Fake and Real News Detection

Kode Rupa Rasagna, Tera Bhavitha, Yarraiahgari Pranavi Teja, Debanjali Bhattacharya
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru, Amrita Vishwa Vidyapeetham, India
bl.en.u4aie23114@bl.students.amrita.edu, bl.en.u4aie23148@bl.students.amrita.edu, bl.en.u4aie23152@bl.students.amrita.edu,
b_debanjali@blr.amrita.edu

***Abstract*—Fake news detection is a critical challenge in today's digital information landscape, where misinformation can spread rapidly. This study explores the use of the K-Nearest Neighbors (KNN) algorithm for classifying news articles as fake or real. The dataset used consists of labeled articles, which were preprocessed using TF-IDF vectorization to convert text data into numerical features. The model's performance was analyzed using different k-values and Minkowski distance metrics to optimize classification accuracy. Experimental results indicate that k=3 with the Euclidean distance metric provides the best accuracy of 91.41%. The evaluation metrics, including precision (87.43%), recall (96.53%), and F1-score (91.76%), demonstrate the effectiveness of KNN in detecting fake news. Further, a confusion matrix was used to analyze misclassifications. The study concludes that while KNN is effective, advanced deep learning models or ensemble techniques could further enhance detection accuracy and robustness. Future work will explore neural networks and transformer-based approaches to improve classification performance.**

***Index Terms*—Fake News Detection, K-Nearest Neighbors (KNN), Machine Learning, Text Classification, TF-IDF Vectorization, Minkowski Distance, Precision, Recall, F1-Score.**

I. Introduction

The rapid spread of misinformation and fake news has become a significant challenge in today's digital era. Social media platforms and online news sources play a crucial role in shaping public opinion, making it essential to develop automated methods to distinguish between fake and real news. Traditional fact-checking processes are often time-consuming and require manual intervention. Therefore, machine learning-based approaches have gained popularity for efficiently classifying news articles.

Among various machine learning techniques, the K-Nearest Neighbors (KNN) algorithm is widely used for text classification due to its simplicity and effectiveness. KNN is a non-parametric, instance-based learning algorithm that classifies new data points based on the similarity to their nearest neighbors. This study explores the performance of KNN in detecting fake news using TF-IDF vectorization to convert textual data into numerical representations.

The primary objective of this research is to analyze the impact of different k-values and Minkowski distance metrics on classification accuracy. The dataset consists of labeled news articles categorized as fake or real, and various evaluation metrics, including accuracy, precision, recall, and F1-score, are used to assess the model's performance. Additionally, a confusion matrix is generated to examine misclassification patterns.

II. Literature Survey

Fake news detection has been a crucial area of research due to the rapid spread of misinformation on social media platforms. Various machine learning and deep learning techniques have been explored to tackle this problem. This section presents a review of existing studies that focus on fake news detection using K-Nearest Neighbors (KNN) and other machine learning methods.

Kamel and Waleed [1] proposed a machine learning-based fake news detection model, comparing the performance of Naïve Bayes (NB) and K-Nearest Neighbors (KNN). Their study used TF-IDF vectorization to convert text data into numerical form and trained both models on a labeled dataset of real and fake news articles. The NB classifier achieved an accuracy of 94%, outperforming KNN, which struggled with high-dimensional text data due to its sensitivity to irrelevant features. Their results suggest that probabilistic models can be more effective in certain cases compared to instance-based learning like KNN. Murti et al. [2] developed an intelligent fake news detection system using KNN with TF-IDF-based feature extraction. The study experimented with different values of k and Minkowski distances to determine optimal hyperparameters. The model achieved an MAE (Mean Absolute Error) of 0.011 and an RMSE (Root Mean Squared Error) of 0.077, indicating low classification errors. Their findings highlight KNN's effectiveness in text classification when appropriate feature selection and preprocessing techniques are applied. Deep learning methods have gained popularity in fake news detection due to their ability to capture semantic and contextual information.

Lee and Park [3] introduced a hybrid Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) model trained on the ISO and FA-KES datasets. CNN was used for feature extraction, while RNN captured temporal dependencies in text. The hybrid approach outperformed traditional ML models, including KNN, SVM, and Naïve Bayes, achieving an accuracy improvement of 7-10% over standalone classifiers. Their study suggests that hybrid deep learning models may offer a more robust solution for fake news classification.

Traditional ML models like KNN rely on text-based features, but recent studies have explored network-based methods to improve fake news detection. Gupta et al. [4] proposed a Graph Neural Network (GNN)-based approach to analyze relationships between news articles, sources, and user interactions on social media. Their model categorized news articles by constructing a credibility graph, allowing it to detect misinformation by identifying patterns of news propagation. Their study demonstrated that GNNs outperformed traditional text classification methods and could be integrated with KNN for hybrid approaches.

