Fake News Detection on Social Media using a Stacking Model

G S Ramesh[1], K.H.S.Supriya[2], P.Akash[3] , A. Rukmananda Reddy[4] , V. Tejaswini[5]

[1] Assistant Professor, Department of Computer Science and Engineering,
VNR VJIET, Hyderabad (Telangana), India.
[2, 3, 4, 5] Student, Department of Computer Science and Engineering,
VNR VJIET, Hyderabad (Telangana), India.

(Corresponding author: G.S. Ramesh)

Abstract- Fake news has been there since before the advent of the Internet. It can be defined as "fictitious articles deliberately fabricated to deceive readers". Some news may have a nugget of truth, but lack any contextualizing details. They may not include any verifiable facts or sources. Some stories may include basic verifiable facts, but are written using language that is deliberately inflammatory, leaves out pertinent details or only presents one viewpoint. In this paper we try to understand these nuances in fake news by it's language. Initially the data set was preprocessed, the article's content was pre-processed using techniques like Lemmatization and TF-IDF. The data was then classified using a Stacking Classifier, using Naïve Bayes, SVM, Decision Tree as base models and Logistic Regression as meta model.

Index Terms- Fake News, Classification, Social Media, SVM, Naïve Bayes, NLP

I. INTRODUCTION

In modern era where the internet is ubiquitous, everyone relies on various online resources for news. Along with the increase in the use of social media platforms like Facebook, Twitter etc. News spread rapidly among millions of users within a very short span of time. Fake news encapsulates pieces of news that may be hoaxes and is generally spread through social media and other online media. Such news items may contain false and/or exaggerated claims, and may end up being made viral by algorithms, and users may end up in a filter bubble. The spread of fake news has far-reaching consequences like the creation of biased opinions to swaying election outcomes for the benefit of certain candidates.

With the current usage of social media platforms, consumers are creating and sharing more information than ever before, some of which are misleading with no relevance to reality. Social media and news outlets publish fake news to increase readership or as part of psychological manipulation. In general, the goal is profiting through click baits to lure users and entice curiosity with flashy headlines or designs to click links to increase advertisements revenues.

A regular examination of the news articles would give us an understanding of what articles are fake and which ones are real. This classification can help in identifying fraud accounts spreading the news and the respective authorities can be alerted to take necessary action.

II. LITERATURE SURVEY

Efeosasere Okoro et al., (2018) in their paper titled "A Hybrid Approach To Fake News Detection On Social Media" stated that there have been a wide range of solutions developed to help humans distinguish between fake and real news however, the solutions rely on either a machine-based approach or a human-based approach to detection. Research in the fields of computer science, artificial intelligence and psychology research has shown the limitations in both approaches. Based on these research findings, this paper proposes a hybrid model for detecting fake news on social media using a combination of both the human-based and machine-based detection approaches.

Ray Oshikawa et al., (2018) in their paper titled "A Survey on Natural Language Processing for Fake News Detection" stated that given the massive amount of Web content, automatic fake news detection is a practical NLP problem useful to all online content providers, in order to reduce the human time and effort to detect and prevent the spread of fake news. In their paper, we describe the challenges involved in fake news detection and they describe related tasks. We systematically review and compare the task formulations, datasets and NLP solutions that have been developed for this task, and also discussed the potentials and limitations of them. Based on their insights, they outline promising research directions, including more fine-grained, detailed, fair, and practical detection models.

MontherAldwairi et al., (2018) in their paper titled "Detecting Fake News in Social Media Networks" stated that clickbaits lure users and entice curiosity with flashy headlines or designs to click links to increase advertisements revenues. This exposition analyzes the prevalence of fake news in light of the advances in communication made possible by the emergence of social networking sites. The purpose of their work is to come up with a solution that can be utilized by users to detect and filter out sites containing false and misleading information. They used simple and carefully selected features of the title and post to accurately identify fake posts. The experimental results showed a 99.4% accuracy using logistic classifier.

