

INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING
2019, ICRTAC 2019

Fake News Detection using Bi-directional LSTM-Recurrent Neural Network

Pritika Bahad^{a,*}, Preeti Saxena^a, Raj Kamal^b

^a *Devi Ahilya University, Indore, 452001, India*

^b *Prestige Institute of Engineering, Management and Research, Indore, 452010, India*

Abstract

Media plays a vital role in the public dissemination of information about events. The rapid development of the Internet allows a quick spread of information through social networks or websites. Without the concern about the credibility of the information, the unverified or fake news is spread in social networks and reach thousands of users. Fake news is typically generated for commercial and political interests to mislead and attract readers. The spread of fake news has raised a big challenge to society. Automatic credibility analysis of news articles is a current research interest. Deep learning models are widely used for linguistic modeling. Typical deep learning models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) can detect complex patterns in textual data. Long Short-Term Memory (LSTM) is a tree-structured recurrent neural network used to analyze variable-length sequential data. Bi-directional LSTM allows looking at particular sequence both from front-to-back as well as from back-to-front. The paper presents a fake news detection model based on Bi-directional LSTM-recurrent neural network. Two publicly available unstructured news articles datasets are used to assess the performance of the model. The result shows the superiority in terms of accuracy of Bi-directional LSTM model over other methods namely CNN, vanilla RNN and unidirectional LSTM for fake news detection.

© 2019 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING 2019.

Keywords: Deep learning; Convolutional Neural Network; Recurrent Neural Network; Long Short-Term Memory; Bi-directional LSTM.

* Corresponding author. Tel.: +91-998-134-4415

E-mail address: bahad.pritika@gmail.com

1. Introduction

Social media facilitates the creation and sharing of information that uses computer-mediated technologies. This media changed the way groups of people interact and communicate. It allows low cost, simple access and fast dissemination of information to them. Majority of people search and consume news from social media rather than traditional news organizations these days. At one side, where social media have become a powerful source of information and bringing people together, on the other side it also put a negative impact on society. Look at some examples herewith;

Facebook Inc's popular messaging service, WhatsApp became a political battle-platform in Brazil's election [1]. False rumors, manipulated photos, de-contextualized videos, and audio jokes were used for campaigning. These kinds of stuff went viral on the digital platform without monitoring their origin or reach.

A nationwide block on major social media and messaging sites including Facebook and Instagram was done in Sri Lanka after multiple terrorist attacks in the year 2019 [2]. The government claimed that "false news reports" were circulating online. This evident the challenges the world's most powerful tech companies face in reducing the spread of misinformation.

Such examples show that Social Media enable the widespread of "fake news" as well. The news disseminated on social media platforms may be of low quality carrying misleading information intentionally. This sacrifices the credibility of the information. Millions of news articles are being circulated every day on the Internet – how one can trust which is real and which is fake? Thus incredible or fake news is one of the biggest challenges in our digitally connected world. Fake news detection on social media has recently become an emerging research domain [3]. The domain focuses on dealing with the sensitive issue of preventing spread of fake news on social media.

Fake news identification on social media faces several challenges. Firstly, it is difficult to collect fake news data. Furthermore, it is difficult to label fake news manually. Since they are intentionally written to mislead readers, it is difficult to detect them simply on the basis of news content [4]. Furthermore, Facebook, Whatsapp, and Twitter are closed messaging apps. The misinformation disseminated by trusted news outlets or their friends and family is therefore difficult to be considered as fake. It is not easy to verify the credibility of newly emerging and time-bound news as they are not sufficient to train the application dataset [4]. Significant approaches to differentiate credible users, extract useful news features and develop authentic information dissemination system are some useful domains of research and need further investigations.

There are several approaches to handle the problem of misinformation on social media. Statistical techniques are used to identify the correlation between various features of the information, analyzing the originator of the information, analyzing patterns of dissemination. Machine learning algorithms are used for classification of unreliable content and analyzing the accounts that share such content. Various approaches focus on developing techniques for authentication of information and on specific case studies as well.

