Machine Learning Algorithm for Fake News Detection

[1]Muntasir Khan, [2]K. Karthik, [3]M. Sushmasri, [4]V. Kalpana

[1]Student, [2]Student, [3]Lecturer, [4]Lecturer
[1]Computer Science Department, [2]Computer Science Department, [3]Electronics and Communication Department,
[4] Computer Science Department
[1]Government Institute of Electronics, Secunderabad, India, [2]Government Institute of Electronics, Secunderabad, India
[3]Government Institute of Electronics, Secunderabad, India. [4]Government Institute of Electronics, Secunderabad, India
[3]hiremath.sushmasri@gmail.com

Abstract— **The rise of digital media has led to an unprecedented spread of misinformation, making fake news a significant challenge in today's society. The Fake News Detection System is developed to address this issue by leveraging machine learning techniques to classify news articles as either legitimate or fake. The system follows a structured approach, starting with data collection, where real and fake news articles are gathered from reliable sources. The collected data undergoes pre-processing, including text cleaning, tokenization, and vectorization using the Term Frequency-Inverse Document Frequency (TF-IDF) method. A machine learning model, specifically Logistic Regression, is then trained on this processed data to classify news articles with high accuracy. Additionally, an interactive Graphical User Interface (GUI) is integrated into the system, allowing users to input news content and receive an instant classification result. This system provides an effective tool for combating misinformation by helping users identify unreliable news sources, contributing to a more informed digital environment.**

***Key Words*—Python, Packages, Machine Learning, Supervised Learning, TF-IDF.**

I. INTRODUCTION

In the digital age, the rapid dissemination of news through social media and online platforms has made it increasingly difficult to distinguish between credible information and misleading content. Fake news, which refers to false or misleading information presented as legitimate news, has the potential to influence public opinion, spread misinformation, and cause social and political unrest. The development of automated fake news detection systems has become essential in mitigating the impact of fake news by identifying unreliable sources and misleading content.

This project focuses on the automatic classification of news articles into real and fake categories using Natural Language Processing (NLP) and machine learning techniques. The system is designed to analyze textual content and extract meaningful features for classification. It employs TF-IDF vectorization, a widely used text representation technique, to transform textual data into numerical form, making it suitable for machine learning algorithms. The classification model is built using Logistic Regression, a supervised learning algorithm known for its efficiency and high accuracy in binary classification tasks. The Fake News Detection System follows a structured pipeline, including data collection, preprocessing, model training, evaluation, and deployment. The trained model is integrated into a user-friendly GUI, allowing users to input news text and receive an instant prediction on whether the news is real or fake. By leveraging machine learning, this system provides an automated and efficient solution to detect misinformation, contributing to a more reliable and trustworthy digital media landscape.

II. LITERATURE REVIEW

In today's generations there is a significant challenge in our society due to the rapid growth in digital media regarding proclamation on false information which is well known as Fake News. This Fake News influence more on public in which there is more chances of changing opinion, manipulation in political outcomes and even my cause financial and social instability. In this regard detecting the Fake News became as critical issue and hence there is rapid enhancement in research area on Fake News detection. Therefore, many Mahine Learning Algorithms, Artificial Intelligence (AI), and Natural Language Processing (NLP), contribute to the development of algorithms capable of detecting fake news with high accuracy.

Many researchers came with different approaches on Fake News detection as the author Shu et al. [1] proposed a fake news detection model based on deep learning techniques that analyze textual and social context features. The author concluded that neural network models like Long Short-Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT) outperform traditional machine learning models like Support Vector Machine (SVM) and Naïve Bayes (NB) in detecting fake news due to their ability to capture deep semantic relationships in text.

The author Zhou et al. [2] conducted an extensive study on various machine learning algorithms, including Logistic Regression, Decision Trees, Naïve Bayes, and Random Forest, for detecting fake news. Their analysis showed that Random Forest achieved an accuracy of 87.2%, outperforming Naïve Bayes, which suffered from higher false positives due to its assumption of independent features. Another author Kumar et al. [3] used TF-IDF (Term Frequency-Inverse Document Frequency) and n-gram-based feature extraction techniques for text classification. They compared different algorithms and found that SVM provided the highest accuracy of 90.5% for text-based fake news detection. However, they noted that SVM struggles with highly imbalanced datasets where real news outnumbers fake news. And by using Machine Learning model there is an enhancement in accuracy.

III. IMPLEMENTATION

In this project, we propose different Supervised Machine Learning algorithms to detect fake news by analyzing textual content and metadata from news articles. Python has been used for building, training, and testing the machine learning models.

A. Introduction to Python

Python is a widely used general-purpose programming language that supports multiple paradigms, including object-oriented, imperative, functional, and procedural programming. It is an interpreted language, meaning it does not require compilation before execution. Python's extensive libraries make it particularly powerful for machine learning, data analysis, and web scraping.

