2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT) | 979-8-3503-3509-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCCNT56998.2023.10307443

HINDI FAKE NEWS DETECTION USING STACKING ENSEMBLE METHOD

Sushama S. Nandgaonkar
Research Scholar
Department of Computer Sci. & Engineering
COEP Technological University Pune, India
sushma.nandgaonkar@gmail.com

Sunil B. Mane
Associate Professor ,
Department of Computer Sci. & Engineering
COEP Technological University Pune, India
sunilbmane.comp@coeptech.ac.in

***Abstract*— Day by day, usage of social media is increasing amongst the users giving rise to certain issues which need attention. Since information can be shared across social media very quickly, it has become a popular medium of information sharing. Users use these platforms to share information to individuals, groups and to communities. As information is spread across social media platforms very quickly, misinformation also gets spread very quickly. Detecting fake contents is very important at the initial stage so that spread of fake information can be prevented. In existing research most of the research carried out is for English news. Very few researchers have attempted to develop a model for other languages. We present a machine learning solution for fake information identification for Hindi News articles. We carried out experiments on three datasets using various machine learning approaches such as Support Vector Machine, Logistic Regression, Decision Tree, Multinomial NB and Random Forest, these learners are used individually, and their performance is evaluated. Then we use these learners in stacking ensemble method and evaluated the performance. To extract features from news TF/IDF is used. Results show that RF performed well over all datasets as individual learner. Ensemble learning resulted in significant improvement in the performance over individual learners in all performance evaluation metrics. In some cases, results of LR are closer to that of RF.**

Keywords*— *fake news detection, Indian languages, machine learning

I. INTRODUCTION

Social media platforms like WhatsApp, Twitter, Facebook, are used to exchange information daily by the users. This information can be of type text, audio, image, and video. These platforms provide the simplest and easiest way for sharing the information. But information spread across these platforms is not always correct. Sometimes false information is spread across these platforms intentionally to misguide the people. Many users forward information without bothering about its authenticity. Such fake information can cause chaos amongst the society hence authenticity of information must be detected at early stages to prevent spread of fake information. Its impact is observed at the time of COVID pandemic where the spread of fake information was at large scale. Antonio Guterres, the United Nation Secretary- General, said that even if our main enemy is COVID19, we have other enemy which is 'infodemic' of misinformation [1]. The authors presented statistics about how the number of mobile users and internet users increased and about how misinformation spread in this period.

According to [2,4,5,6,7] the US presidential election was believed to be influenced by the impact of fake news. As per the statistics provided in [5], In 2016 around 62% of US adults acquired the information from social media, whereas this percentage was 49% in 2012. It shows how usage and popularity of social media is increasing. X. Zhang et al. in [11] summarized characteristics of fake news in terms of volume, variety, velocity. Massive fake contents are present on the internet related to various domains. According to the authors, identification of online fake contents is difficult due to its real time nature. It is also difficult to identify the source of fake information as many accounts are short lived. Recently a lot of research is carried out on identifying fake contents. Linguistic features are considered in text-based models for detecting authenticity of the news. Many researchers have provided machine and deep learning-based solutions for misinformation detection. In this paper we have used algorithms such as Decision Trees, Extreme gradient boosting, Linear Support Vector Machine, Naïve Bayes, Random Forest, and Ensemble method to identify fake contents. Lot of research is carried out for identification of fake contents, but very limited work has been done in the case of regional languages. User can post content on social media in different languages, so it is essential to identify fake contents in different languages. We propose a ML based ensemble approach for detecting fake contents for Hindi news data. There are some facts checking news sites available for checking the authenticity of the news but to stop the spread of fake information across the media, an automated tool is required. We tested our model on three datasets collected from Kaggle and GitHub and a third dataset obtained by merging two datasets. Fake news detection systems can be implemented by considering content-based features and propagation-based methods. Content features evaluate the

features from the news text and these features are used to decide the authenticity of the news. Linguistic and visual features are considered in content features. News Content features are once again categorized into two types: based on knowledge which handles the news content features and other on style which is related to writing style of the news.