The rest of this paper is structured as follows: Section II defines the fake news detection problem and summarizes related work to detect fake news using machine learning and deep learning methods. Section III presents a fake news detection model based on Bi-directional LSTM-recurrent neural network. In Section IV experimental settings and evaluation results are presented. Section V presents the conclusions and outlines future work.

2. Related Work

Fake news is "a news article that is intentionally and verifiably false" [5]. A news article is a sequence of words. Hence in past, many authors propose the use of text mining techniques and machine learning techniques to analyze news textual data to predict the news credibility. With more computational capabilities and to handle massive datasets, deep learning models present a finer performance over traditional text mining techniques and machine learning techniques. Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN), are widely explored Deep Neural Network (DNN) architectures to handle various NLP tasks [6][7]. The current work is related to number of research areas such as text classification, rumor detection, spammer detection, and sentiment analysis.

Fake news can be identified using different machine learning methods. Authors [8] proposed a simple approach for fake news detection using naive Bayes classifier is tested against a data set of Facebook news posts. Performance

evaluation of multiple classification algorithms namely Support Vector Machines, Gradient Boosting, Bounded Decision Trees and Random Forests on a corpus of about 11,000 news articles are presented in [9].

The word embedding is the neural representation of a word as a real-valued vector. The word embedding enables to measure word correlation by calculating the distance between two embedding vectors. Neural networks reveal their best performance in many NLP tasks with the pre-trained word embeddings [7]. A Neural network architecture based on term frequency-inverse document frequency (TF-IDF) and bag-of-words (BOW) representations as input to a multi-layer perception (MLP) are used for stance detection in news articles[10]. Automatic fake news detection based on surface-level linguistic patterns is proposed in [11] and a novel convolutional neural network to integrate meta-data with text is designed. A recurrent neural network based model [12] is proposed used for dialogue act classification. Context or sequence information is also explored in this model. A novel recurrent neural network based method is designed in [13] that learn continuous representations of microblog events for identifying rumors. Researchers proposed a deep learning approach for aspect-specific sentiment analysis [14].

3. Bi-directional LSTM-RNN Model for fake news detection

Fake news detection problem is a binary classification problem. The proposed solution detects fake news by determining whether the information present in the article is correct on the basis of quantifying the bias of a written news article and by analyzing the relationship between the news article headline and article body. This section describes the structure of model and the data processing method used for the experimentation.

3.1. Convolutional Neural Network

A convolutional neural network is feed-forward artificial neural networks and is normalized versions of multilayer perceptions. A CNN is a neural-based approach which represents a feature function applied to constitute words or n-grams to extract higher-level features [15]. The resulting features are successfully used for sentiment analysis, machine translation, text classification, and other NLP tasks.

The sentences are first tokenized into words to perform sentence modeling and classification with a basic CNN. Words are converted into a word embedding matrix (input embedding layer) of d dimension. An input embedding layer can be represented as function $y=f(x)$ where both the input x and output y are tensors. Various word embedding models available such as word2vec, GloVe[16], or FastText may be used for converting sentences into sentence matrices. Convolutional filters of different window sizes are applied to this input embedding layer to generate a new feature representation. Pooling method is applied on new features and pooled features from different filters are concatenated with each other to form hidden representation. These representations are then followed by one or multiple fully connected layers to make the final prediction.

3.2. Recurrent Neural Network

Recurrent Neural Network is a feed-forward artificial neural network. RNNs handle a variable-length sequence input by comprising a recurrent hidden layer whose activation at each time is dependent on the previous time. Hence RNNs are better choice for long-distance contextual information [17].

A basic RNN can be formalized as follows: given an input sequence (x_1, x_2, \dots, x_T) and for each timestamp, the model update the hidden state (h_1, h_2, \dots, h_T) and output (y_1, y_2, \dots, y_T) . Figure 1 shows the general architecture of basic RNN. The vectors x_t and y_t are input and output vector at timestamp t . The three connection weight matrices are W_{IH} , W_{HH} and W_{OH} represent the weight corresponding to input, hidden and output vectors respectively.