A.1 Extensive Set of Packages

Python has an extensive and powerful set of packages ready to be used in various domains, including machine learning and data science. The key libraries used in this project include:

- **NumPy**: Used for numerical computations and handling large datasets efficiently.
- **pandas**: Used for data manipulation, preprocessing, and analysis.
- **scikit-learn**: Provides machine learning algorithms like logistic regression, SVM, and decision trees.
- **NLTK & BeautifulSoup**: Used for Natural Language Processing (NLP) and web scraping.
- **TF-IDF Vectorizer**: Converts textual data into numerical representations for ML models.
- **Matplotlib & Seaborn**: Used for data visualization and evaluation metrics.

B. Machine Learning for Fake News Detection

Fake news detection relies on Supervised Machine Learning models, which learn patterns from labeled datasets and classify news as real or fake. The main steps involved in building a Machine Learning model for fake news detection are:

- **Data Collection**: Gathering articles from various news sources.
- **Text Preprocessing**: Cleaning and tokenizing the text.
- **Feature Extraction**: Converting text into numerical features using methods like TF-IDF (Term Frequency-Inverse Document Frequency).
- **Model Training**: Using algorithms such as Logistic Regression, Support Vector Machine (SVM), and Random Forest.
- **Evaluation**: Measuring model performance using accuracy, precision, recall, F1-score, and confusion matrix.

B.1 Web Scraping for Data Collection

Web scraping is used to collect news articles from different websites. The Beautiful Soup and newspaper3k libraries are used to extract article text and metadata (title, source, publication date). The collected data is stored in a MySQL database for further analysis.

B.2 Text Preprocessing

Raw text data must be cleaned before being fed into the machine learning model. The following preprocessing steps are applied:

- Removing HTML tags, URLs, punctuation, and numbers.
- Converting text to lowercase.
- Tokenizing words and removing stopwords using NLTK.

B.3 Feature Extraction

For machine learning models to process text, it must be converted into numerical representations. TF-IDF Vectorization is used to assign importance scores to words based on their frequency in the dataset.

C. Installation and Execution

To install necessary libraries, in our computer the following are the packages used for Machine Learning. In this regard first install numerical python called numpy. Scikit learn (sklearn), natural language tool kit(nlkt) and pandas (pd). The packages as follows:

```
import numpy as np
import pandas as pd
import sklearn
import nltk
```

IV. PROPOSED WORK

A. Proposed System:

To address the challenges, we proposed a Machine Learning-based Fake News detection system using TF-IDF vectorization and Supervised Learning Algorithms. Our system automates fake news classification by improving efficiency and accuracy.

B. Supervised Learning Algorithm

Supervise learning is a type of Machine Learning where the model is trained on labelled data. In this the training of dataset consists of input and output pairs where the input data are considered as features and the output data are considered as labels. Basically, Supervised Learning algorithms are broadly classified into two types namely; Classification and Regression. The Classification is a categorical output where it meant for detecting fake news whereas the Regression is a continuous output used for predicting features. The most commonly used Supervised Learning algorithms are:

- Logistic Regression
- Support Vector Machine
- Random Forest

- Naïve Bayes

B.1 Logistic Regression

Logistic Regression is a Supervised Machine Leaning Algorithm is widely used for classification problems and the name suggests that Regression does not belongs to Regression type of Supervised Learning but extended to the multi-class classification. The major applications arise in Fake News Detection, Spam Detection, Medical diagnosis etc. It is more suitable for smaller and medium datasets.

B.2 TF-IDF

The Term Frequency-Inverse Document Frequency (TF-IDF) is a numerical statistic used in Natural Language Processing (NLP) to evaluate the importance of a word in a document relative to a collection of documents. It is commonly used in Text Mining, Fake News Detection, and Information Retrieval to transform text into numerical values for Machine Learning Models.

V. Model Training and Evaluation with Performance Metrics

The dataset is split into training (80%) and testing (20%) sets. The models are trained on the training set and evaluated using the test set.

A. Performance Metrics

The following metrics are used to evaluate model performance:

- **Accuracy**:

The accuracy is used to measure, predict or estimate the object from an image. İt is the ratio between number of true predictions to the total number of predictions is seen in "Eq. 1".

$$Accuracy = \frac{Number\ of\ true\ predictions}{Total\ number\ of\ predictions} \quad (1)$$

Here the number of true predictions is considered as sum of true positive and true negative predictions and total number of predictions are sum of all true positive & negative and false positive & negative is seen in "Eq. 2". The true positive and true negative are number of correct predictions of instances in positive and negative respectively and false positive and false negative are number of incorrect predictions of instances negative and positive respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

- **Precision**:

The precision value used to measure the ratio between True Positive (TP) to the sum of TP and False Positive (FP) as seen in "Eq. 3".