In a linguistic approach linguistic features are extracted. Knowledge features are of the form subject, predicate, and object. These features can be used to verify the news authenticity and can be represented in terms of knowledge graphs [22].

II. Related Work

A. Text features-based approach for English news

ML approach using n-gram is proposed by H. Ahmed et al. [6] using the frequency of the term (TF) and Term Frequency and Inverse Document Frequency (TF/IDF) for extracting the features for English News data They used ML algorithms such as SVM, LSVM, KNN, SGD, LR, and DT. Experiments were performed on political news dataset collected from Kaggle. They showed the performance of linear classifiers is better as compared with nonlinear classifiers. For evaluating the model n-grams of different sizes were used. In [7] Ahlem Drif et al. used CNN and LSTM based approach for identifying the fake news. Features are extracted from CNN layer and are passed through LSTM layer to find the relatedness between the terms in the input sequence. GloVe is used for word embedding. GloVe is an unsupervised algorithm for vectorization from the text data. For evaluating the model, they used Liar dataset and on the News Articles Dataset with reasonable accuracy. Julio C. S. Reis et al. [3] used models such as KNN, NB, RF, SVM, and XGB Buzzfeed dataset. Along with the textual features of the news they used news source features and environment features. Authors showed that the RF and XGB results are best amongst all the classifiers. Shuo Yang et al. [15] presented an unsupervised learning-based approach using a hierarchical user engagement model based on tweet, retweet and reply. Two datasets LIAR and BuzzFeed were used to test the model.

Monther Aldwairi et al. [10] used BayesNet, Logistic Regression, Random Tree and Naive Bayes classifiers on data collected from URLs of the clickbait by crawling the web. Accuracy with Logistic is highest over the dataset. Clickbait is a text used to get the attention of a user so that the user can follow the link and read the article. Authors proposed an approach to detect and remove the web pages containing false information. Z Khanam et al. [12] used different ML approaches such as LR, RF, XGB, NB, KNN, DT, SVM along with TfidfVectorizer or feature extraction. The LIAR dataset was used for experimentation. XGB resulted in 75% of accuracy which is highest amongst all classifiers. Iftikhar Ahmed et al [4] proposed an ensemble learner such as RF, voting classifier, bagging and boosting for identifying the fake news with different ML algorithms such as LR, MLP, KNN. Results obtained on ISOT are 99% of accuracy on ISOT dataset using RF. Feature extraction is performed using LIWC. Classifier XGB performed well on all evaluation metrics as compared to individual classifiers. Ensemble approach using stacking is proposed in [2] with ML and DL algorithms LR, KNN, RF, DT, SVM, CNN, LSTM, and GRU. TF/IDF is used for tokenization for ML algorithms and word embedding for deep learning algorithms. Experiments were performed on ISOT and KDnuggets data sets with accuracy of 99.9% and 96% respectively. The results were obtained using stacking method for ensemble. In [8], Nida Aslam et al presented a deep learning approach to deal with fake news. They used ensemble method. They used Bi-LSTM, and GRU ensemble learners on the LIAR dataset for dealing with textual attributes. For other attributes, a deep dense network is used. Dataset was divided in three groups based on features as statements, categorical and numerical and the third one containing all features. Accuracy of 89.8% is achieved over the LIAR dataset on the statement feature.