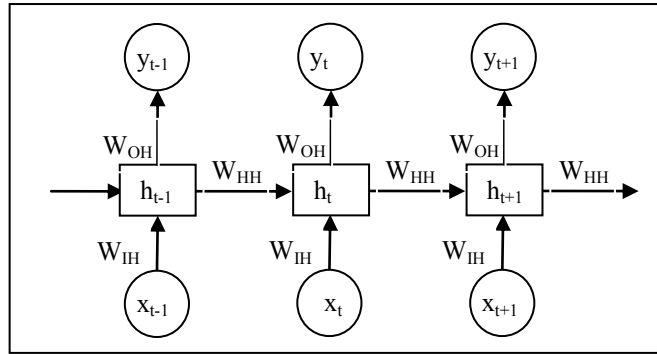


Fig 1. General architecture of basic RNN [17]

The operational behavior of RNN can be described as a dynamic system by a pair of non-linear matrix equations based on the function of the hidden and output unit activation functions f_H and f_O as follows:

$$h_t = f_H(W_{IH}x_t + W_{HH}h_{t-1} + b_h) \quad (1)$$

$$y_t = f_O(W_{OH}h_t + b_y) \quad (2)$$

where b_h and b_y are bias associated with hidden and output vector respectively.

RNN maintains state information across time steps that allow for processing of variable-length inputs and outputs. In the context of credibility analysis of a news article on the whole news articles are of variable length. To access the credibility whether a news article is real or not, a word as token the resultant score of previous state will be considered as input to current state. RNNs perform well for short context news articles.

RNN uses the back-propagation algorithm for training. During neural network training, the minimum of the error function is set up by iteratively taking small steps in direction of the negative error derivative with respect to networks weights. The magnitude of the gradients gets exponentially smaller for each consequent layer. As a result RNN suffers from vanishing gradient problem in the lower layers of a deep network. The three possible solutions to handle vanishing gradient problem is to use Rectified Linear Unit (ReLU) activation function; to use RMSProp adaptive learning optimization algorithm and/or to use different network architecture such as Long Short-Term Memory networks or Gated Recurrent Unit. ReLU activation function gives unit output while calculating gradient.

3.3. Long Short-Term Memory - Recurrent Neural Network

Long Short-Term Memory networks (LSTM) are a special type of RNN competent in learning long-term dependencies [18]. LSTM is a very effective solution for addressing the vanishing gradient problem. In LSTM-RNN the hidden layer of basic RNN is replaced by an LSTM cell as in Figure 2.

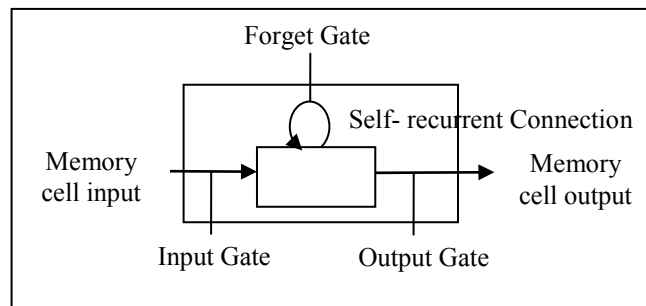


Fig 2. Structure of LSTM cell [18]

3.4. Bi-directional Long Short-Term Memory - Recurrent Neural Network

LSTMs help to preserve the error that can be back-propagated through time and in lower layers of a deep network [19]. Bi-directional processing is an evident approach for a large text sequence prediction and text classification. As shown in Figure 3, a Bi-Directional LSTM network steps through the input sequence in both directions at the same time.