Kai Shu et al., (2017) in their paper titled "Fake News Detection on Social Media: A Data Mining Perspective" started that first,

news is intentionally written to mislead readers to believe false information, which makes it difficult and nontrivial to detect based on news content; therefore, we need to include auxiliary information, such as user social engagements on social media, to help make a determination. Second, exploiting this auxiliary information is challenging in and of itself as users' social engagements with fake news produce data that is big, incomplete, unstructured, and noisy. In this survey, they present a comprehensive review of detecting fake news on social media, including fake news characterizations on psychology and social theories, existing algorithms from a data mining perspective, evaluation metrics and representative datasets. They also discuss related research areas, open problems, and future research directions for fake news detection on social media.

M. F. Mridha et al., (2021) in their paper titled "A Comprehensive Review on Fake News Detection With Deep Learning" stated that a protuberant issue of the present time is that, organizations from different domains are struggling to obtain effective solutions for detecting online-based fake news. It is quite thought-provoking to distinguish fake information on the internet as it is often written to deceive users. Compared with many machine learning techniques, deep learning-based techniques are capable of detecting fake news more accurately. Previous review papers were based on data mining and machine learning techniques, scarcely exploring the deep learning techniques for fake news detection. However, emerging deep learning-based approaches such as Attention, Generative Adversarial Networks, and Bidirectional Encoder Representations for Transformers are absent from previous surveys. This study attempts to investigate advanced and state-of-the-art fake news detection mechanisms pensively The prominent evaluation metrics in fake news detection are also discussed. Nevertheless, they suggest further recommendations to improve fake news detection mechanisms in future research directions.

Iftikhar Ahmad et al., (2020) in their paper titled "Fake News Detection Using Machine Learning Ensemble Methods" stated that even an expert in a particular domain has to explore multiple aspects before giving a verdict on the truthfulness of an article. In this work, we propose to use machine learning ensemble approach for automated classification of news articles. Their study explores different textual properties that can be used to distinguish fake contents from real. By using those properties, they train a combination of different machine learning algorithms using various ensemble methods and evaluate their performance on 4 real world datasets. Experimental evaluation confirms the superior performance of our proposed ensemble learner approach in comparison to individual learners.

Syed Ishfaq et al., in their paper titled "Fake News Detection Using Machine Learning approaches: A systematic Review" state that the easy dissemination of information by way of sharing has added to exponential growth of its falsification. The credibility of social media networks is also at stake where the spreading of fake information is prevalent. Thus, it has become a research challenge to automatically check the information viz a viz its source, content and publisher for categorizing it as false or true. Machine learning has played a vital role in classification of the information although with some limitations. This paper reviews various Machine learning approaches in detection of fake and fabricated news. The limitation of such and approaches and improvisation by way of implementing deep learning is also reviewed.

GihwanKim et al., (2020) in their paper titled "Effective fake news detection using graph and summarization techniques" stated that there is an increasing need to accurately detect fake news to prevent such damage. In this paper, they propose a novel method that uses graph and summarization techniques for fake news detection. Their proposed method represents the relationship of all sentences in a graph structure to accurately understand the context information of the document. Accordingly, the relationship between sentences in the graph is calculated as a score through the attention mechanism. Then, the summarization technique is used to reflect the sentence subject information in the graph update process. Our proposed method shows better performance than Karimi's and BERT based models by approximately 10.34%p and 3.72%p, respectively.

Jiawei Zhang et al., (2019) in their paper titled "FAKEDETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network" stated that investigating the principles, methodologies and algorithms for detecting fake news articles, creators and subjects from online social networks and evaluating the corresponding performance. This paper addresses the challenges introduced by the unknown characteristics of fake news and diverse connections among news articles, creators and subjects. This paper introduces a novel automatic fake news credibility inference model, namely FAKEDETECTOR. Based on a set of explicit and latent features extracted from the textual information, FAKEDETECTOR builds a deep diffusive network model to learn the representations of news articles, creators and subjects simultaneously. Extensive experiments have been done on a real-world fake news dataset to compare FAKEDETECTOR with several state-of-the-art models, and the experimental results have demonstrated the effectiveness of the proposed model.