Rohit Kumar Kaliyar et al. in [13] presented a BERT based approach along with One-layer Convolutional Neural Network (CNN) with different filter and kernel and sizes to address the ambiguity in natural language processing. The experiment was performed on Political data collected from Kaggle and is related to the 2016 US President Election using CNN and LSTM and with a BERT based model with GloVe and BERT. These are pre-trained methods for word embeddings. FAKE BERT approach has the highest accuracy. In [9], a CNN-RNN based approach was used by the authors. Features are extracted using the CNN and RNN is used for estimating long term dependency. Experiments were performed on FA-KES and ISOT dataset. Different sized filters were used with CNN. These filters are used to extract number of feature maps containing the local features. The RNN layer consisting of LSTM Cells takes features produced by CNN as the input. Results on both the models were better than baseline models. Shaina Raza et al. [14] presented a method to identify the fake contents using transformers. Along with the news content they also considered social context features. Zero shot learning is used to decide a user's credibility. Model was tested on the datasets NELA-GT-2019 and Fakeddit. The accuracy obtained was 74.89%.

B. Fake news detection for Indian Languages

Normally we prefer to communicate with others using regional languages. Fake contents are also spread in regional languages. Detecting fake news in these languages is also essential. Few attempts were made by the researchers to build fake content identification models for other languages. This study reflects research carried out for Indian regional languages. Debanjana Kar et al. in [16] presented a technique using mBERT for multilingual fake content identification. They created a Hindi and Bengali news dataset for COVID19 news. They extracted different text and statistical features from the news data. They used a zero-shot learning approach in which tweets from one of the languages is used only as testing data. Considering only text features resulted in better performance. They achieved 79% F-Score for the Hindi tweets and 81% F-score over the Bengali Tweets using cross domain data augmentation. Using zero shot learning accuracy for Hindi tweets increases to 81%. Sudhanshu Kumar et al. in [19] presented a fake content identification method on Hindi

dataset. They used classifiers such as NB, LR along with DL model LSTM were used for experimentation purposes. TF/IDF was used for feature extractions. Since very few Hindi datasets are available, authors created Hindi fake news detection dataset. Experiments show that LSTM gives best performance with 92.36% accuracy. Authors in [18] used a modified version of BERT from Monsoon-nlp/hindi-tpu-electra which is Hindi BERT. They used a dataset from the Jelwin13afc repository. They achieved 52% of accuracy using their approach. In [20], Dilip Kumar Sharma, presented a technique using ML algorithms to identify the fake contents. They used BoW and TF/IDF for feature extraction. Authors created a dataset for Hindi News. ML classifiers such as SVM, Gaussian NB, KNN, RF, and DT were tested over the dataset. SVM resulted in the best accuracy of 91% over all the models. Shivangi Singhal et al. [17] presented a multilingual fact checking dataset containing fact checking data in different Indian languages. Jathin Badam et al. in [21] presented a technique to detect fake contents. They used TF/IDF and Word2Vec for feature extraction from news data. Authors used classifiers such as Decision Tree, Logistic Regression, Random Forest, Gradient Boost, and Recurrent Neural Network. Along with it pretrained Hindi BERT and Hindi TPU-Electra models were used. Experimental results show that the Hindi BERT and Hindi TPU-Electra have better results over all the models.

III. METHODOLOGY

First, we will discuss the datasets used for system implementation. Then we will discuss the algorithms that are used as individual learners and then the stacking ensemble approach used to implement the model.

A. Dataset

Main challenge in building a fake content identification model for Hindi news is the availability of datasets since very few datasets are publicly available. Three datasets were used for experimentation purposes. First dataset is from Kaggle the other from GitHub and the third dataset is formed by combining the first two datasets. First is the Hindi Fake news dataset from Kaggle Hindi fake news detection challenge. It contains 3864 fake news articles out of which 2824 news articles are unique and 3242 true news articles out of which 2647 articles are unique. Second dataset is collected from Jelwin13afc/Fake News Detection [18] containing 1250 fake articles and 760 true news articles. Third dataset is created by merging these two datasets.