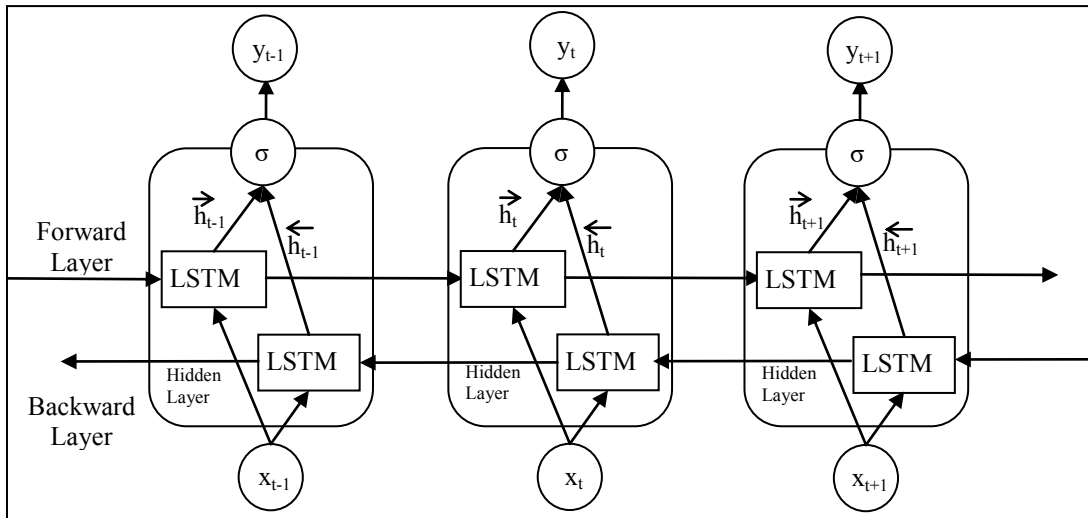


Fig 3: General architecture of Bi-directional LSTM- RNN [18]

The proposed fake news detection model based on Bi-directional LSTM-recurrent neural network is shown in Figure 4. The news articles are first pre-processed. A binary label is set to each news article as 1 for fake news and 0 for real news. The input news articles are converted to UTF-8 format and punctuation and stop words are removed. The news articles' title and content text are turned into space-separated padded sequences of words. These sequences are further split into lists of tokens.

Global Vectors for Word Representation (GloVe) embeddings is provided by Stanford NLP team [16]. It is an unsupervised learning algorithm for obtaining vector representations for words. Pre-trained GloVe word embeddings are used to deal with the high dimensional news articles. The embedding layer will load the weights from GloVe instead of loading random weights. GloVe applies globally aggregated co-occurrence statistics across all words in the news article corpus. The resulting representations formalize significant linear substructures of the word vector space. The transformed vector represented data is partitioned into train, validation and test data. The training is carried out on the news article corpus. Validation data set is used for fine-tuning the model. Further, the test data is used to know the predicted label of news article based on trained model. The model selection among CNN, variation of RNN as Vanilla RNN, LSTM-RNN, and Bi-directional LSTM-RNN is carried out to detect fake news. If CNN is selected the model, each embedding layer corresponding to training data is fed into CNN. Different filter sizes are used for evaluation of CNN performance. For example, if the filter size is 3, the filter will stride through the document computing the calculation above with all the trigrams. Global Max Pooling layer is used to extract the maximum value from each filter. The resultant is passed through several dense hidden layers with dropout. Finally softmax layers are used to make a binary decision of whether or not the article is credible. Similarly in Bi-Directional LSTM network each embedding layer corresponding to training data is inspected in both orders at the same time.

Model is trained iteratively to minimize loss function and to improve accuracy. The cross-entropy loss is considered to detect fake news article in proposed model. Various adaptive learning optimization algorithms namely AdaGrad, RMSProp, and Adam are considered to improve the model performance.

