Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

FactBuddy: Browser Website for Fake News Detection Using Machine Learning Algorithm

Avijit Datta
Department of Computer Science & Engineering
Netaji Subhash Engineering College
Kolkata, India

Arpan Banerjee
Department of Computer Science & Engineering
Netaji Subhash Engineering College
Kolkata, India

Aritra Naskar
Department of Computer Science & Engineering
Netaji Subhash Engineering College
Kolkata, India

Shilpi Bose
Department of Computer Science & Engineering
Netaji Subhash Engineering College
Kolkata, India

Riya Majumdar
Department of Computer Science & Engineering
Netaji Subhash Engineering College
Kolkata, India

Chandra Das
Department of Computer Science & Engineering
Netaji Subhash Engineering College
Kolkata, India

***Abstract*— Social media has developed into a haven for fake news, which is essentially misleading information disseminated to the public. This issue is not new; dishonest publications have always attempted to distort the truth to benefit themselves. Social media's rapid news dissemination and ease of use are the main reasons for its appeal. But the same ease of use has also made it simple for false information to proliferate like wildfire, harming both people and society as a whole. The widespread dissemination of false information poses a substantial obstacle: the identification and effective prevention of inaccurate data. Sensational stories are often written by those disseminating phony news to draw attention and sway public opinion. This has affected both online and traditional journalism by making people doubt the validity of social media as a news source. In this paper, a simple, fast and easy to implement method is presented based on Logistic Regression. The system is validated and tested with both known and unknown data-sets and the results obtained are in good agreement with the actual one and within permissible limits.**

***Keywords*— Fake News Prediction, Machine Learning, Logistic Regression, Accuracy Score**

I. INTRODUCTION

In the current modern age where information dissemination occurs at an unprecedented pace, the prevalence of fake news poses a significant challenge to the credibility of online content. The research in this direction introduces a novel solution in the form of a Smart System for fake news detection, leveraging the capabilities of machine learning [1,2]. As users navigate the vast landscape of the internet, this system serves as a vigilant guardian, employing advanced algorithms to discern between authentic and deceptive information. By harnessing the power of machine learning, the extension analyzes content intricacies, linguistic nuances, and contextual cues to provide real-time assessments of the veracity of online information. Through the seamless integration of technology into the everyday web-browsing experience, this research not only focuses to the ongoing efforts to tackle misinformation but also offers a user-friendly toolkit for individuals seeking to navigate the digital realm with increased discernment and trust. False data is provided by a large number of websites. They make an effort to spread intentional misinformation, lies, and publicity as if it were actual news.

FactBuddy automates following two tasks:

(1) Selecting the claims which are fact check worthy

(2) Verify the genuineness of claims based on the evidence found online.

FactBuddy initiates users to do fact checking on the Website. If users are not sure on which they should claim to fact check, FactBuddy can automatically filter fact check worthy claims for them. FactBuddy can automatically search for online evidence via web search engines like google to verify claims. In addition to classifying the claims, FactBuddy also provides evidence snippets highlighting the importance of both words and sentences relevant for classifying the claim using attention weights.

II. LITERATURE SURVEY

A lot of research has been done on how to automatically spot fake news and misleading posts. Fake news can be found in many ways, such as by looking for things like robots that spread false information or clickbaits that spread rumours. There are a lot of clickbaits on social media sites like Facebook that make people more likely to share and like posts, which spreads false information. A lot of work has been put into finding false information.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

To overcome the scenario, Sahoo [1] proposed an automatic technique for Fake News Detection for the chrome browser environment using this detection on Facebook. Moreover, the technique contains multiple features linked with a Facebook Account in addition to another news content features to analyze the characteristics of the account across deep learning. In another repository of fake news data, created by Shu [2] there are two complete data with different features, information and social contexts facilitates fake news related research. This extensive exposition of FakeNewsNet [2] presents a scholarly examination of two datasets from diverse perspectives and deliberates on the merits of FakeNewsNet in relation to prospective implementations in the study of false news on social media. SAF/S exhibits superior performance in both accuracy and F1 score. SAF/A yields an equivalent outcome to SAF/S, achieving an accuracy of 66.7%. This suggests that user engagements, in conjunction with news articles from the PolitiFact [3]dataset, may contribute to the detection of false news. Concurrently, the selection strategy may be implemented to reduce disturbance in the data acquisition procedure by utilizing web search results.