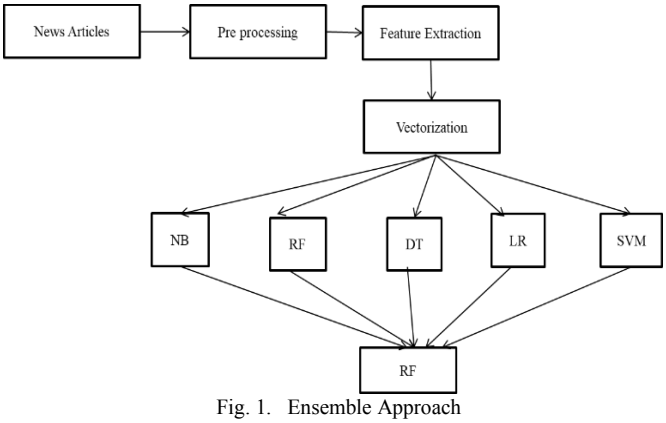

Fig. 1. Ensemble Approach

B. Machine Learning Algorithms

First the pre-processing is done on the news data by removing unwanted characters. It is followed by two well-known preprocessing operations for removing the stop words and the stemming operations. Stop words are the common words that do not carry much meaning in a sentence hence they are removed. Whereas stemming is a process of reducing a word to its root form. TF/IDF is used after preprocessing for feature extraction. TF is a Term Frequency and IDF is known as Inverse Document Frequency. It is used to find the importance of the term related to the document.TF is a frequency of the occurrence of the term in a document as shown in equation (1). IDF is used to find how uncommon the word is to the document as shown in equation (2). TfidfVectorizer is used to vectorize the text data by using TF-IDF. TF/IDF is calculated by using equation (3).

$$tf(t,d) = \frac{count(t,d)}{N} \tag{1}$$

$$idf = \log\left(\frac{N}{(df(t)+1)}\right) \tag{2}$$

$$tf.idf(t,d) = tf(t,d) * \log\left(\frac{N}{(df(t)+1)}\right) \tag{3}$$

Here *t* is the term and *d* is the document, *df(t)* is the occurrence of term *t* in document *d*, and *N* is the count of terms in the corpus.

Data is divided into train and test sets and different machine learning algorithms are applied as individual learners. Then these learners are used in an ensemble approach. Two experimental setups are used. In the first setup equal number of fake and real news articles are used from dataset and in another method dataset of original size is used. Following are the algorithms used to implement the model.

Support Vector Machine (SVM): Linear SVM is popularly used in binary classification problems where an optimal hyperplane separates data in two classes with maximum margin. Margin is the distance of the closest data point from the hyperplane. Hyperplane is a linear function of the input features. SVM finds the optimal coefficients which can maximize the margin between two classes. Due to this SVM algorithm has good, and generalized performance on unseen data.

Logistic Regression (LR): It is another popular algorithm used frequently for 2 class classification problems. It determines the probability of an instance to the class. LR is a maximum likelihood estimation algorithm which finds the likelihood of a data with respect to a certain class.

Decision Tree (DT): It is a hierarchical structure where non leaf nodes of the tree represent a test condition; the outcome of the test condition is represented on the branch and the class label is decided by the leaf node. Tree is built by selecting an attribute at each node based on information gain or Gini index.

Once a tree is built then it can be used for classification on unseen data items. However, it may not perform well if datasets are very small.

Random Forest (RF): is an ensemble learner which builds multiple decision trees. The attribute selection for the non-leaf node is in done based on random subset of input features. At each node the splitting attribute is selected randomly from a set of input features. Predictions are made by considering the predictions of individual learners in order to improve accuracy and reduce overfitting.

Naïve Bayes (NB): Naïve Bayes can be effectively used with text classification. It is the probabilistic classifier based on Bayes' theorem with the underlying assumption about the independence of the features. Multinomial Naïve Bayes is used on word vectors formed by TF/IDF vectorization. It uses maximum likelihood principle is used to perform the classification.

XGBoost (XGB): XGB stands for extreme gradient boosting. This algorithm is the optimized gradient boosting classifier which is very efficient for classification. It uses parallel tree boosting to do the classification. It is a scalable algorithm and can handle large datasets. It is one of the popular algorithms used in Kaggle competitions. In boosting the models are added sequentially and they are trained using iterations until it rectifies errors incurred in previous model.