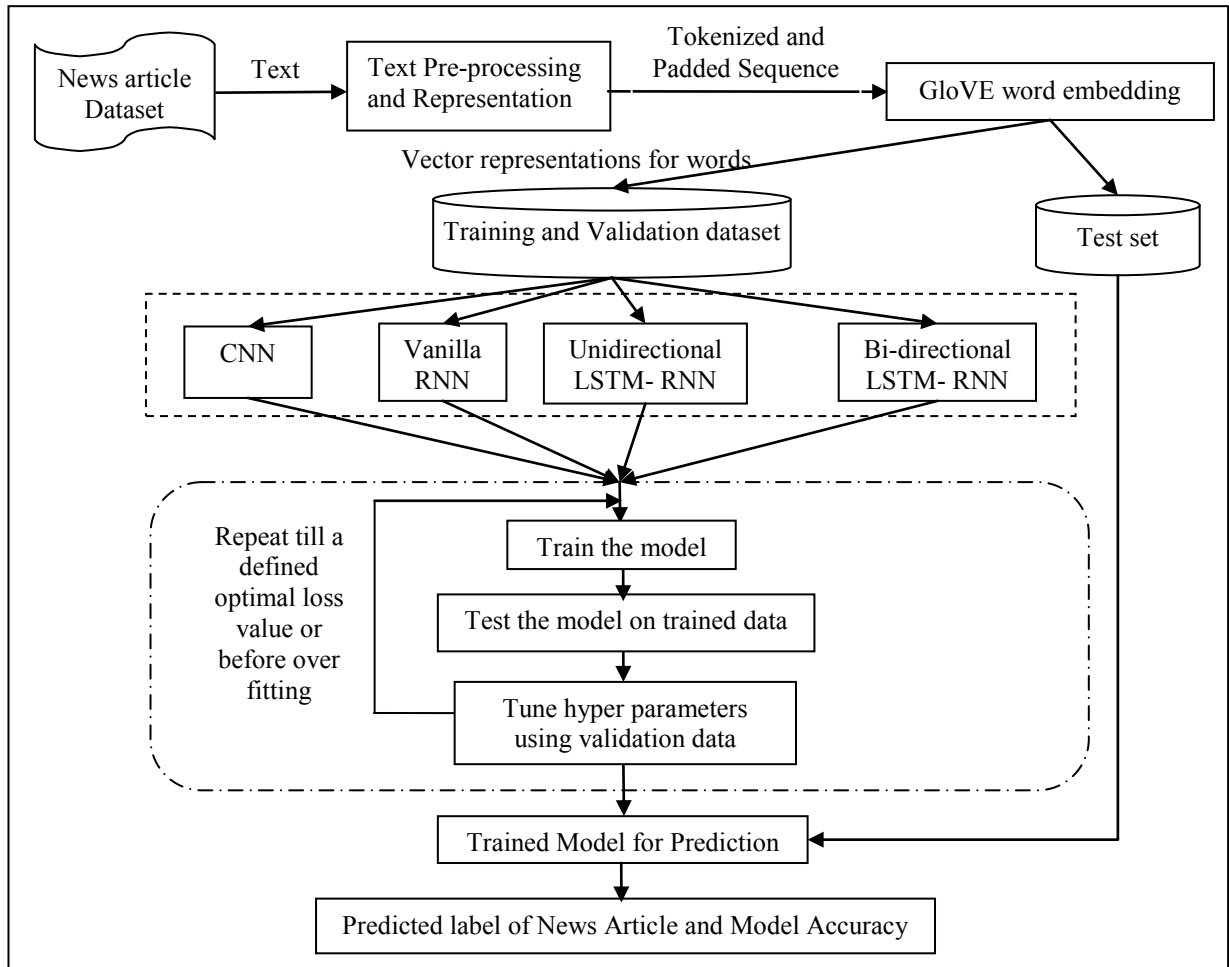


Fig 4. Proposed fake news detection model based on Bi-directional LSTM-recurrent neural network

4. Experiments and Results

All code are written in Python 2.7, using TensorFlow r0.12. All experiments have been performed on a Core™ processor Intel® CPU i7-4790U 3.60 GHz with 16 GB RAM. The two datasets utilized in this study are obtained from open Machine Learning Repository accessible at Kaggle [20][21].

DS1: Each news article in this dataset [20] consists of its id, title, content and text label as REAL or FAKE. The vocabulary size for this dataset is approximately 29 MB. DS1 Dataset specification is shown in Table1.

Table 1. Specification of News Dataset DS1

Attribute	Type
# (id)	Numeric
title	Text
text	Text
label	Text (either REAL or FAKE)

DS2: Each news article in this dataset [21] consists of article source, headline, the body of article and binary label as 1/0 is allotted to the real/fake news articles, respectively. The vocabulary size for this dataset is approximately 12 MB. DS2 Dataset specification is shown in Table 2.

