International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Fake News Detection using LSTM

Tejaswini Yesugade[1], Shrikant Kokate[2], Sarjana Patil[3], Ritik Varma[4], Sejal Pawar[5]

[1,3,4,5]Computer Department, Pimpri Chinchwad College of Engineering, Pune, India
[2]Asst. Professor, Department of Computer Engineering, Pimpri, Chinchwad College of Engineering, Pune, India

---***---

Abstract —We are in the age of information, every time we read a piece of information or watch the news on TV, we look for a reliable source. There are so many fake news spread all over the internet and social media. Fake news is misinformation or manipulated news that is spread across the social media with an intention to damage a person, agency and organization. The spread of misinformation in critical situations can cause disasters. Due to the dissemination of fake news, there is need for computational methods to detect them. So, to prevent the harm that can be done using technology, we have implemented Machine Learning algorithms and techniques such as NLTK, LSTM. Our contribution is bifold. First, we must introduce the datasets which contain both fake and real news and conduct various experiments to organize fake news detector. We got better results compared to the existing systems.

Index terms—Embedding, LSTM, NLTK.

1. INTRODUCTION

Fake News is news, stories, or hoaxes created to d eliberately misinform or deceive readers. Usually, these stories are created to either influence people' s views, push a political agenda, or cause confusio n and can often be a profitable business for online publishers. The purpose of choosing this topic is because it is becoming a serious social challenge.It is leading to a poisonous atmosphere on the web and causing riots and lynching on the road. Examples: political fake news, news regarding sensitive topics such as religion, covid news like salt and garlic can cure corona and all such messages we get through social media. We all can see the damage that can be caused because of fake news which is why there is a dire need for a tool that can validate particular news weather it is fake or real and give people a sense of authenticity based on which they can decide whether or not to take action, amongst so much noise of fake news and fake data if people lose faith in information, they will no longer be able to access even the most vital information that can even sometimes be life- changing or lifesaving. Our approach is to develop a model wherein it will detect whether the given news is false or true using LSTM (long short-term memory) and other machine learning concepts such as NLP, word embedding, one hot representation, etc. The model will give us the results for the dataset provided. It gives accuracy up to 91.5%

2. RELATED WORK

In today's era spread of misinformation is become a very easy task because of social media. To stop this we need to find out news is fake or real. For which we are going to build a model which will identify that given news is fake or not using some ML and NLP concepts and algorithms.

Bag Of Words(BoW):

Bag Of Words is most commonly used in the methods of document classification[11][13]. BoW is Natural Language Processing method and Information Retrieval method. NLP model are used on the numbers we cannot use text data into our model. Therefor BoW model is used to preprocess text data by converting it into bag of words. In this method Frequency of the every word is used as feature in the classification.

***N**-grams:*

BoW is an order less documentation representation model in which only frequency of words is important[13]. n-grams is text classification model which mostly used in NLPand text mining[3]. N-grams is actually is a set of co-occurring words in given data and whencomputing n-grams move one word forward.

TF-IDF:

Term Frequency–Inverse Document Frequency is a numerical statistic that is intended to understand importance of a word is to a document in a dataset[5][3]. It is often used as a weighting factor in searches of information retrieval, text mining.

Term Frequency (TF):

Term Frequency is count of words present in the d ocument or an find out the inequality between the document[5][13]. Each document is characterized in a vector that contains the word count. This term i s calculated by Number of times term appears in a document divided by Total number of terms in the Document[3].

Inverse Document Frequency (IDF):

Inverse Document Frequency is the how many co mmon or rare words in the whole document or dat aset. This term is calculated by total number of doc uments, dividing it by the number of documents th at contain a word[5][3]. If the word is very common and appears in so many documents, then this will result as 0. Otherwise1.

Naïve Bayes:

Naïve Bayes uses probabilistic approaches and are ba sed on Bayes theorem[8]. They deal with probability distribution of variables in the dataset and predicting t he response variable of value. They are mostly used f or text classification. Bayes theorem is

$P(a|b) = p(b|a)p(b)/p(a)$

There are mainly 3 types of naïve base model as - Gaussian Naïve Bayes, Multinomial naïve Bayes and Bernoulli Naïve Bayes.We have used Multinomial Naïve Bayes model for our project to detect fake news[5][13]. An advantage of naïve Bayes classifieristhat o nly requires less training data for classification.