TABLE I. EXPERIMENTAL SETUP-1

	DS I				DS II				DS III			
Algorithm	P	R	F	A	P	R	F	A	P	R	F	A
RF	**86.3**	88.4	87.3	87.8	88.8	88.3	88.5	88.2	90.7	94	92.4	92.5
LSVM	85	90	87.5	87.7	84.2	89.3	86.7	85.8	88.8	93	90.8	91
DT	83.1	81.8	82.5	83.4	83.5	77.2	80.2	80.3	86.7	88.8	87.7	88
LR	84.6	91.9	88.1	88.1	86.1	91.9	88.9	88.2	85.7	93.2	89.3	89.2
NB	90.4	85	97.5	88.2	87.7	86.6	87.2	86.8	90	86.9	88.4	88.9
XGB	84.8	92.7	88.6	88.3	88	79.6	83.6	83.9	83.6	91.2	87.4	87.2
Ensemble	88.6	91.7	90.2	90.2	91.6	89.8	90.7	90.5	90.7	94.6	92.6	92.7

TABLE II. EXPERIMENTAL SETUP-2

	DS I				DS II				DS III			
Algorithm	P	R	F	A	P	R	F	A	P	R	F	A
RF	87.3	86	86.7	87.7	89.1	81.8	85.3	88.3	93.6	93.4	93.5	94.2
LSVM	84.8	88	86.4	87.1	87.5	80.4	83.8	87.1	88.1	92.1	90.1	91
DT	80.8	75.6	78.1	80.3	82.7	75.6	79	83.3	88.7	87.7	88.2	89.6
LR	84.7	91.3	87.9	88.3	88.7	75.1	81.3	85.7	85.6	93.6	89.4	90.2
NB	89.4	86.2	87.8	88.6	87	87	87	89.6	92.5	86.3	89.3	90.5
XGB	84.3	91	87.5	87.9	83.9	83.9	83.9	87.1	87.7	91.8	89.7	90.4
Ensemble	87.1	92.9	89.9	90.3	89	90.7	89.8	91.8	93.7	95.2	94.4	94.8