Mohamed K. Elhadad et al., in their paper titled "Fake News Detection on Social Media: A Systematic Survey" stated that these days there are instabilities in many societies in the world, either because of political, economic, and other societal issues. The advance in mobile technology has enabled social media to play a vital role in organizing activities in favour or against certain parties or countries. Many researchers see the need to develop automated systems that are capable of detecting and tracking fake news on social media. In this paper, they introduce a systematic survey on the process of fake news detection on social media. The types of data and the categories of features used in the detection model, as well as benchmark datasets are discussed.

Chun-ming Lai at al., in their paper titled "Fake News Classification Based on Content Level Features" stated that due to the openness and easy accessibility of online social media (OSM), anyone can easily contribute a simple paragraph of text to express their opinion on an article that they have seen. Without access control mechanisms, it has been reported that there are many suspicious messages and accounts spreading across multiple platforms. Accordingly, identifying and labeling fake news is a demanding problem due to the massive amount of heterogeneous content. In essence, the functions of machine learning (ML) and natural language processing (NLP) are to enhance, speed up, and automate the analytical process. Therefore, this unstructured text can be transformed into meaningful data and insights. In this paper, the combination of ML and NLP are implemented to classify fake news based on an open, large and labeled corpus on Twitter. They compare several state-of-the-art ML and neural

network models based on content-only features. To enhance classification performance, before the training process, the term frequency-inverse document frequency (TF-IDF) features were applied in ML training, while word embedding was utilized in neural network training. By implementing ML and NLP methods, all the traditional models have greater than 85% accuracy. All the neural network models have greater than 90% accuracy. From the experiments,they found that the neural network models outperform the traditional ML models by, on average, approximately 6% precision, with all neural network models reaching up to 90% accuracy.

Nasir et al., in their paper titled "Fake news detection: A hybrid CNN-RNN based deep learning approach" stated that the amplified the old problem of fake news, which became a major concern nowadays due to the negative impact it brings to the communities. In order to tackle the rise and spreading of fake news, automatic detection techniques have been researched building on artificial intelligence and machine learning. The recent achievements of deep learning techniques in complex natural language processing tasks, make them a promising solution for fake news detection too. This work proposes a novel hybrid deep learning model that combines convolutional and recurrent neural networks for fake news classification. The model was successfully validated on two fake news datasets (ISO and FA-KES), achieving detection results that are significantly better than other non-hybrid baseline methods. Further experiments on the generalization of the proposed model across different datasets, had promising results.

Isabel Bezzaoui et al., in their paper titled "Distinguishing Between Truth and Fake Using Explainable AI to Understand and Combat Online Disinformation" stated that since social media users and even larger platform operators are currently unready to clearly identify disinformation, new techniques for detecting online disinformation are urgently needed. In this paper, they present DeFaktS, an Information Systems research project, which takes a comprehensive approach to both researching and combating online disinformation. The project develops a data pipeline in which (i) messages are extracted in large quantities from suspicious social media groups and messenger groups with the help of annotators. Based on this corpus, a Machine Learning-based System (ii) is trained that can recognize factors and stylistic devices characteristic of disinformation, which will be used for (iii) an explainable artificial intelligence that informs users in a simple and comprehensible way about the occurrence of disinformation. Furthermore, in this paper an interdisciplinary multi-level research approach focusing on media literacy and trust in explainable artificial intelligence is suggested in order to operationalize research on combating disinformation.

Dorota Marquardt in their paper titled "Linguistic Indicators in the Identification of Fake News" stated that the issue of fake news identification was approached from the corpus linguistics and discursive studies perspective. The texts of both actual and fake news have been analysed in search of dependences that would permit the increase of the ability to determine the probability of the given news being real or fake, taking into account the discursive characteristics of the particular texts.