Daun [4] in the paper, approached this challenge by explaining two features, one of them is linguistic and another one is sentiment feature from the user's tweet feed as well as retrieving the presence of hashtags, emojis and political bias in their tweets. Subsequently, these characteristics were employed to classify operators into those who disseminated false information or not. This proposed methodology achieved an accuracy rate of 72% among the top four results obtained by systems performing the task in English. In applications employing a variety of classification algorithms and combining the various representations, however, not every representational combination improved accuracy. When combined with other representations, NER is not compatible with SVMs or ANNs. Additionally, this limit required multiple increases. Probably a significant number of features (416,834) necessitate this.

In their study, Kumar [5] introduced a CNN+bidirectional LSTM ensembled network to collect new instances, including those from PolitiFact, in order to construct several pieces of information that can distinguish between true and deceptive news, as well as to compare with various state-of-the-art methods. Among the numerous state-of-the-art approaches are Long Short-Term Memories (LSTMs), convolutional neural networks (CNNs), attention mechanisms, and ensemble methods. Multiple datasets distinguishing between authentic and fabricated news stories are compiled from 1356 news instances collected by this research from various users via Twitter and media sources including PolitiFact. In contrast, Ko et al. accomplished a detection rate of 85% while addressing the issue of false news identification. The study's findings indicate that the CNN+bidirectional LSTM ensembled network with attention mechanism demonstrated the highest accuracy at 88.78%. Achieving a maximal accuracy of 88.78%, the ensemble of CNN and bidirectional LSTM with focus mechanism was the outcome. Although gratifying, the outcomes failed to inspire confidence. With respect to the alternatives under investigation, the CNN architecture exhibited the least accuracy. Compared to a basic CNN architecture, the performance of the LSTM and bidirectional LSTM structures was considerably enhanced. By integrating more intricate models into our methodology, we increased our desire for greater precision.

False news identification was proposed by Nasir [6] using a hybrid deep learning architecture that combines recurrent neural and convolutional networks. This model was effectively certified on two datasets containing false news (ISO and FA-KES). Its detection results are significantly superior to those of other non-hybrid foundation techniques. The statistical significance of the results was assessed using a paired t-test; the experiments were replicated five times (utilizing 5-fold cross-validation, which involved a division of 80%–20%); and the reported accuracy was accompanied by 95% confidence intervals. ISOT is selected for training purposes due to its substantial size and limited room for development, as numerous models achieve classification accuracy levels exceeding the 0.9 threshold. Furthermore, intricate neural network architectures should not be incorporated into the research.

Choudhary [7] introduced a linguistic model that aims to discern the characteristics of the content and subsequently generate language-driven features. This linguistic prototype extracts syntactic, sentimental, grammatical, and legibility-specific news features. The language-driven paradigm necessitates a strategy for handling handcrafted feature issues and requires considerable time to address the challenge of dimensionality issues. Consequently, in order to improve the accuracy of false news detection, a neutrality-based continuous learning model is implemented. The outcomes are compiled in order to validate the significance of the extracted features from the linguistic model. Ultimately, the integrated linguistic feature-driven model demonstrates an average accuracy of 86% in the identification and classification of fraudulent messages. However, the model lacks comprehensive features and parameters that govern its performance. Analyze the model that detects false news based on latent semantic features, and investigate different iterations of convolution neural networks that are designed to detect fake news images.

There is also a browser extension named BRENDA [8] which facilitated a extension support for browsers but the work is limited to some browsers.

The leading industry players are currently prioritizing shielding themselves from unfounded rumors, instead emphasizing the importance of genuine news and verified articles. Information extraction techniques heavily depend on Machine Learning & Natural Language Processing (NLP). Collaboration among classifiers, models, and analytical algorithms is essential in verifying the authenticity of news articles. From the previously mentioned studies, the gap between researches motivates to study the hybrid and fast method for fake news detection. From the existing methods, the combination of LSTM and CNN has shown impressing results. However, till now LSTMs have been used for embeddings of words and CNN for doing the job of final classification. The analysis of the pertinent literature concludes that numerous real-time catastrophes have been significantly influenced by false news. A variety of datasets have been utilised to test machine learning techniques, whereas deep learning techniques for false news detection and related tasks have yet to be exhaustively evaluated. Table 1 provides a comparison of the most recent and advanced techniques.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