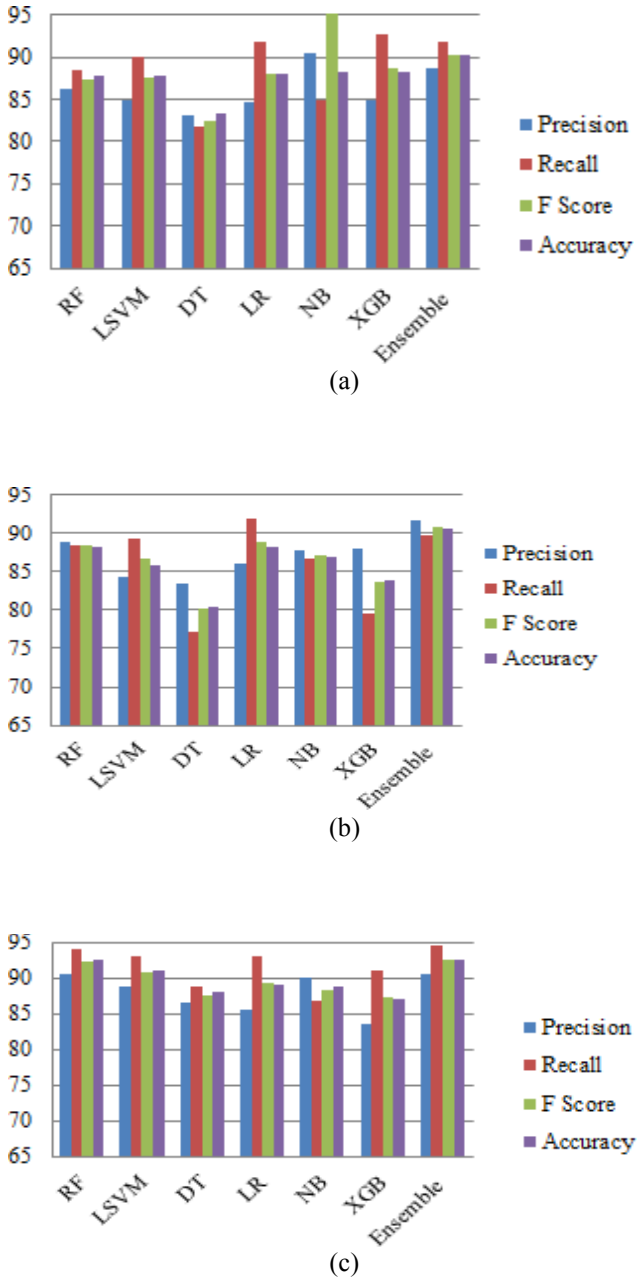

Fig. 2. Results for experiment setup I on DSI (a), DSII (b), and DSIII(c)

Ensemble Method: All the above learners are used individually over the datasets and their performance is observed. Then these individual learners are combined into ensemble approach to predict the authenticity of the news. Result analysis show that there is significant improvements in the performance when ensemble approach is used. Ensemble approach can be implemented using bagging, boosting, and stacking approach. In this approach we have used stacking ensemble method to combine individual learners. RF, LSVM, DT, LR and Multinomial NB are used as base classifiers in layer 1. The stacking approach is used to combine results from these learners. In layer 2 RF is used to implement the stacking ensemble method as shown in figure 1.