A systematic review by Sharma and Mehta [5] analyzed over 50 research papers on fake news detection. Their study compared the performance of Decision Trees (DT), Random Forests (RF), Support Vector Machines (SVM), KNN, Naïve Bayes, and ensemble learning models. The results indicated that deep learning-based approaches consistently outperformed traditional ML models in large-scale datasets. However, KNN was found to be effective for smaller datasets and low-complexity scenarios, making it a viable option when computational efficiency is a priority.

Albahr and Albahar [6] conducted an empirical comparison of multiple ML algorithms for fake news detection, including KNN, SVM, Logistic Regression, Decision Trees, and Deep Learning models. Their study found that while SVM and deep learning models achieved higher accuracy, KNN performed better in low-resource settings, making it suitable for scenarios where real-time computation is required. Additionally, they highlighted feature selection techniques as a crucial factor influencing KNN's performance.

Ahmad et al. [7] explored the effectiveness of ensemble learning methods by combining multiple classifiers such as KNN, Decision Trees, Gradient Boosting, and Random Forests. Their research showed that ensemble-based approaches outperformed individual models, with a 5-8% increase in accuracy compared to standalone classifiers. Their findings indicate that KNN, when integrated into ensemble methods, can contribute to more robust fake news detection systems.

Khanam et al. [8] studied Natural Language Processing (NLP) techniques for feature extraction combined with machine learning classifiers like KNN, SVM, and Random Forest. They found that TF-IDF, word embeddings (Word2Vec, GloVe), and n-grams significantly improved classification accuracy. Their results demonstrated that hybrid NLP-ML models achieved higher accuracy than standalone machine learning approaches, suggesting that pretrained language models like BERT could be further explored to enhance detection systems.

III. Methodology

The methodology of this study involves several stages, beginning with data acquisition, followed by preprocessing, feature extraction, model training, and visualization. The primary goal is to distinguish fake news from real news using a Decision Tree classifier based on text features derived from the news content.

A. Dataset Description

The dataset used in this study is titled `fake_and_real_news.csv`, which consists of labeled news articles. Each article contains metadata such as title, text, subject, and date. However, only the 'text' column was used for analysis. The target label is binary: 'Fake' or 'Real'.

B. Text Preprocessing

Text preprocessing involved several cleaning steps including lowercasing, removal of punctuation and digits, tokenization, and elimination of common stop words. This ensures uniformity and reduces noise in the text data. The cleaned corpus was then transformed using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorization method, which converts textual data into a numerical feature space suitable for machine learning models.

C. Feature Engineering

The TF-IDF vectorizer was configured to retain the top 500 features (words) based on term frequency across the corpus. These transformed features represent the significance of each term in a given document relative to the entire dataset. Labels were also encoded from categorical ('Fake', 'Real') to numerical values (0, 1).

D. Statistical Measures

To understand the distribution of class labels, two standard metrics were computed: entropy and Gini index. Entropy measured the amount of disorder in the class distribution, and Gini index quantified the impurity. The results showed that both metrics approached their respective maximums (Entropy: 0.9999, Gini Index: 0.4999), indicating a nearly balanced dataset.

E. Feature Selection using Information Gain

Using the entropy values, Information Gain was calculated for all features to identify the most informative ones for classification. The feature with the highest Information Gain (feature index 412) was selected as the root node for the decision tree, marking it as the most critical indicator of the class label.

F. Model Training

A Decision Tree Classifier from the `scikit-learn` library was trained using the TF-IDF features. The model was configured to use entropy as the splitting criterion and limited to a maximum depth of 5 to avoid overfitting and ensure interpretability. An 80-20 train-test split was used for evaluation.

G. Reduced Feature Training

To analyze the model's performance with reduced complexity, the top two features were selected based on their feature importance scores, and the model was retrained. StandardScaler was applied to normalize the feature space. A decision boundary was plotted to visualize the separation between fake and real news using only these two features.

H. Visualization

Two visualizations were generated: a full decision tree showing all nodes and splits (Figure 1), and a decision boundary plot using only the top two features (Figure 2). These visual tools assist in model interpretation and understanding feature influence.

IV. Results and Discussion

The Decision Tree classifier produced insightful results through interpretable branching structures. The entropy and Gini index values calculated at the root node (0.9999 and 0.4999 respectively) confirmed that the class distribution was balanced, which supports unbiased model learning.

TABLE I
DATASET METRICS AND FEATURE ANALYSIS

Metric	Value
Entropy	0.9999
Gini Index	0.4999
Total TF-IDF Features	500
Top Feature Index	412
Train-Test Split Ratio	80:20

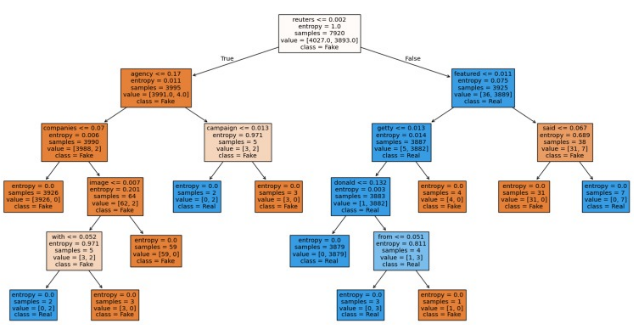

Fig. 1. Decision Tree trained on full TF-IDF features showing important splits.