TABLE 1

Reference	Goals	Pros	Cons
Sahoo [1]	A method for detecting false news automatically in the Chrome environment that employs deep learning and machine learning classifiers to identify phony news on Facebook.	Both user profile and news content features are analyzed.	
Shu [2]	Repository for disinformation data FakeNewsNet comprises two extensive data sets that encompass a wide range of characteristics, including news content, social context, and spatiotemporal information.	The research community would benefit from FakeNewsNet's investigations into a variety of topics, including (early) fake news detection, the evolution of fake news, fake news mitigation, and malicious account detection.	This link displays the metadata of only 5000 users as a result of space constraints.
Daun [4]	Multiple machine learning and deep learning algorithms are combined to detect false news patterns with the utmost degree of precision.	Highest-performing model achieved 70% accuracy on the testing set on TIRA.	Not all representational combinations resulted in improved accuracy. When combined with other representations, NER is not compatible with SVMs or ANNs.
Kumar [5]	Conduct a comparative analysis of CNN, LSTM, the bidirectional LSTM model, a CNN+LSTM ensembled network, and a bidirectional LSTM+LSTM ensembles model in order to collect new instances, such as PolitiFact, and construct a diverse set of information to discern between authentic and fabricated news.	Apply this research to counteract misinformation and mitigate the far-reaching consequences of false news.	This study primarily examined the sentiments expressed in news articles, neglecting to consistently assess the credibility of the news sources as a result of resource constraints.
Nasir [6]	An innovative deep learning model is suggested, which integrates convolutional and recurrent neural networks to classify false news.	Almost 100% accuracy on the ISOT Dataset	Overfitted models reveal a great deal of complexity and analyze a great deal more data than is likely required to reach a conclusion.
Chaudhary [7]	A neural-based sequential learning language model is suggested for the identification of bogus news.	Analyse the significance of the extracted feature sets as well; of all the retrieved features, readability is thought to be the least frequently employed.	Not many features or parameters for model performance
Brenda [8]	A browser extension to detect fake news over various news sites	Suggests more news article related to the news	Doesn't work in all browsers, limited features, Time consuming

Disadvantages of Existing System: Existing works and systems contains some drawbacks i.e. Information was not clear and not able to extract the correct information in the bulk of news, defamation is among the drawbacks of fake news. False Perception is also present in the existing system. Fake news may lead to Social Unrest.

III. PROPOSED MODEL

Online Passive-aggressive algorithms [9] is easy to implement and coefficients of the features can be easily interpreted to understand their impact on the probability of a news being categorized as fake or real. It is faster to train on large datasets

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

and it gives probabilistic output. Moreover Online Passive-aggressive algorithm i.e Logistic Regression prevents overfitting issues, thus enhancing the model's performance and robustness. In this work, the logistic Regression Model is therefore selected for its implementation. However, the proposed model can also be implemented using other machine learning algorithms.

News aggregator [10] sites allow users to conveniently access different news and updates from various sources in single place. They gather the information, categorize it, and present it in a well-organized manner for easier use. There are several popular websites that offer semi-structured news data, such as Google News, Feedly, and News360. There are RSS Aggregator plugins available to simplify things. Aggregators enhance the quality and accuracy of news. The primary objective of any news aggregator is to gather data. One approach involves regularly monitoring RSS Feeds, extracting articles from different news sites, and collecting information. Commonly used methods to find related articles are keyword-based approaches. Once all the processes are complete, they display relevant and up-to-date news on the page.

News Authenticator employs a series of steps to determine the veracity of news articles. It will analyze news provided by us by comparing it with content from different websites and News articles. If the News is found on a reputable News website, it indicates that the news is true. However, if there is no mention of such news in the past few days, it suggests that the news does not exist. This can be beneficial in combating misinformation. In today's digital age, the rapid spread of misinformation has become a pressing issue, largely due to the prevalence of social media and the internet. The news authenticator is a valuable tool for determining the authenticity of news articles.

Advantages of Proposed System: Information was very clear and understandable**.** It gives accurate predictions which is very clear to the user. FactBuddy has user friendly and faster time compatibility because of Logistic Regression, Decision Tree Classifier, Gradient Bound Classifier & Random Forest Classifier. Whereas existing system produces maximum 90% accuracy rate, but the Algorithms used in proposed model produces 100% accuracy rate. The results are shown in the Implementation & Result section. Existing systems and researches uses dataset containing 5000 data, but in FactBuddy, more than 20,000 data containing dataset is used to train and test the model. Also the proposed model has user-friendly interface which is very simple and easy to understand.

