection with deep diffusive network model." *arXiv preprint arXiv:1805.08751*.

[7] Umer, Muhammad, Zainab Imtiaz, Saleem Ullah, Arif Mehmood, Gyu Sang Choi, and Byung-Won On. "Fake news stance detection using deep learning architecture (cnn-lstm)." *IEEE Access* 8 (2020): 156695-156706.

[8] Ahmed, H., Traore, I. and Saad, S., 2017, October. "Detection of online fake news using n-gram analysis and machine learning techniques." In *International conference on intelligent, secure, and dependable systems in distributed and cloud environments* (pp. 127-138). Springer, Cham.

[9] Granik, M. and Mesyura, V., 2017, May. "Fake news detection using naive Bayes classifier." In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)* (pp. 900-903). IEEE.

[10] Balwant, M.K., 2019, July. "Bidirectional LSTM based on POS tags and CNN architecture for fake news detection." In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.

[11] Kaur, S., Kumar, P. and Kumaraguru, P., 2020. "Automating fake news detection system using multi-level voting model." *Soft Computing*, *24*(12), pp.9049-9069.

[12] Braşoveanu, Adrian MP, and Răzvan Andonie. "Integrating Machine Learning Techniques in Semantic Fake News Detection." *Neural Processing Letters* (2020): 1-18.

[13] Karimi, H., Roy, P., Saba-Sadiya, S. and Tang, J., 2018, August. “Multi-source multi-class fake news detection.” In *Proceedings of the 27th international conference on computational linguistics* (pp. 1546-1557).

[14] Thota, A., Tilak, P., Ahluwalia, S. and Lohia, N., 2018. “Fake news detection: A deep learning approach.” *SMU Data Science Review*, *1*(3), p.10.

[15] Singhal, Shivangi, Rajiv Ratn Shah, Tanmoy Chakraborty, Ponnurangam Kumaraguru, and Shin'ichi Satoh. "Spotfake: A multi-modal framework for fake news detection." In *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*, pp. 39-47. IEEE, 2019.

[16] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). “*The Performance of LSTM and BiLSTM in Forecasting Time Series.” 2019 IEEE International Conference on Big Data (Big Data).* doi:10.1109/bigdata47090.2019.9005997.

[17] Morris, D.S., 2018. “Twitter versus the traditional media: A survey experiment comparing public perceptions of campaign messages in the 2016 US presidential election.” *Social Science Computer Review*, *36*(4), pp.456-468.

[18] Dwoskin, E., Dewey, C. and Timberg, C., 2016. “Why Facebook And Google Are Struggling to Purge Fake News.” *Washington Post*, *1*.

[19] Preslav Nakov., Da San Martino., Giovanni., Tamer Elsayed., Alberto Barron Cedeno., Ruben Miguez., Shaden Shaar., Firoj Alam., Fatima Haouari., Maram Hasanain., Nikolay Babulkov., Alex Nikolov., Shahi, Gautam Kishore., Struß, Julia Maria., Thomas Mandl., Sandip Modha., Mucahid Kutlu., Yavuz Selim Kartal., “Overview of the {CLEF}-2021 {CheckThat}! Lab on Detecting Check-Worthy Claims, Previously Fact-Checked Claims, and Fake News" CLEF~'2021 ,2021.

[20] Shahi, Gautam Kishore., Struß, Julia Maria., Thomas Mandl., “Overview of the {CLEF}-2021 {CheckThat}! Lab Task 3 on Fake News Detection” CLEF~'2021, 2021.

[21] Shahi, Kishore, G., Strub, Maria, J., Mandl, T., “Task 3: Fake News Detection at {CLEF}-2021 {CheckThat}!” 10.5281/zenodo.4714517, https://doi.org/10.5281/zenodo.4714517, 2021.