III. PROPOSED METHODOLOGY

Due to the complexity of fake news detection in social media, it is evident that a feasible method must contain several aspects to accurately tackle the issue. This is why the proposed method is a combination of Naïve Bayes classifier, Support Vector Machines, Logistic Regression and Lemmatization. This method is a combination between Machine Learning algorithms that subdivide into supervised learning techniques, and natural language processing methods. Although each of these algorithms can be separately used to classify and detect fake news, in order to increase the accuracy they have been combined into an integrated algorithm as a method for fake news detection. As a first step the data will be preprocessed and any NULL valued columns were dropped.

A. Lemmatization

Before applying this step the news articles are broken down to tokens. The NLP technique find the relationships among the word, Lemmatization is the process where we take individual tokens from a sentence and we try to reduce them to their base form. The process that makes this possible is having a vocabulary and performing morphological analysis to remove inflectional endings. The output of the lemmatization process is the lemma or the base form of the word. Lemmatization was preferred over stemming to take into account the sentence context. After performing lemmatization stop words like the, a, is were removed since they do not add value to the data.

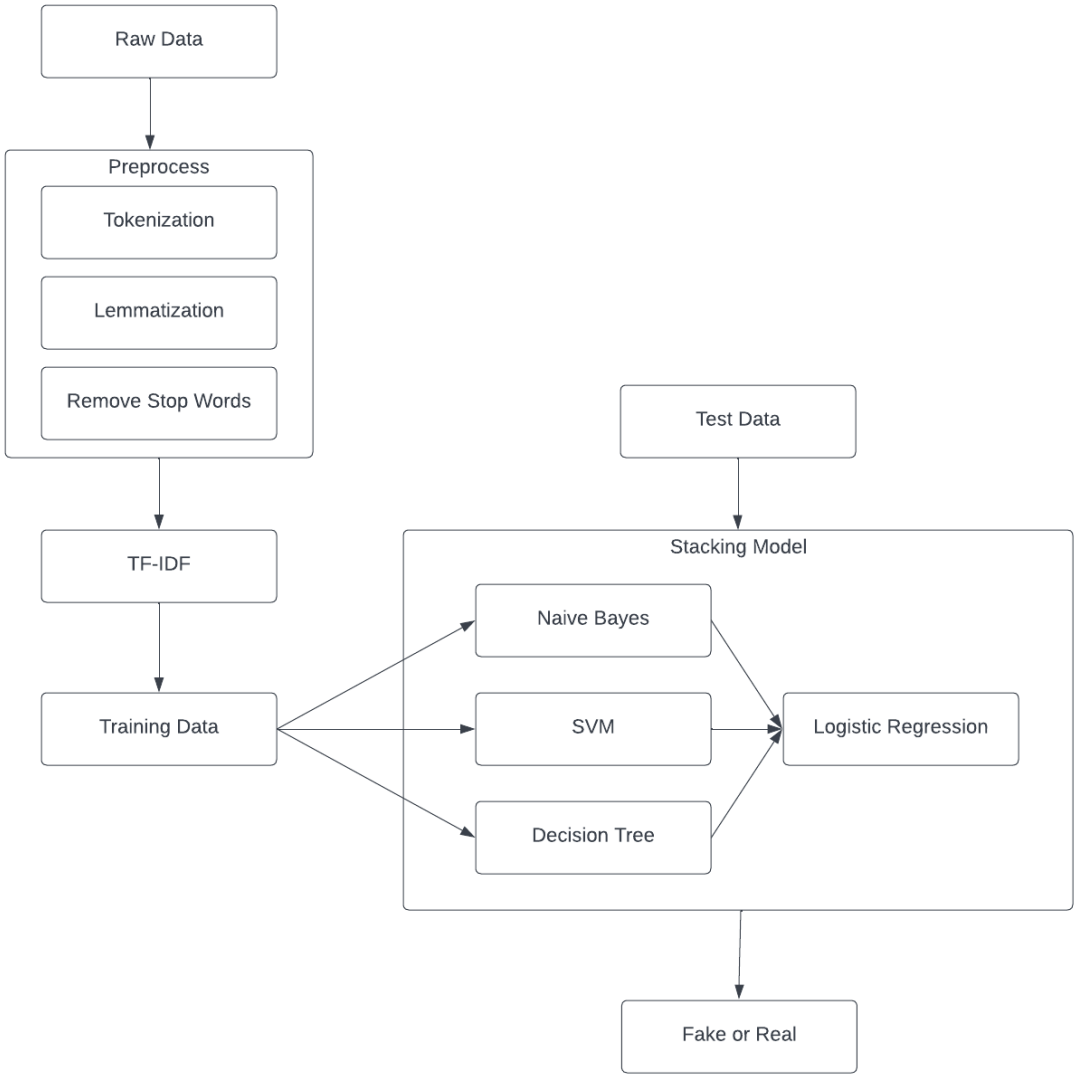
B. TF-IDF

TF-IDF is the importance of a term is inversely related to its frequency across documents.TF gives us information on how often a term appears in a document and IDF gives us information about the relative rarity of a term in the collection of documents. Since machine learning algorithms deal with numerical data and not textual, we use this method to convert text to numbers. Each token is mapped to a number between 0-1 which indicates its importance in the document.

C. Classification Models

The SVM and Naïve Bayes classifier and Decision Tree tend to "rival" each other due to the fact they are supervised learning algorithms that are efficient at classifying data. Three techniques are moderately accurate at categorizing fake news in experiments, which is why this proposed method focuses on combining SVM and Naïve Bayes classifier and Decision Tree to get even more accurate results. The biggest drawback of Naïve Bayes classifier