IV. METHODOLOGY

4.1 Workflow: The Project contains three main segments of the methodology. First, the core Machine learning model. Secondly the web interface, and then the common platform that brings the model and the interface together. There are two parts to the ML Model Building. Two types of model used in the project.

For first, Online – Passive aggressive classifiers are used. The steps begins with the collection of a labeled dataset comprising thousands of news articles, categorized as real or fake, alongside metadata such as text and class. The dataset which was used to test and train the model is simple. The dataset is downloaded from Kaggle. It contains the titles of the news, the body text and a label field, which, if the news is authentic, shows REAL and if inauthentic, shows FAKE. This dataset undergoes preprocessing to convert textual data into a numerical format suitable for analysis. Following preprocessing, the dataset is divided into training and test subsets for Logistic Regression, Decision Tree Classifier, Gradient Bound Classifier & Random Forest Classifier model training and evaluation, respectively. Passive-aggressive algorithms are a family of great learning algorithms. They are similar to Perceptron because it does not require a reading scale. However, unlike Perceptron, they include parameter correction. Passive is used when the prediction is correct and there is no change in the model. But if there is any kind of change in the model, that is if the prediction is not correct then the aggressive part is called, which changes the model accordingly. The aggressive part of the model changes the model according to its wish on the backend. The model is built through the train and test of the dataset, by ensuring that the training is done for 75% of the dataset and testing is done in the rest of the 25% of the dataset.

In the second part, LSTM is used. In this part, the first setp is loading the Data which is same as the passive-aggressive one. Then Data is loaded from a CSV file. This consists of the body of selected news articles. It then contains a label field that indicates whether the news is real or fake. In this code block, we scan the CSV and clean the titles to filter out stop words and punctuation. Later, The tokenizer is used to assign indices to words, and filter out infrequent words. This allows us to generate sequences for our training and testing data. The Embedding Matrix is applied later. The important point to be considered in the project is “Stemming”. Basically, stemming is the process of reducing a word into its root word. Stemming removes the prefix and suffixes from words because they don’t give such meanings in the dataset. An embedding matrix is used to extract the semantic information from the words in each title. The final step is building the model and finding out the accuracy via confusion matrix. The model is created using an Embedding layer, LSTM, Dropout, and Dense layers. The data is going to run on 20 epochs. It is observed that the LSTM model is vastly inaccurate in predicting the authenticity of the news. So to show the output by running it through the Passive-aggressive classifier model is good.

Post-training, model performance is assessed using test data, with the trained model then utilized for predicting the authenticity of new news articles, thereby completing the project workflow.

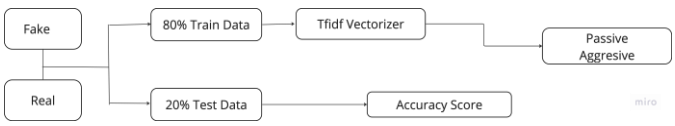

Figure 1. Passive-aggressive model.

4.1.1 Logistic Regression: Logistic regression [11] is a statistical method used for binary classification problems, where the goal is to predict the probability of a binary outcome (whether the news is fake or not) based on one or more input variables (also known as independent or predictor variables).

It models the relationship between the input variables and the binary output variable using sigmoid function, which transforms the output of the model to a probability between 0 and 1.

Sigmoid Function**:** The sigmoid function [11] is one of the most commonly used activation functions. Its mathematical

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

representation is $\sigma(x) = 1/1+e^{-x}$

In this sigmoid function we have x. And we replace the x by the equation x = w.X + b which is same as the straight-line function i.e., y = mx + c.

Here, X = input features
Y = Prediction Probability
w = weights
b = biases

When x is small, the value of the sigmoid function is close to 0. When x is large, the value of sigmoid is close to 1. The sigmoid function transforms the continuous real number into a range of (0,1), so that the input value of the next layer is within a fixed range and the weight is more stable.

4.1.2 Decision Tree Classifier: A decision tree classifier is a type of algorithm used for classification tasks that builds a tree-like model by recursively partitioning the data based on the input variables and their values to predict the class label of new data.