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

- **Recall**:

The recall value measures the ratio between True Positive (TP) to the sum of True Positive (TP) and Fasle Neagtive (FN) is seen in "Eq. 4".

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

- **F1-Score**:

It is a performance metric used for classification models that balances both Precision and Recall. This metric mainly useful in dealing with imbalanced datasets and basic formula is shown in below "Eq. 5". As F1-score closer to 1 indicates a better model.

$$F1 - Score\ = 2x\ \frac{Precision\ x\ Recall}{Precision + Recall} \quad (5)$$

- **Confusion Matrix**:

The Confusion Matrix is used for performance evaluation tool for classification models. It consists of a table which summarizes the number of correct and incorrect predictions made by the model and it majorly useful for analyzing Accuracy, Precision, Recall and F1-score metrics. The structure of a Confusion Matrix as follows:

Table 1 Confusion Matrix

	Predicted Positive	**Predicted Negative**
Actual Positive	TP (True Positive)	FN (False Negative)
Actual Negative	FP (False Positive)	TN (True Negative)

VI. RESULTS

In our project the complete design is proposed with Logistic Regression which is a Supervised Machine Algorithm and TF-IDF. The "Fig 5.1" represents the Machine Learning steps like data collection and training model which are displayed in Interface Window.

Figure 5.1 Interface Window

When the input data (here the news article) is fed in the interface window then it predicts the article whether it is true article or not fake article as seen in below figures. From the "Fig 5.2" it represents the detection when it is true article by showing confidence level whereas the "Fig 5.3" represents about the fake article.

Figure 5.2 True Prediction

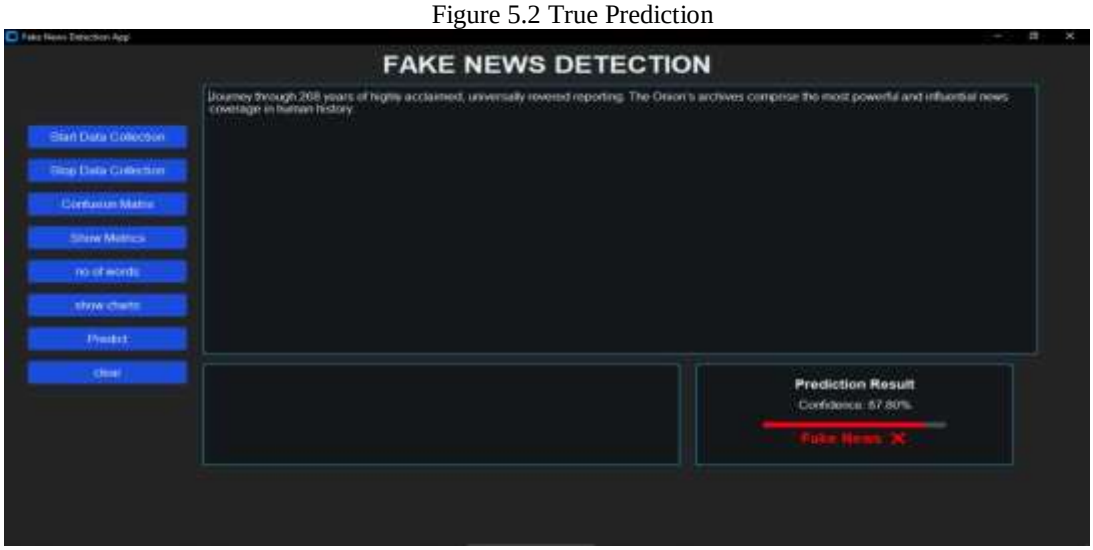

Figure 5.3 False Prediction

By overall model prediction we acquire 92% of accuracy and the following table 2 represents the performance metrics related to our model. And also, the "Fig 5.4" represents the graph plotted for below mentioned metrics.

Table 2 Performance metrics

Algorithm	Evaluation Metrics			
	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.92	0.88	1.00	0.94