is that it deems all features of a document, or whichever textual format being used, to be independent even though most of the time that is not the situation. This is a problem due to lowered accuracy and the fact that relationships are not being learned if everything is assumed to be unrelated. Combining these with supervised learning algorithms with the concept of Ensemble Learning will yield us more accuracy than the algorithms working individually. For this purpose we use a stacking model, which takes uses SVM, Naïve Bayes and Decision Tree as base models and Logistic Regression as the Meta model. Using the outputs of the base models, the meta model is trained and hence its accuracy is always greater than or equal to the individual models.

Fig. 1: The Proposed Methodology



IV. RESULTS

Explorative data analysis was performed on the Fake News Classification data set. It was analyzed by means of Histogram, Box plots, Bar plots, and Word Clouds. These plots helped us understand the nature of the data at hand.

One of the biggest difference between fake and real news sources is the title. Specifically, we find, fake news titles are longer than real news titles and contain simpler words in both length and technicality. Fake titles also used more all capitalized words, significantly more proper nouns, but fewer nouns overall, and fewer stop-words (examples: the, and, a, an). In addition, we find that fake titles use significantly more analytical words, and in our data set, fake titles use significantly more verb phrases and significantly more past tense words. Inorder exploit these properties, we measured the lengths of the titles.

The mean length, min length and max length of the news article titles in terms of number of characters are as shown in Table. 1. It is indicating that fake articles tend to have titles that are either too long or too short compared to articles that contain real or legitimate news.

Table 1: Article Title Lengths

Label	Mean Length	Min Length	Max Length
All	74.7023	3	456
Real	79.6142	22	175
Fake	68.2798	3	456

Upon analyzing the title lengths via a bar plot as show in Fig. 1, we observe that most articles tend to have a length between 50-100 characters. We observe that there seem to be outliers for the length of the articles, so we have plotted a box plot as Fig. 2, which plots the article title lengths. Outliers tend to lie more on the greater side to the median in this particular data set.