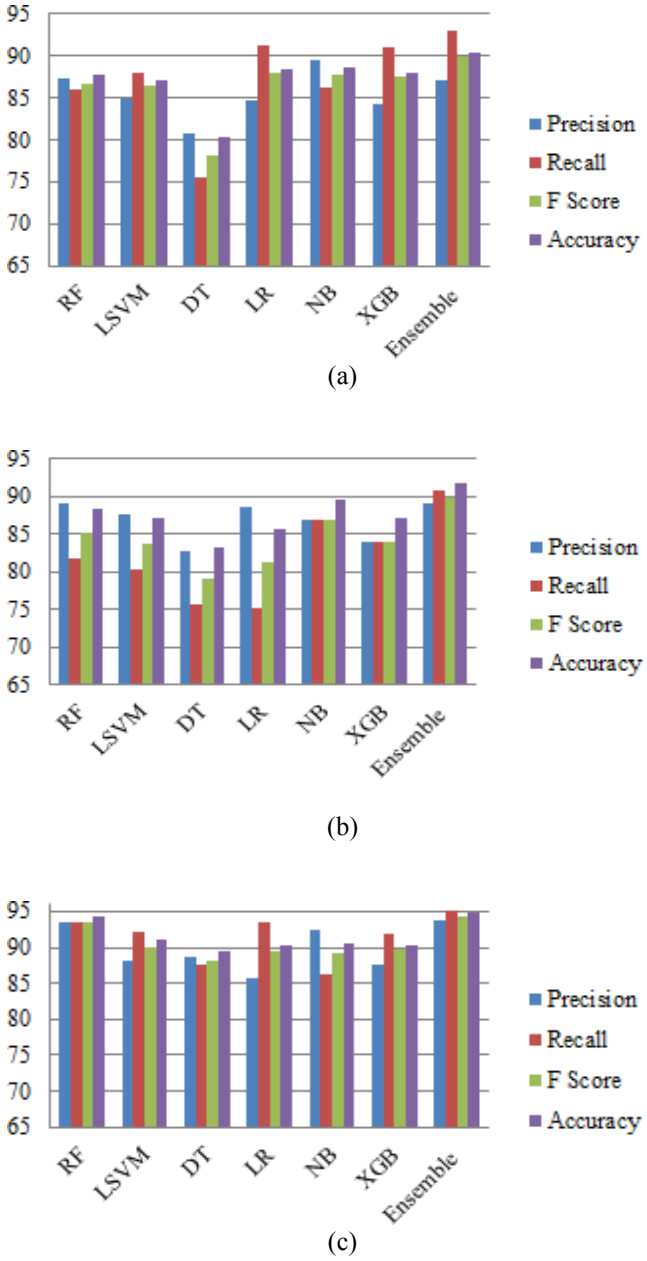

Fig. 3. Results for experiment setup II on DSI (a), DSII (b), and DSIII (c)

IV. Results and Discussions

All the experiments were performed on Google Colab. 80% of news articles are used for the training purpose and 20% for the testing purpose. Performance of the system is measured using evaluation metrics Precision (P), Recall (R) F1_Score (F) and accuracy(A). The results for both the experimental setup are represented in table 1 and 2. Accuracy is not always a good measure to decide the effectiveness of the classifier. For example, in case of unbalanced data we need other measures. In such cases Precision, Recall and F Score can be used. In an ideal classifier we want the value of precision and recall being 1 so that False Positives and False Negatives will be zero. So F1 score, harmonic mean of Precision and Recall, is calculated

based on both Precision and Recall. Following are the formulas (4-7) to calculate Precision, Recall, F1 Score and Accuracy.

$$\text{Precision}=\frac{TP}{TP+FP} \tag{4}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{5}$$

$$F1\_\text{Score}=2*\frac{\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{6}$$

$$\text{Accuracy}=\frac{TP+TN}{TN+FP+TP+FN} \tag{7}$$

Here TP represents True Positives TN is used for True Negatives, FP for False Positives and FN for False Negatives. Results in Fig. 2 and 3 show that linear models have better performance as compared to nonlinear models. It also reflects that ensemble approach has overall good performance as compared to individual learners. In the case of individual learners Random Forest has resulted in better performance. Results also show the significant improvement in experiment setup II where the number of real and fake news are equal.

V. Conclusion and Future Work

We have used various machine learning algorithms for model building for Hind news data. Experiments were carried out on datasets collected from Kaggle, GitHub and third by combining these two datasets. Evaluation metrics Precision, Recall, F1 Score and Accuracy are used to test the performance of the system. Results show that RF has overall performed well over all individual learning algorithms. Ensemble approach using stacking method is implemented to improve the performance. It resulted in significant improvement over the results. In future work we would like to work over large datasets. Since very few Hindi news data for authenticity detection are available publicly, we used datasets that are available publicly. Right now, the framework is implemented over small available dataset. In future we will extend it to a large dataset and evaluate the model over it. Second, we will also build and evaluate model using deep learning algorithms.

References

[1] Zapan Barua, Sajib Barua, Najma Kabir,, Mingze Li, Salma Aktar, "Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation", Elsevier, Volume 8, DOI - http://dx.doi.org/10.1016/j.pdisas.2020.100119 (2020)

[2] Tao Jiang, Jian Ping Li, Amin Ul Haq, Abdus Saboor,Amjad Ali , "A Novel Stacking Approach for Accurate Detection of Fake News", IEEE Access, vol. 9, DOI - 10.1109/ACCESS.2021.3056079 (2021)

[3] Reis JC, Correia A, Murai F, Veloso A, Benevenuto F, "Supervised learning for fake news detection", IEEE Intelligent Systems, IEEE Computer Society DOI- 10.1109/MIS.2019.2899143 (2019)

[4] Ahmad I, Yousaf M, Yousaf S, Ahmad MO, " Fake news detection using machine learning ensemble methods", Hindawi Complexity, DOI - 10.1155/2020/8885861 (2020)

[5] Kai Shu, Xinyi Zhouy Suhang Wangz, Reza Zafaraniy, and Huan Li, "The Role of User Profiles for Fake News Detection", arXiv:1904.13355v1 (2019)