LSTM:

Long Short Term Memory is a kind of recurrent neur al network. In RNN output from the last step is fed as input in the current step. It tackled the problem of lo ngterm dependencies of RNN in which the RNN can not predict the word stored in the long term memory but can give more accurate predictions from the recen t information[5]. LSTM can by default retainthe infor mation for long period of time. It is used for processi ng, predicting and classifying on the basis of time series data.

Word Embedding:

Word embedding is a set of language modelling and f eature extraction techniques in Natural Language Pro-cessing (NLP). In word embedding words from vocab ulary are converted into the vectors of real numbers. Word embedding is type of word representation that allows words with similar meaning to have a similar r epresentation.

3. EXISTING SYSTEM

Detecting fake news is believed to be a comple x task and much harder than detecting fake product reviews.The open nature of the web and soci al media, inaddition to the recent advance incompu ter technologies, simplifies the process ofcreating and spreading fake news. While it's easier to unde rstand and trace the intention and the impact of fake reviews, the intention and the impact of creating propaganda by spreading fake newscannot be meas ured or understood easily.

For instance, it is clear that fake review affects the product owner, customers, and online stores; on th e other hand, it is not easy to identify the entitiesaffected by the fake news.

This is because identifying these entities requires measuring the news propagation, which has shownto be complex and resource intensive.

Working of Existing System:

Each is a representation of inaccurate or deceptive reporting. Furthermore, the authors weight the different kinds of fake news and the pros and cons of using different text analytics and predictive mod-el ling methods in detecting them. In their paper, the y separated the fake news types into 3 groups:-

1. Serious fabrications are news not published in m ainstream or participant media, yellow press, or ta bloids, which, as such, will be harder to collect [3].

2. Large-Scale hoaxes are creative and unique and often appear on multiple platforms. The authors argued that it may require methods beyond text anal ytics to detect this type of fake news.

3. Humorous fake news is intended by their writersto be entertaining, mocking, and even absurd. Acc ording to the authors, the nature of the style ofthis type of fake news could have an adverse effect on the effectiveness of text classification techniques.

It starts with preprocessing the dataset by removin g unnecessary characters and words from the data.The n-gram features are extracted, and a matrix of features is formed representing the documents inv olved. The last step in the classification process is to train the classifier. We investigated different classifiers to predict the class of the documents. We specifically investigated 6 different machine learning algorithms, namely, stochastic gradient descent (SGD), SVM, linear support vector machines (LS VM), K-nearest neighbor (KNN), LR, and decisio n trees (DT).

Term Frequency is a method that uses word count from texts to find similarities between texts[5]. Ea ch document is represented by a vector of equal le ngth that contains word counts. Next, each vectoris made in such a way that the sum of its element s will be added to the other. Each number of word s is converted into opportunities for such a word that is present in the documents. For example, if the word is something document, will be represented as 1, and if any not in the document, itwill be set t o 0. So, each the document is represented by group s of names. The typical TF of the word w in terms of document d is defined as follows: Standard Ti me = Value for Documentary / Total Number of D ocumentary Opposition (IDF) term w in reference to document corpus D, definedas IDF(w) D[5], by logarithm of the total number of documents in the corpus divided by the number of letters in which t he particular name appears, and is calculated as fo llows:

Inverted document TF = 1+log (total documents

/no of documents with particular term)

TF-IDF is a weighting metric often used in inform ation retrieval and NLP[3]. It is a statistical metric used to measure how important a term is to a docu ment in a dataset. Around 80% of the dataset is us ed for training and 20% for testing. After extractin g the features using either TF or TF-IDF, we train a machine learning classifier to decide whether a s ample's content is truthful or fake.

Naïve Bayes Model:

- Among the fields, that are present in the dataset, only few of them were used. They are link to the Facebook post with the text of the news article a nd the label of the text.

- Text of the news articles was retrieved using Fac ebook API[8]. News articles with labels “mixture of true and false” and “no factual content” were not considered. Couple of the articles in the datas et are broken they do not contain any text at all (or contain “null” as a text). These articles were ig nored as well. After such filtering data set with 1 771 news articles was obtained.

- The dataset was randomly shuffled, and after that divided into three subsets: training dataset, valida tion dataset, test dataset. Training dataset was use d for training the naive Bayes classifier[8]. Valid ation dataset was used for tuning some global par-ameters of the classifier. Test dataset was used t oget the unbiased estimation of how well the clas sifier performs on new data (it is a well known f act, that it is not correct to only have training and test datasets when parameter tuning is perfo-med, because received results on test set will be biased in this case).