Fig. 1: Article Title Lengths vs Frequency

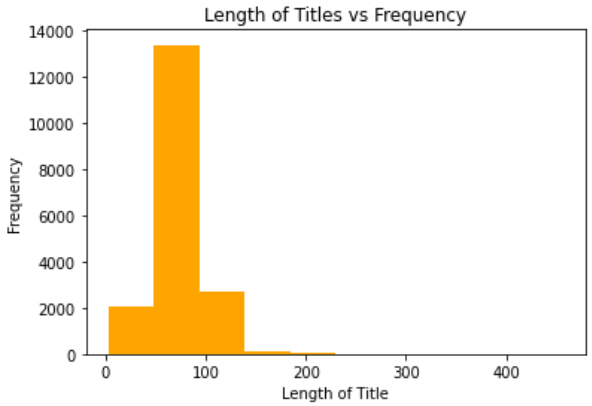
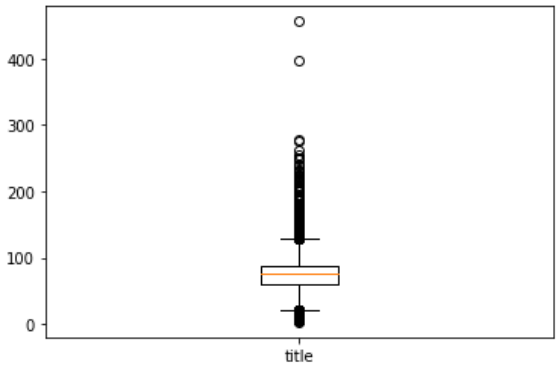


Fig. 2: Box plot for article lengths



In order to understand what words could be probable indicators to fake news and real news, a Word Cloud was plotted. A word cloud is a collection, or cluster, of words depicted in different sizes. The bigger and bolder the word appears, the more often it's mentioned within a given text and the more important it is.

Fig. 3(a). Words like Trump, Hillary, VIDEO, New, Comment, Russia, Election are some of the frequent words among fake articles. We also see words like "will" suggesting most fake article titles are in future tense. Whereas words like New York, Times, Breitbart, Donald, Trump are some of the frequent words among real articles Fig.3(b).

Fig. 3(a): Fake News Title Word Cloud

Fig. 3(b): Real News Title Word Cloud

Fig.4 is an example of how a news articles text may be converted before it is vectorized. We observe that stop words like "the", "in", "a" and many more are missing in the processed text. Similarly, some words are reduced to their root form, like remains is converted to remain. All words are not reduced since some of them need to remain in their current form to maintain language consistency.

Fig. 4: Text Transformation after processing

Raw Text	Processed Text
Ever get the feeling your life circles the roundab out rather than heads in a straight line toward th e intended destination? [Hillary Clinton remains t he big woman on campus in leafy, liberal Wellesley , Massachusetts. Everywhere else votes her most li kely to don her inauguration dress for the remaind er of her days the way Miss Havisham forever wore that wedding dress. Speaking of Great Expectation s, Hillary Rodham overflowed with them 48 years ag o when she first addressed a Wellesley graduating class. The president of the college	Ever get feeling life circle roundabout rather hea d straight line toward intend destination Hillary Clinton remain big woman campus leafy liberal Well esley Massachusetts Everywhere else vote likely in auguration dress remainder day way Miss Havisham f orever wear wed dress Speaking Great Expectations Hillary Rodham overflow 48 year ago first address Wellesley graduating class The president college