Table 2. Specification of News Dataset DS2

Attribute	Type
URLs	Text
Headline	Text
Body	Text
Label	Numeric (either 1 for REAL or 0 for FAKE)

The proposed model is trained on both the datasets DS1 and DS2. To train the models AdaGrad, RMSProp and Adam adaptive learning rate algorithms are chosen. RMSProp is unpublished adaptive learning optimization algorithm that clips the gradients when it goes higher than a defined threshold. RMSProp shows better performance for large size news articles as it does not suffers from vanishing gradient problem.

The performance accuracy of the candidate model can be done using various available evaluation metrics. To check the prediction accuracy whether the news article is correctly classified as REAL or FAKE, accuracy as evaluation metric is used. Accuracy is calculated as ratio of number of correctly predicted samples to total number of samples for a given data set.

Considering the size of unique token set of the news datasets and to avoid over fitting and under fitting, the datasets are partitioned as 60:20:20 to train, validate and test the model. The experiment results corresponding to adaptive learning rate algorithm RMSProp with train, validation and test partitioning of dataset as 60:20:20 is shown in Table 3.

Table 3. Fake News Detection Results

(a) For News Dataset DS1

Model	Training Accuracy	Validation Accuracy	TestAccuracy
CNN	0.9942	0.8942	0.9077
Vanilla RNN	0.9971	0.7601	0.7822
Unidirectional LSTM-RNN	0.9939	0.8903	0.9148
Bi-directional LSTM-RNN	0.9992	0.8974	0.9108

(b) For News Dataset DS2

Model	Training Accuracy	Validation Accuracy	TestAccuracy
CNN	0.9992	0.9738	0.9833
Vanilla RNN	0.9971	0.9701	0.9638
Unidirectional LSTM-RNN	0.9996	0.9738	0.9863
Bi-directional LSTM-RNN	1	0.9825	0.9875

The data set DS1 partitioned in 60:20:20 and consist of real and fake news in training, validation and testing partition as [1934, 1867], [636, 631] and [601, 666] respectively. It consists of 104308 unique tokens and total 400000 word vectors in GloVe. Similarly during evaluation of model for dataset DS2, it is again partitioned in 60:20:20 and consist of real and fake news in training, validation and testing partition as [1293, 1112], [422,380] and [422, 380] respectively. It consists of 58324 unique tokens and total 400000 word vectors in GloVe.

Figure 5 and Figure 6 show the accuracy of the original CNN, Vanilla RNN, Unidirectional LSTM-RNN and Bi-directional LSTM on training, validation and test split on news dataset DS1 and DS2 respectively.