4.1.3 Gradient Bound Classifier: A gradient bound classifier is a type of machine learning algorithm used for classification tasks. It updates a set of weights to minimize the difference between predicted and actual outcomes. The "gradient bound" term means it has a constraint that prevents the weights from becoming too large. This helps prevent numerical instability or overfitting.

4.1.4 Random Forest Classifier: A random forest classifier is a type of machine learning algorithm used for classification tasks. It combines many decision trees to make a prediction. Each tree looks at a random subset of the data and features, which helps to prevent overfitting. The final prediction is based on the most commonly predicted class by all the trees. Random forest classifiers are accurate and can handle large and complex datasets.

4.2 Useful feature extraction: Feature extraction addresses the challenge of identifying the most concise and informative collection of features. When dealing with classification and regression issues, the most prevalent and easiest way to describe data is by creating feature vectors. The feature extraction approach frequently used is determined on the size of the data table. As data storage becomes more efficient, the size of data tables is growing. The key to experimental study is in extracting impactful elements from text and avoiding unnecessary data processing.

Text information may exist in several sizes and formats. Structured data is a crucial factor to consider while extracting characteristics from text. Typically, unprocessed raw data in text format is transformed into structured data. Information extraction refers to the use of techniques to get accurate and reliable information. Similarly, in this study, the process of extracting valuable characteristics from the actual news material is a difficult challenge due to the ability of fake news propagators to make their information seem authentic.

4.2.1. Term frequency and inverse document frequency (TF-IDF): In this study, the strategy used term frequency and inverse document frequency (TF-IDF) to detect the valuable characteristics of news material. Prior research on the identification of false news may be categorized into two distinct groups: unimodal and multimodal. In order to acquire a broader and more inclusive joint representation, a minimax game is established between the event discriminator and the feature extractors.

4.3 System Architecture:

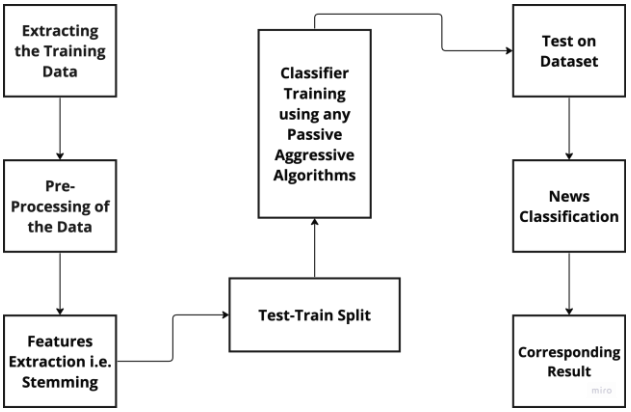

Figure 2: Flow-chart - Classifier Training

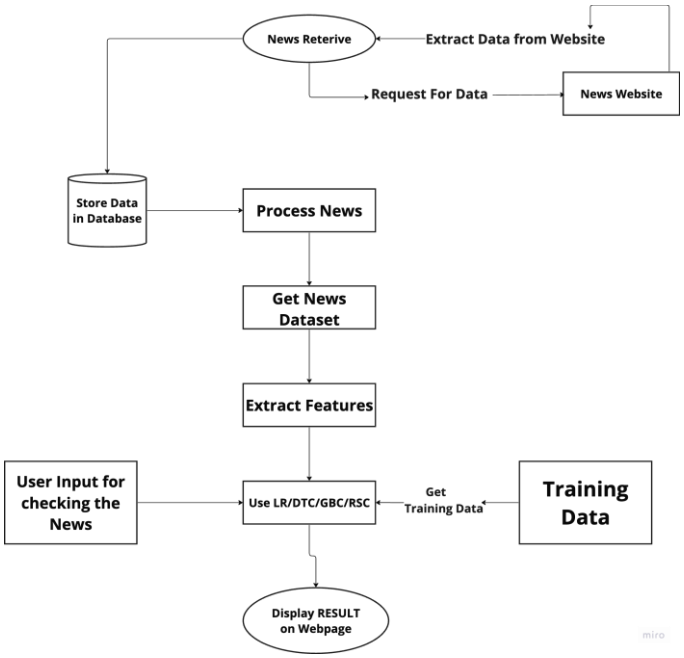

Figure 3: Flow-chart – The Proposed Model

4.4 System Design FactBuddy is a browser-based web application designed to detect fake news using machine learning algorithms. The system architecture comprises various components that interact seamlessly to provide a robust and user-friendly experience. **Here's a System Design overview:**