[6] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques,", ISDDC 2017, Proceedings 1, Springer (2017)

[7] Ahlem Drif, Zineb Ferhat Hamida, Silvia Giordano, "Fake News Detection Method Based on Text-Features", The Ninth International Conferenve on Advances in Information Mining aand Management, ISBN-: 978-1-61208-731-3 (2019)

[8] Nida Aslam, Irfan Ullah Khan, Farah Salem Alotaibi , Lama Abdulaziz Aldaej, and Asma Khaled Aldubaikil, "Fake Detect: A Deep Learning Ensemble Model for Fake News Detection", Hindawi, DOI - 10.1155/2021/5557784 (2021)

[9] Jamal Abdul Nasir, Osama Subhani Khan, Iraklis Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach", International Journal of Information Management Data Insights, (2020)

[10] Monther Aldwairi, Ali Alwahedi, "Detecting Fake News in Social Media Networks", The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks, Elsevier DOI - 10.1016/j.procs.2018.10.171 (2018)

[11] X. Zhang and A. A. Ghorbani, ''An overview of online fake news: Characterization, detection, and discussion'', Information Processing and. Management, vol. 57, no. 2, Elsevier DOI - doi.org/10.1016/j.ipm.2019.03.004 (2020)

[12] Z Khanam, B N Alwasel, H Sirafi and M Rashid, "Fake News Detection Using Machine Learning Approaches", IOP Conference Series: Materials Science and Engineering, DOI-:10.1088/1757-899X/1099/1/012040(2020)

[13] Rohit Kumar Kaliyar, Anurag Goswami, Pratik Narang , " FakeBERT: Fake news detection in social media with a BERT-based deep learning approach" Multimedia Tools and Applications DOI - 10.1007/s11042-020-10183-2. (2021)

[14] Shaina Raza, Chen Ding, "Fake news detection based on news content and social contexts: a transformer-based approach", International Journal of Data Science and Analytics DOI- 10.1007/s41060-021- 00302-z. (2022)

[15] Yang S, Shu K, Wang S, Gu R, Wu F, Liu H , " Unsupervised fake news detection on social media: a generative approach". In:Proceedings of the AAAI conference on artificial intelligence, vol 33, no. 01, (2019).

[16] Debanjana Kar, Mohit Bhardwaj, Suranjana Samanta, Amar Prakash Azad, "No Rumours Please! A Multi-Indic-Lingual Approach for COVID Fake-Tweet Detection", arXiv:2010.06906v1 (2020)

[17] Shivangi Singhal, Rajiv Ratn Shah, Ponnurangam Kumaraguru, "Factorization of Fact-Checks for Low Resource Indian Languages", arXiv:2102.11276v1 (2021)

[18] Enjoy Maity, Ankush Tomar, Ruhi Peter, "Fake News Detection System: In Hindi Data set using BERT", International Research Journal of Modernization in Engineering Technology and Science Volume: 04 Issue: 05 (2022)

[19] Sudhanshu Kumar, Thoudam Doren Singh, "Fake news detection on Hindi news dataset", Global Transitions Proceedings DOI- 10.1016/j.gltp.2022.03.014 (2022)

[20] Dilip Kumar Sharma, Sonal Garg, "Machine Learning Methods to identify Hindi Fake News within social-media", 12th ICCCNT, IEEE (2021)

[21] Jathin Badam, Akash Bonagiri, KVLN Raju, and Dipanjan Chakraborty, "Aletheia: A Fake News Detection System for Hindi", 5th Joint International Conference on Data Science & Management of Data, ACM DOI- 10.1145/3493700.3493736 (2022)

[22] Xinyi Zhou, Atishay Jain, Vir V. Phoha, and Reza Zafarani.. "Fake News Early Detection: A Theory-driven Model". Digit.Threat.: Res. Pract. 1, 2, Article 12 (June 2020), DOI - https://doi.org/10.1145/3377478. (2020)