- For the unconditional probability of the fact, that a ny news article is correct all of the values from int erval [0.2; 0.75] with step 0.01 were considered. F or the true probability threshold all of the values fr om interval [0.5; 0.9] with the same step were con sidered. The best results on the validation dataset were received with the unconditional probabilityof the fact, that any news article is correct being equ al to 0.59 and the true probability threshold being equal to 0.8.

- The global parameters, that were tuned, are the unc-onditional probability of the fact, that any news ar ticle is correct and the true probability threshold.

  The true probability threshold is such a value, that every article with probability to be true news articl e bigger than the threshold would be considered by the classifier as a true news article, and all other a rticles – as a false news article.

- Consider the classification procedure of the naive Bayes classifier. When iterating through the wordsof the news article that is being classified, a corner case is possible: some specific word might not be present in the training dataset at all. For all such words it was decided to define the probability of t he news article being fake given that it contains th is word as 0.5. Equation (4) won’t be affected in s uch case: indeed, both nominator and denominator get multiplied by 0.5. Basically, current implementation just ignores such words.

- If all of the words in the news article are unknown to the classifier (never occurred in the training da taset), the classifier reports, that it can not classify given news article.

- If some word occurred in the news article several times, it contributed to the total probability of the f act, that a news article is fake exactly the same number of times.

- Equation (4) is computationally unstable if calculat ed directly. This is caused by the fact, that lots of probabilities get multiplied, and the result of such multiplication becomes close to zero really fast. Most of programming languages do not provide th e needed degree of precision, and that’s why they interpret the result of multiplication as exactlyzero [8]. Let p be the probability of the fact, that given news article is fake. One can calculate the value 1/p-1 instead, and after that receive the value of p q uite easily. The following equation holds:

$1/p - 1 = p2 / p1$ (5),

where p, p1, p2 are the same as in (2), (3) and (4)

. p1 and p2 can be calculated in more stable way using algorithms and exponentiation.

Articles information loading

Articles filtering based on the presence of the content and relevant label

Separating data in the training, validation and test datasets

Training the naive Bayes classifier

Testing and accuracy evaluation

Received Result:

If classifier find news article as fake, then:

The true positive are the correctly classified as fake news articles.

The false positive examples are the incorrectly classi fied as fake articles.

The true negative are the correctly classified as real news articles.

The false negative examples are incorrectly classified as real news articles

4. PROPOSED SYSTEM

LSTM model:

Long short-

term memory (LSTM) units are a building block f or the layers of a recurrent neural network (RNN). A LSTM unit is composed of a cell, an input gate

, an output gate and a forget gate [12]. The cell is responsible for "remembering" values over a vast t ime interval so that the relation of the word in the starting of the text can influence the output of the word later in the sentence. Traditional neural netw orks cannot remember or keep the record of what a ll is passed before they are executed this stops the desired influence of words that comes in the sente nce before to have any influence on the ending wo rds, and it seems like a major shortcoming.

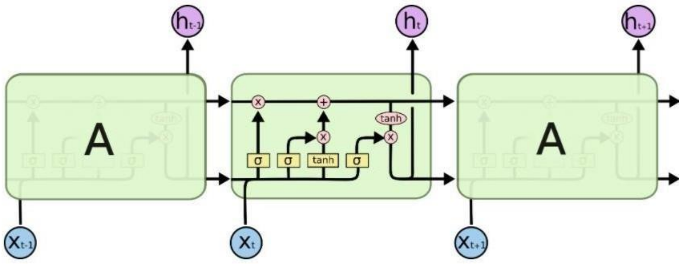

Figure 1: Architecture of LSTM

Overview of Dataset:

Dataset is taken from Kaggle platform. It has the fo llowing attributes: id: unique id for a news article, title: the title of a news article, author: author of the news article, text: the information of news article. Dataset consist of total 18285 news articles for trai ning and testing of model. Dataset is formed with c ombination of real and fake news.

Implementation details:

PREPROCESSING: To transform data into the rele vant format the data set needs pre-

process. Firstly, we removed all the NAN values fr om the dataset. Vocabulary size of 5000 words is d ecided. Then NLTK (Natural Language Processing) Tool Kit is used to remove all the stop words from the dataset. Stop words is list of punctuations + sto p words from nltk toolkit i.e. Words such as 'and' ' the' and 'I' that don't convey much information co nverting them to lowercase and removing punctuation. For each word in documents if it is not a stop word then that words tag is taken from postag. Then, this collection of words is appended to document.

WORD INDEX OF TOKENIZE DATASET:

Word tokenizing, appends text to a list and the list be named as documents. The output for this stage isthe list of all the words in the narration.