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

The system for fake news detection begins with data collection and preprocessing. A labeled dataset is collected containing news articles marked as real or fake. The text data is then cleaned by removing stop words and punctuation, followed by tokenization. The cleaned text is converted into numerical representations using TF-IDF vectorization. For machine learning models, a Passive-Aggressive Classifier is employed, which involves loading the preprocessed data, applying TF-IDF vectorization, and training the classifier on the training data. Additionally, an LSTM model is used, which includes tokenizing and padding sequences for uniform input size, creating an embedding matrix for the words, and building and training the LSTM model with the processed data. The system features a web interface with a frontend developed using HTML and CSS to create a user-friendly interface where users can input news articles and submit them for analysis. The backend utilizes Flask to handle requests and interface with the machine learning models, loading the trained models, and preprocessing user input to make predictions. Integration with Flask involves creating routes to handle user inputs and display results, using the Pickle module to serialize and deserialize the trained models, and displaying the prediction results on the web interface.

V. IMPLEMENTATION & RESULT

The results of the analysis of the datasets using the Logistic Regression has been depicted using the confusion matrix. The confusion matrix is automatically obtained by Python code using the cognitive learning library when running the algorithm code in Google Colab Platform. The confusion matrix so obtained from the analysis shows the False Positives and False Negatives as shown in fig. (4).

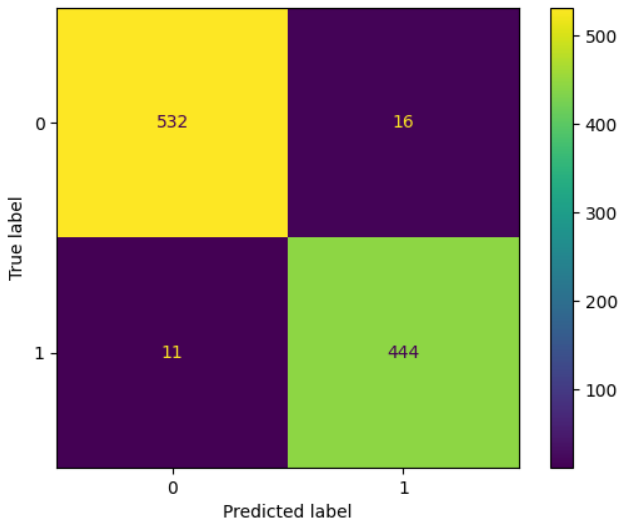

Figure 4: Confusion Matrix

The figure shows 11 False Positives and 16 False Negatives using the Logistic Regression algorithm in a balanced dataset with a TF-IDF vectorizer. The train data contains about 3000 number of data-sets which is nearly 75% of data from the data-set. The data is subsequently filtered and stemmed out using functions of the natural language processing library and other functions used for processing text data in an organized form.

TABLE 1

Evaluation Parameters	Obtained Scores
Correct Results	976 out of 1003
Accuracy Score	0.973
Precision Score	0.965
Recall Score	0.976
F1 Score	0.970

The evaluation parameters obtained from the different metrices library of Python are tabulated in Table- I. These shows that the results obtained from the test data are in good agreement with the actual value. The ROC curve shown in fig 5 clearly depicts the fact that the model is well working within permissible limit.

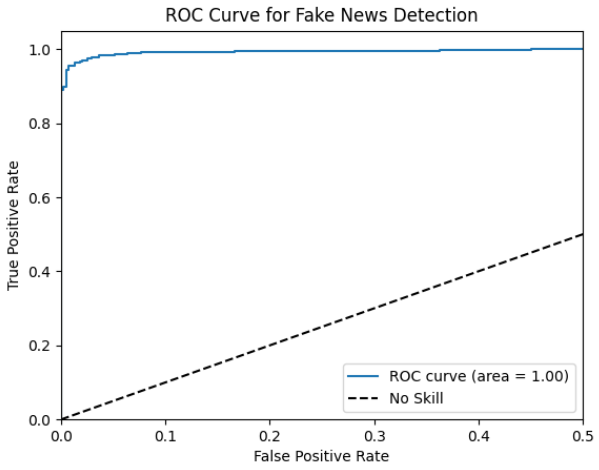

Figure 5: ROC Curve

The scheme is tested with the different commonly used Classification algorithms. The accuracy scores for all those cases are tabulated in Table – 2. Comparison among all these schemes shows that the Logistic Regression and the SVM gives better and nearly equal results. However, among these two, the Logistic Regression should be the preferred choice as it work relatively faster and is also easy to implement which clearly justifies the choice of model in this work.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 5, May 2024