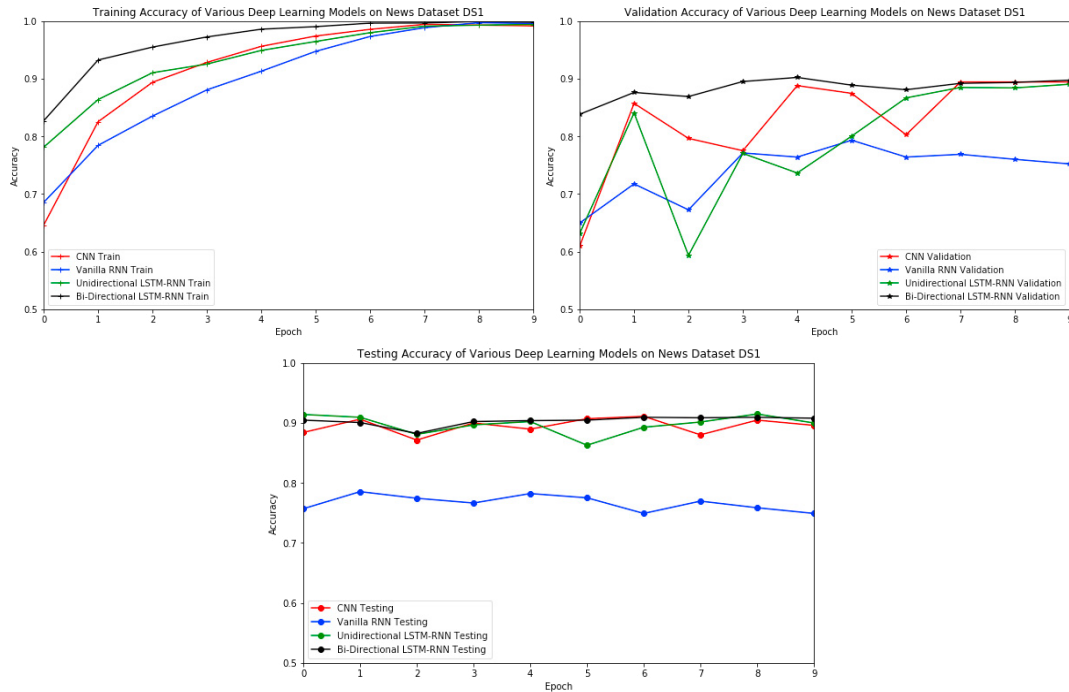


Fig.5. Accuracy of Fake News Detection Model on News Dataset DS1

In the model training, the RNN model suffers vanishing gradient problem due to the deep network hierarchy. LSTM-RNN solves the vanishing gradient problem. LSTM-RNN model has higher prediction accuracy than the original CNN model.

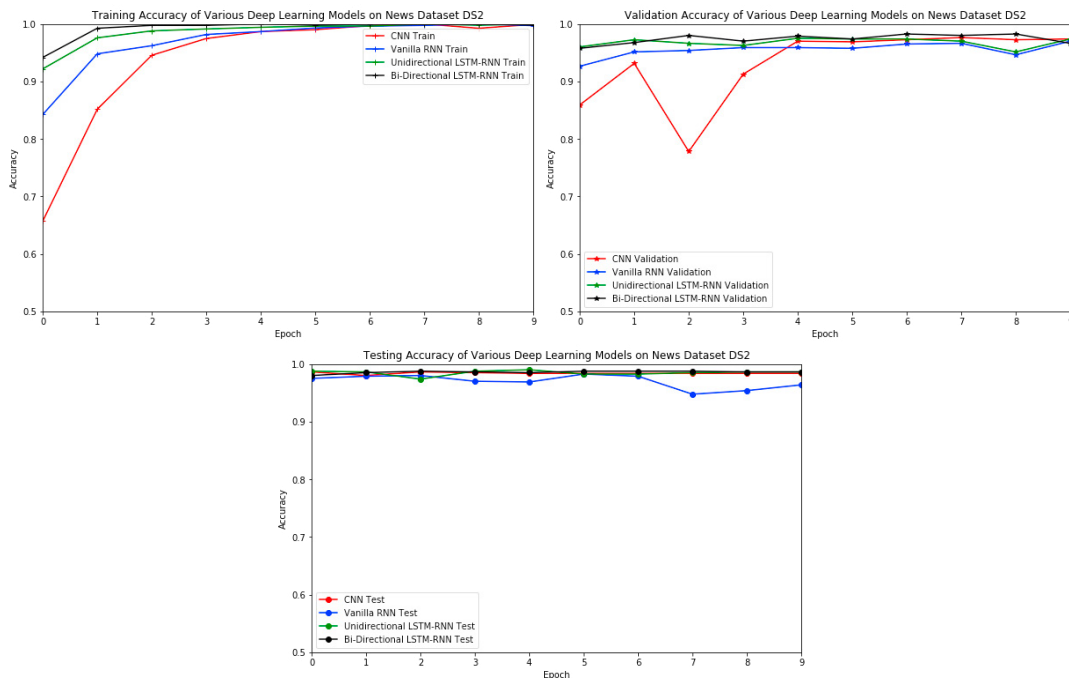


Fig.6. Accuracy of Fake News Detection Model on News Dataset DS2

5. Conclusion

The aim of this work is to predict fake news article using deep learning models. In the proposed model, GloVe word embedding is applied to measure relationship between the news article headline and body. The co-occurrence probabilities for headline words and words from the article body vocabulary weigh up the credibility of news. GloVe deals well with variance of the high dimensional news articles. The accuracy of Bi-directional LSTM-RNN model with CNN, vanilla RNN, and unidirectional LSTM-RNN are evaluated and compared. CNN performs better for extracting local and position-invariant features while LSTM-RNN is well suited for a long-range semantic dependency based classification. RNN work better for tasks where sequential modeling is more important. The results show Bi-directional LSTM-RNN model is significantly more effective than unidirectional models. The experiments also show that the choice of the adaptive learning rate algorithm plays a major role in the output to handle vanishing gradient problem of RNN. The proposed model works well for the balanced and imbalanced high dimensional news data set. More thorough experiments will be required in the future to further understand how deep learning model with attention can help to evaluate the automatic credibility analysis of News.