WORD EMBEDDING:

Onehot Representation: We cannot give input in th e form of text format to the algorithm so we have to convert them into the numeric form, for which we are using one hot representation. In onehot represen tation each word in the dataset will be provided its index from the define vocabulary size and these inde xes are replaced in sentence. While giving input to t he word embedding, we have to provide it with the fix length. To convert each sentence into the fix len gth padding sequences is used. We have considered max length of 20 words while padding title. Either we can provide padding before the sentence (pre) orafter the sentence (post), and then these sentences pass as input to the word embedding. Word embedd ing apply feature extraction on the provided input v ector. In total 40 vector features are considered.

MODEL:

Output from the word embedding is provided to the model. The machine learning model implemented here is a sequential model consisting of embedding as first layer which consist of values vocabulary siz e, number of features and length of sentence. The n ext is LSTM with 100 neurons for each layer, follo wed by Dense layer with sigmoid activation functio n as we need one final output. We have used binary cross entropy to calculate loss, Adam optimizer for adaptive estimation, finally adding drop out layer in between so that overfitting is avoided. Then trainin g and testing of model id done.

CLASSIFICATION:

For both preprocessed testing data the result is predi cted. If the predicted value>0.5 Classified as 1 is real and 0 is fake. Accuracy = (TP + TN) / Total. Th e following terms were used: True Negative (TN), i.e., the prediction was negative and test cases, too, were actually negative; True Positive (TP) i.e., the pre diction was positive and test cases, too, were actually positive; False Negative (FN) i.e., the prediction was negative, but the test cases were actually positi ve; False Positive (FP), i.e., the prediction was posi tive, but the test cases were actually negative

Model:

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        (None, 20, 40)            200000
_________________________________________________________________
lstm (LSTM)                  (None, 100)               56400
_________________________________________________________________
dense (Dense)                (None, 1)                 101
=================================================================
Total params: 256,501
Trainable params: 256,501
Non-trainable params: 0
_________________________________________________________________
None
```

Accuracy:

```
accuracy_score(y_test,y_pred)
```

```
0.9105824446267432
```

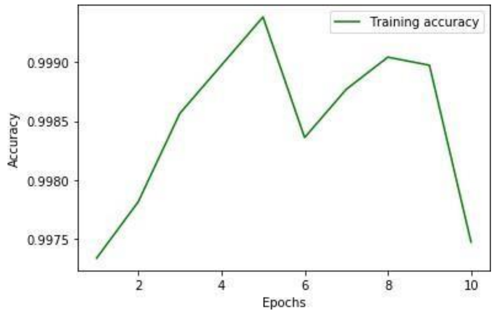

Figure 4: Accuracy chart.

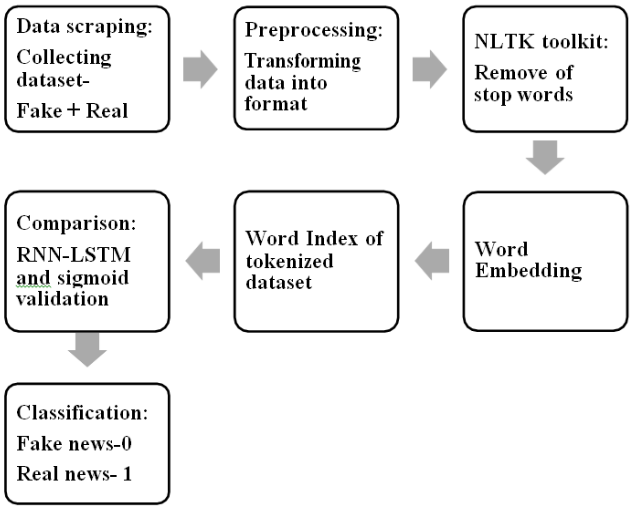

Figure 2: Architecture flow of proposed system.

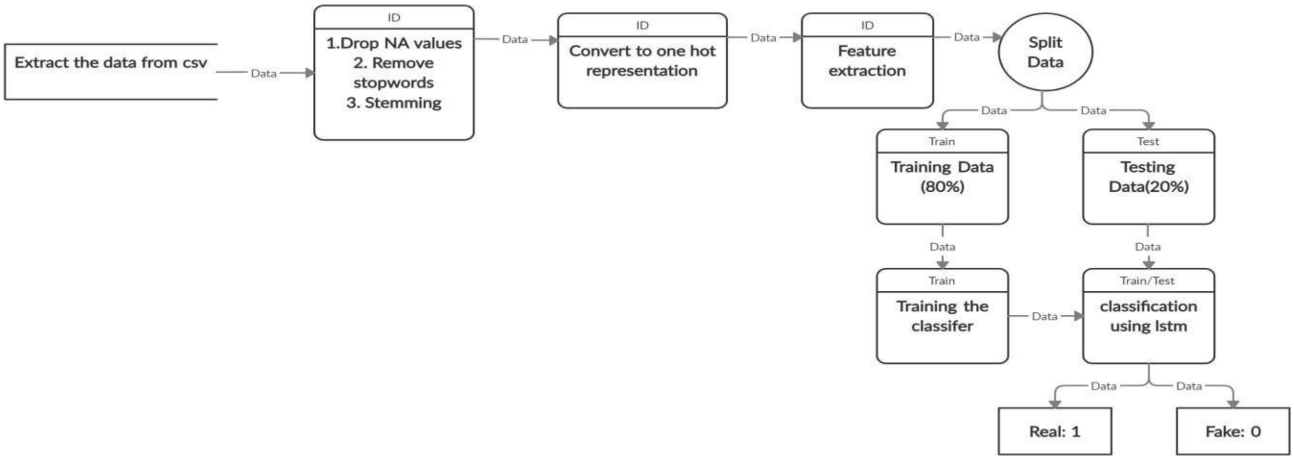

Figure 3: Proposed system module.

5. RESULTS

The classification accuracy for true news articles a nd false news articles is roughly the same, but classi fication accuracy for fake news is slightly deviated. By using the confusion matrix and the classification report further the accuracy of each individual mode l is measured.

N=3657	Predicted:NO	Predicted:YES
Actual: NO	TN= 1900	FP= 182
Actual: YES	FN= 145	TP=1430

- Get more data and use it for training. In mach ine learning problems it is often the case whe n getting more data significantly improves the performance of a learning algorithm. The data set, that was described in this article contains only around 18285 total news. From which 8 0% is taken for training i.e. 14628 and 20% is taken for testing i.e. 3657. Accuracy can bei ncreased by training the model on more data.

- Use the dataset with much greater length of t he news articles. The news articles, that were presented in the current dataset, usually were not that long. Training a classifier on a dataset with larger news articles should improve its p erformance significantly.

- Remove stop words from the news articles. S top words are the words, that are common toa ll types of texts (such as articles in English). These words are so common, that they don't really affect the correctness of the information in the news article, so it makes sense to get rid of them [14].

- Use stemming. In linguistic morphology and information retrieval, stemming is the process of reducing inflected (or sometimes derived) words to their word stem, base or root form – generally a written word form[15]. Such techn

  -ique helps to treat similar words (like "write" and "writing") as the same words and may i mprove classifier's performance as well.

6. EXISTING SYSTEM VS PROPOSED SYSTEM

Navie bayse classifier gives accuracy around 75% [16] which shows that LSTM is much more reliable with accuracy of 91%

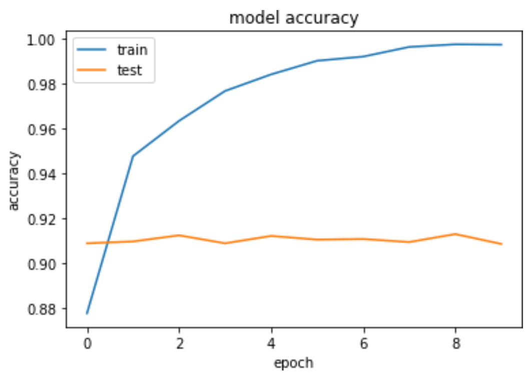

Figure 5: Model accuracy chart.

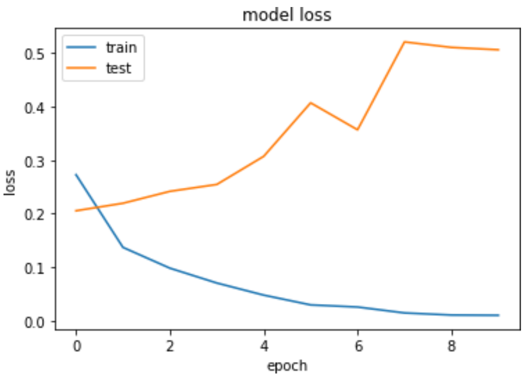

Figure 6: Model loss chart.

7. LIMITATIONS

While the results discussed herein suggest for model some external features like source of the news, author of the news, place of origin of the news, time stamp of news were not considered in our model whi ch can be influence the outcome of the model. Avai lability of datasets and literature papers are limited for fake news detection. The length of the news tha t is heading or whole news is less which affects the result in terms of accuracy. In Fake News with inc reasing in layer of module training time increases.

8. APPLICATION

Journalism:

The major spread of information and truste d source is through newspapers and news channels, so this detection can be used to verify the news before broadcasting it.

Social Media:

In today's world of social media, it is easyto manipulate any information or news. Such ma nipulated news misguides the readers. It is import ant to identify that news is fake or real. This paper provides various techniques that can be used in detection and classification of information.

9. CONCLUSION