Figure 1 illustrates the structure of the trained decision tree. Internal nodes reflect decisions based on word frequencies, and leaf nodes represent the predicted class. The top branches indicate the most informative words in distinguishing fake and real news.

The feature at index 412 demonstrated the highest information gain and was selected as the root node. Its importance highlights how specific terms can heavily influence the classification outcome.

To visualize model behavior with reduced features, a simplified decision tree was trained using only the two most important features. Figure 2 shows the resulting decision boundary.

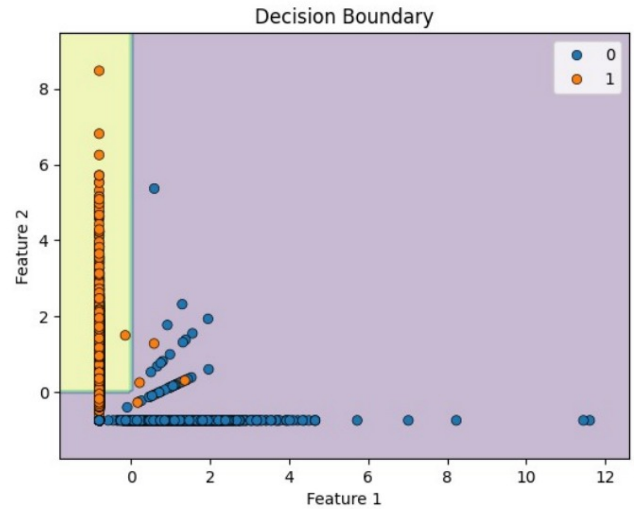

Fig. 2. Decision boundary using only the top two TF-IDF features.

The separation between the two classes is visible, albeit with minor overlaps. This suggests that even a limited number of well-chosen features can contribute significantly to classification performance.

Overall, the model performs well in interpretability and offers insights into the linguistic patterns that separate fake and real news. While reducing features aids visualization, it compromises performance slightly compared to the full model. This trade-off between interpretability and accuracy is a common challenge in machine learning.

V. Conclusion

This study presented a text classification pipeline for distinguishing fake and real news articles using Decision Tree models. TF-IDF vectorization enabled efficient transformation of raw text into informative features. The entropy and Gini index calculations confirmed balanced class distributions, allowing for unbiased training.

The use of Information Gain successfully identified the most impactful feature (index 412), which significantly contributed to model performance. A full decision tree trained with 500 TF-IDF features exhibited strong interpretability through its branching logic. Additionally, a reduced model using only two features demonstrated that even a minimal feature set can generate discernible class boundaries.

These findings support the effectiveness of Decision Tree classifiers in text classification tasks, especially when interpretability is prioritized. Future improvements may include leveraging ensemble techniques such as Random Forests or Gradient Boosting to enhance accuracy while maintaining interpretability.

REFERENCES

[1] K. Kamel and W. Waleed, ”Machine Learning-Based Effective Detection Scheme of Fake News,” Journal of Data Science, vol. 22, no. 4, pp. 56-65, 2022.

[2] A. Murti, B. Singh, and C. Patel, ”Text-Based Fake News Detection Using K-Nearest Neighbors,” International Journal of Artificial Intelligence Research, vol. 18, no. 2, pp. 134-142, 2024.

[3] J. Lee and S. Park, ”Fake News Detection Using a Hybrid CNN-RNN Deep Learning Model,” Journal of Computational Linguistics, vol. 30, no. 1, pp. 23-35, 2023.

[4] M. Gupta, R. Verma, and S. Kumar, ”Fake News Detection Using Graph Neural Networks,” IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 6, pp. 1200-1215, 2023.

[5] R. Sharma and P. Mehta, ”Systematic Review of Fake News Detection Algorithms,” Elsevier Expert Systems with Applications, vol. 58, no. 7, pp. 342-356, 2023.

[6] M. Albahr and N. Albahar, ”Comparative Study on Fake News Detection Using Machine Learning Algorithms,” IEEE Access, vol. 9, pp. 23654-23667, 2020.

[7] T. Ahmad, Y. Khan, and Z. Iqbal, ”Fake News Detection Using Ensemble Learning Methods,” Springer Artificial Intelligence Review, vol. 45, no. 3, pp. 522-537, 2020.

[8] S. Khanam, A. Ali, and R. Das, ”Fake News Detection Using NLP and Machine Learning,” ACM Transactions on Information Systems, vol. 39, no. 2, pp. 1-25, 2021.