References

- [1] Reuters World News, October 21, 2018, <https://www.reuters.com/article/us-brazil-election-whatsapp-explainer/facebooks-whatsapp-flooded-with-fake-news-in-brazil-election-idUSKCN1MU0UP>, last accessed 2019/07/13
- [2] CNN Business April 21, 2019, <https://edition.cnn.com/2019/04/21/tech/sri-lanka-blocks-social-media/index.html>, last accessed 2019/07/13.
- [3] Fake News Challenge, <http://www.fakenewschallenge.org/>, last accessed 2019/07/13.
- [4] William Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang and Huan Liu. (2017). "Fake News Detection on Social Media: A Data Mining Perspective", SIGKDD Explorations: **19(1)**.
- [5] H. Allcott and M. Gentzkow. (2017) "Social Media and Fake News in the 2016 Election," *Journal of Economic Perspectives*, **31(2)**: 211–236.
- [6] Wenpeng Yin, Katharina Kann, Mo Yu, and Hinrich Schütze. (2017). "Comparative Study of CNN and RNN for Natural Language Processing".
- [7] Siwei Lai, Liheng Xu, Kang Liu, Jun Zhao. (2015). "Recurrent Convolutional Neural Networks for Text Classification", *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*
- [8] Granik, Mykhailo, Volodymyr Mesyura. (2017). "Fake News Detection using Naive Bayes Classifier." *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*: 900-903.
- [9] S. Gilda. (2017). "Evaluating Machine Learning Algorithms for Fake News Detection," *IEEE 15th Student Conference on Research and Development (SCORED)*, Putrajaya: 110-115.
- [10] Bourgonje, Peter, Moreno Schneider, Julian and Rehm, Georg. (2017). "From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles". *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism*: 84-89.
- [11] Wang, William Yang. (2017). "'Liar, Liar Pants on Fire': A New Benchmark Dataset for Fake News Detection." *ACL*.
- [12] Liu, Yang & Han, Kun & Tan, Zhao & Lei, Yun. (2017). "Using Context Information for Dialog Act Classification in DNN Framework", *Proceedings of the Conference on Empirical Methods in Natural Language Processing*: 2170–2178.
- [13] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Jim Jansen, Kam-Fai Wong and Meeyoung Cha. (2016). "Detecting Rumors from Microblogs with Recurrent Neural Networks", *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*: 3818-3824.
- [14] Lakkaraju, Himabindu, Richard Socher, and Chris Manning, (2014). "Aspect Specific Sentiment Analysis using Hierarchical Deep Learning", *NIPS Workshop on deep learning and representation learning*.
- [15] Kim, Yoon. (2014). "Convolutional Neural Networks for Sentence Classification". *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. 10.3115/v1/D14-1181.
- [16] GloVe: Global Vectors for Word Representation, <https://nlp.stanford.edu/projects/glove/>, last accessed 2019/07/13.
- [17] Hassan and A. Mahmood. (2018). "Convolutional Recurrent Deep Learning Model for Sentence Classification." *IEEE Access* **6**:13949-13957
- [18] Hochreiter, J. Schmidhuber. (1997). "Long Short-Term Memory", *Neural Computation*, **9(8)**:1735-1780.
- [19] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, J. Schmidhuber. (2015). "LSTM: A search space odyssey." *IEEE Transactions on Neural Networks and Learning Systems*.
- [20] real_or_fake, <https://www.kaggle.com/rchitic17/real-or-fake>, last accessed 2019/07/13.
- [21] Fake News detection, <https://www.kaggle.com/jruvika/fake-news-detection>, last accessed 2019/07/13.