In this digital age, where hoax news is present every where in digital platforms, there is an ultimate need f or fake news detection and this model serves its purp ose by being the need of the hour tool. Fake News re garding sensitive topics leads to a toxic environment on the web. Fake News Detection is the analysis of s ocially relevant data to distinguish whether it is real o r fake. Here in this paper we compared various meth ods like Bag Of Words(BoW),Ngrams, TF-IDF, Naïve Bayes etc. LSTM to be most effectiveof all we used various techniques like stop word removal, one hot r epresentation, word embedding and howLSTM can be used to get better results. Model mentioned in this p aper is very effective, Also compared toexisting system the model proposed here gives better results with a ccuracy of 91.05% which is very promising, we can f urther increase results by increasing training data.

10. REFERENCES

[1.][IEEE 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (A VSS) - Auckland, New Zealand (2018.11.27-2018.11.30)]2018 15th IEEE International Confere nce on Advanced Video and Signal Based Surveilla nce (AVSS) Fake Information and News Detection using Deep Learning.

[2.]International Institute of Information TecInternational Institute of Information Technolog y [2018], Bangalore, India. 3HAN A Deep Neural Network for Fake News Detectionhnology [2018], Bangalore, India. 3HAN -A Deep Neural Network for Fake News Detection.

[3.]ECE-Department University of Victoria, Canada-School of Computer Science, University of Windso r, Canada -Detecting Opinion spams and Fake News using Text Classification.

[4.]Fabula AI(UK), USI Lugano (Switzerland), Imperia l college of London-Fake News Detection on Social Media using Geometric deep learning.

[5.][IEEE 2018 4th International Conference on Computing Communication and Automation(ICCCA) - Greater Noida, India (2018.12.14-2018.12.15)] 2018 4thInternational Conference on Computing Communication and Automation (ICCCA) Fake News Detection Using A Deep Neural Network.

[6.]Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Au tomatic deception detection: Methods for finding fa ke news. Proceedings of the Association for Information Science and Technology.

[7.]Wu, Liang, and Huan Liu. "Tracing Fake- News Footprints: Characterizing Social Media Messages by How They Propagate."

[8.]Granik, Mykhailo, andVolodymyr Mesyura. "Fake news detection using naive Bayes classifier." Electrical and Computer Engineering (UKRCON), 2017 IEEE First Ukraine Conference on. IEEE, 2017.

[9.]Buntain, Cody, and Jennifer Golbeck. "Automati cally Identifying Fake News in Popular Twitter Threads." Smart Cloud 2017 IEEE International Conference on. IEEE, 2017.

[10.]Shu, Kai,"Fake news detection on social media: A data mining perspective." ACM SIGKDD Ex plorations Newsletter 19.1 (2017).

[11.]Bhatt, Gaurav "Combining Neural, Statistical a nd External Features for Fake News Stance Iden tification." Companion of the Web Conference 2018 on the Web Conference 2018. International World Wide Web Conferences Steering Committee, 2018.

[12.] S.Ananth, Dr.K.Radha, Dr.S.Prema, K.Nirajan International Journal of Innovative Research in Computer and Communication Engineering "Fa ke News Detection using Convolution Neural N etwork in Deep Learning".

[13.]Supervised Learning for Fake News Detection.[14.] Stop words. (n.d.) Wikipedia. [Online].

Available: https://en.wikipedia.org/wiki/S top_words. Accessed Feb. 6, 2017.

[14.] Stemming. (n.d.) Wikipedia.
[Online].
Available: https://en.wikipedia.org/wiki/ Stemming. Accessed Feb. 6, 2017

[15.] Mykhailo Granik, Volodymyr Mesyura-2017 IEEE First Ukraine Conference on Electri cal and Computer Engineering (UKRCON) – "Fake News Detection Using Naive Bayes Classifier"