TABLE 2

Classification Algorithm	Accuracy Score
Decision Tree	0.962
Random Forest	0.966
K Nearest Neighbour	0.894
Support Vector Machine	0.989
Logistic Regression	0.974
Naïve Bayes	0.841

VI. CONCLUSION

The creation of a machine learning system that can distinguish between authentic and fraudulent news articles was investigated in this study. This system was built on the robust binary classification method known as Logistic Regression. Nevertheless, news articles' unstructured content made things difficult because machine learning models need organized data to be analyzed.

The task of classifying news manually requires in-depth knowledge of the domain and expertise to identify anomalies in the text. The data used in work contains news articles from various domains to cover most of the news rather than specifically classifying political news. The primary aim of the research is to identify patterns in text that differentiate fake articles from true news. Here different textual features extracted from the articles and used the feature set as an input to the models. The learning models were trained and parameter-tuned to obtain optimal accuracy. FactBuddy has user friendly and faster time compatibility because of Logistic Regression, Decision Tree Classifier, Gradient Bound Classifier & Random Forest Classifier. Whereas existing system produces in the range of 70-90% accuracy rate, but the Algorithms used in proposed model produces 99% accuracy rate. The results are shown in the Implementation & Result section. Existing systems and researches uses dataset containing 5000 data, but in FactBuddy, more than 20,000 data containing dataset is used to train and test the model. Also the proposed model has user-friendly interface which is very simple and easy to understand.

REFERENCES

[1] S. R. Sahoo and B. B. Gupta, "Multiple features based approach for automatic fake news detection on social networks using deep learning," Appl Soft Comput, vol. 100, p. 106983, Mar. 2021, doi: 10.1016/J.ASOC.2020.106983.

[2] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media," https://home.liebertpub.com/big, vol. 8, no. 3, pp. 171–188, Jun. 2020, doi: 10.1089/BIG.2020.0062.

[3] "PolitiFact." Accessed: May 19, 2024. [Online]. Available: https://www.politifact.com/

[4] "RMIT at PAN-CLEF 2020: Profiling Fake News Spreaders on Twitter | Damiano Spina." Accessed: May 19, 2024. [Online]. Available: https://www.damianospina.com/publication/duan-2020-rmit/

[5] S. Kumar, R. Asthana, S. Upadhyay, N. Upreti, and M. Akbar, "Fake news detection using deep learning models: A novel approach," Transactions on Emerging Telecommunications Technologies, vol. 31, no. 2, Feb. 2020, doi: 10.1002/ETT.3767.

[6] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach," International Journal of Information Management Data Insights, vol. 1, no. 1, p. 100007, Apr. 2021, doi: 10.1016/J.JJIMEI.2020.100007.

[7] A. Choudhary and A. Arora, "Linguistic feature based learning model for fake news detection and classification," Expert Syst Appl, vol. 169, p. 114171, May 2021, doi: 10.1016/J.ESWA.2020.114171.

[8] B. Botnevik, E. Sakariassen, and V. Setty, "BRENDA: Browser Extension for Fake News Detection," in SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, Inc, Jul. 2020, pp. 2117–2120. doi: 10.1145/3397271.3401396.

[9] M. BNBabar Anushka Khandagale, P. More, Y. Malvade, and S. Kokane, "A SMART SYSTEM FOR FAKE NEWS DETECTION USING MACHINE LEARNING," JOIREM, 2023. [Online]. Available: www.joirem.com

[10] A. Jain, A. Shakya, H. Khatter, and A. K. Gupta, "A smart System for Fake News Detection Using Machine Learning," in IEEE International Conference on Issues and Challenges in Intelligent Computing Techniques, ICICT 2019, Institute of Electrical and Electronics Engineers Inc., Sep. 2019. doi: 10.1109/ICICT46931.2019.8977659.

[11] "Logistic Regression in Machine Learning - GeeksforGeeks." Accessed: May 18, 2024. [Online]. Available: https://www.geeksforgeeks.org/understanding-logistic-regression/