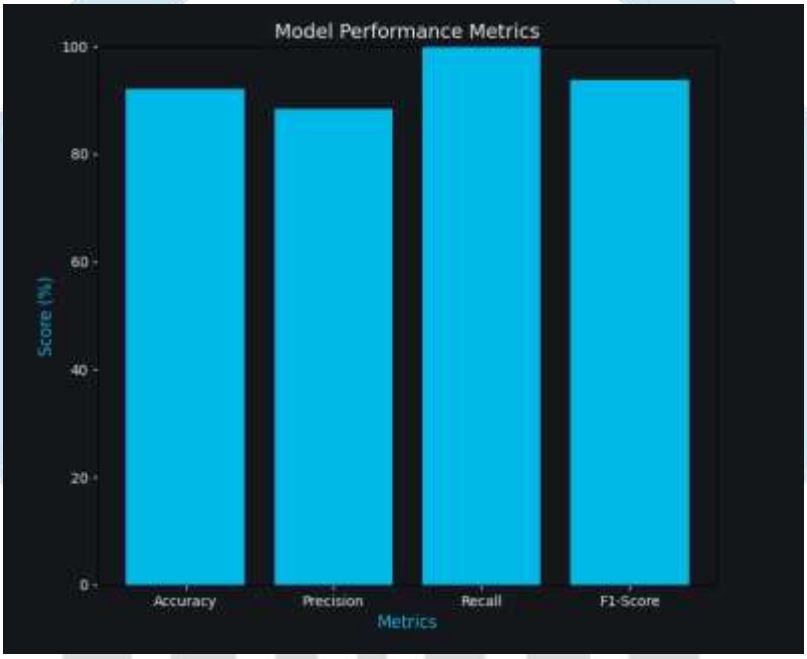

Figure 5.4 Graph of Performance Metrics

The following image "Fig 5.5" shows about the confusion matrix taken among TF-IDF and Logistic Regression. Based on this we can know the actual content of the information which is being predicted.

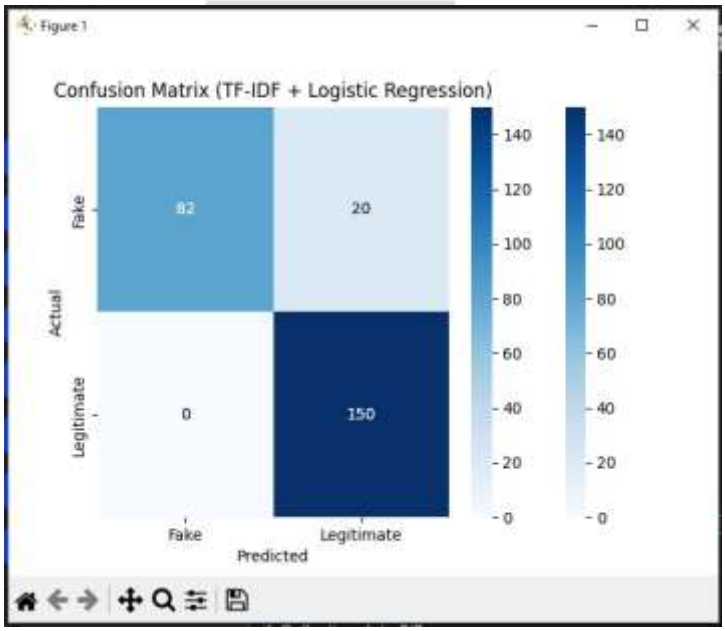

Figure 5.5 Confusion Matrix

VII. Conclusion and Future Scope

Fake news has become a widespread issue, influencing public opinion and misleading individuals. This project implemented by using Machine Learning algorithm to detect fake news, analyzing text features and evaluating classification performance using TF-IDF and supervised learning techniques. The results are demonstrated based on Logistic Regression and TF-IDF which have effectively classified news articles by predicted with greater accuracy.

For future scope, integrating deep learning models like BERT and LSTMs could enhance contextual understanding and improve detection rates. Additionally, incorporating social network analysis and real-time detection systems would enable more efficient identification of misinformation at scale.

REFERENCES

[1] SHU ET AL. (2017) - "FAKE NEWS DETECTION ON SOCIAL MEDIA: A DATA MINING PERSPECTIVE", ACM SIGKDD EXPLORATIONS. DOI: 10.1145/3137597.3137600.

[2] Zhou et al. (2020) - "Fake News Detection: A Survey of Machine Learning Techniques", ACM Computing Surveys. DOI: 10.1145/3395046.

[3] Kumar et al. (2018) - "False Information on Web and Social Media: A Survey", arXiv preprint: 1804.08559.

[4] Ruchansky et al. (2017) - "CSI: A Hybrid Deep Model for Fake News Detection", Proceedings of CIKM. DOI: 10.1145/3132847.3132877.

[5] Wang et al. (2018) - "EANN: Event Adversarial Neural Networks for Fake News Detection", Proceedings of KDD. DOI: 10.1145/3219819.3219903.

[6] Castillo et al. (2011) - "Information Credibility on Twitter", Proceedings of WWW.

[7] Gupta et al. (2014) - "TweetCred: Real-time credibility assessment of content on Twitter", Social Network Analysis and Mining. DOI: 10.1007/s13278-014-0163-5.

[8] Sharma et al. (2019) - "Combating Fake News: A Survey on Identification and Mitigation Techniques", ACM Transactions on Intelligent Systems and Technology (TIST). DOI: 10.1145/3305260.