Fig. 5: Authors with most fake articles

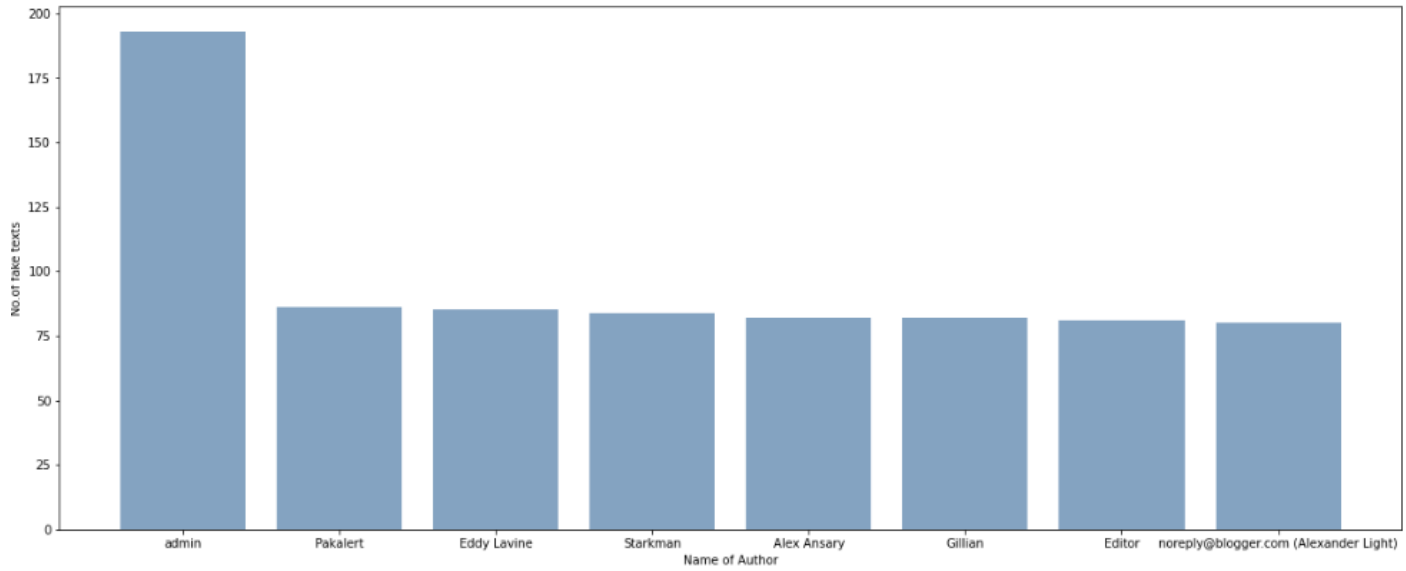


Fig. 6: Articles with word Trump

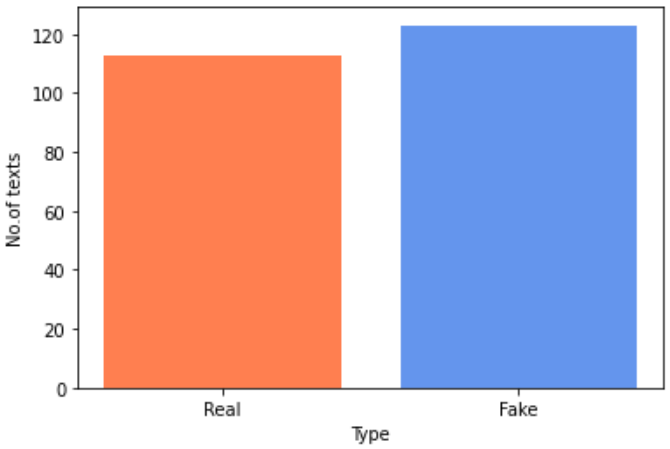


Fig.5 indicates that most of the fake articles comes from authors with names, admin, Pakalert, Eddy Lavine, Starkman, Alex Ansary, Gillian, Editor, noreaply@blogger.com (Alexander Light). Since Trump is one of the most frequent words, on plotting a bar graph as shown in Fig.6, we observe that there are equal number of fake and real texts with the word Trump. Hence for any article on Trump there may be a 50% chance of it being fake or real.

Table. 2: Accuracy Scores

DataType	Naïve Bayes	SVM	Decision Tree	Stacked Model
Train	92.5307	99.1562	96.5432	99.7734
Test	91.0682	96.9559	94.9325	97.8855

Table. 2 shows the classification accuracy scores for the different models. We observe that SVM has the most accurate classification whereas Decision Tree is the least accurate. Stacked Model which uses Naïve Bayes, SVM, Decision Tree as base models and Logistic Regression as the Meta model always has an accuracy greater than the individual models.

V. CONCLUSION

With the increasing popularity of social media, more and more people consume news from social media instead of traditional news media. However, social media has also been used to spread fake news, which has strong negative impacts on individual users and broader society. In this project, we explored the fake news problem by reviewing existing literature. The proposed methodology has two main components: processing + vectorization and usage of a stacking model. We have also explored the data set by creating necessary plots to analyze the fake news we are dealing with. We also further discussed the evaluation metrics of our model and why it shows increased accuracy.

VI. FUTURE SCOPE

There are promising future directions in fake news detection research. Further work includes trying out other machine learning models in the stacked model to check which combination works better, tune the model parameters to better adjust to the data set. Developing an easy to access GUI would make the model more accessible.

ACKNOWLEDGMENT

Special thanks to our team guide, Mr. G.S.Ramesh, for all the technical support and guidance which led to the completion of this project with a good result in the end.

REFERENCES

[1] Efeosasere Okoro, Benjamin Abara, Umagba Alex, and Anyalewa Alan Ajonye (2018) “A hybrid approach to fake news detection on social media” Nigerian Journal of Technology 37(2) DO - 10.4314/njt.v37i2.22

[2] Ray Oshikawa, Jing Qian, and William Yang Wang (2018) “A Survey on Natural Language Processing for Fake News Detection” arXiv:1811.00770 [cs.CL]

[3] MontherAldwairi and AliAlwahedi (2018) “Detecting Fake News in Social Media Networks” Procedia Computer Science,Volume 141, 1877-0509

[4] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu (2017) “Fake News Detection on Social Media: A Data Mining Perspective” arXiv:1708.01967 [cs.SI]

[5] M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar and M. S. Rahman, (2021) "A Comprehensive Review on Fake News Detection With Deep Learning," in IEEE Access, vol. 9, pp. 156151-156170

[6] Ahmad Iftikhar, Yousaf Muhammad, Yousaf Suhail, and Ahmad Muhammad Ovais (2020) “Fake News Detection Using Machine Learning Ensemble Methods”, Hindawi, 1076-2787

[7] S. I. Manzoor, J. Singla and Nikita, (2019) "Fake News Detection Using Machine Learning approaches: A systematic Review," 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 230-234, doi: 10.1109/ICOEI.2019.8862770.

[8] Gihwan Kim and Youngjoong Ko (2020) “Effective fake news detection using graph and summarization techniques” Pattern Recognition Letters, Volume 151, 0167-8655

[9] Jiawei Zhang , Bowen Dong , and Philip S. Yu (2019) “FAKEDETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network”, arXiv:1805.08751v2 [cs.SI]

[10] Mohamed K. Elhadad, Kin Fun Li, and Fayez Gebali (2019) “Fake News Detection on Social Media: A Systematic Survey”, IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM), 2325-0445

[11] Lai, C.-M.; Chen, M.-H.;Kristiani, E.; Verma, V.K.; Yang, C.-T. (2022) “Fake News Classification Based onContent Level Features”. Appl. Sci.,12, 1116. https://doi.org/10.3390/

[12] Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-rnnbased deep learning approach. International Journal of Information Management DataInsights,1(1), Article 100007

[13] Fegert, Jonas AU, Weinhardt, Christof PY (2022) “Distinguishing Between Truth and Fake Using Explainable AI to Understand and Combat Online Disinformation”. The 16th International Conference on Digital SocietyAt: Porto, Portugal. DO - 10.5445/IR/1000148097 ER

[14] Marquardt, Dorota PY (2019) “Linguistic Indicators in the Identification of Fake News” VL - 3 DO - 10.17951/ms.2019.3.95-114 JO - Mediatization Studies ER