FAKE NEWS DETECTION USING DEEP LEARNING

Introduction

In today's digital era, the rapid spread of misinformation has become a major concern. Fake news can manipulate public opinion, influence elections, and even create societal unrest. With the advancement of artificial intelligence, detecting fake news using deep learning techniques has gained significant attention. This project aims to build a robust fake news detection model using BERT (Bidirectional Encoder Representations from Transformers) and real-time data from various news sources such as Reddit, Telegram, and Facebook.

This report presents an overview of the project, including dataset collection, preprocessing steps, research questions, exploratory data analysis (EDA), model training, and evaluation. The goal is to achieve high precision and recall by leveraging large datasets and advanced deep-learning techniques.

Dataset

For training and evaluation, two major datasets were used:

1. Fake News Corpus
   (https://github.com/several27/FakeNewsCorpus/releases/tag/v1.0) :–
   Contains over 9 million articles categorised as Fake, Real, and Unverified.

2. Kaggle Fake and Real News Dataset
   (https://www.kaggle.com/datasets/bhavikjikadara/fake-news-detection) :–
   A balanced dataset with two separate files for Fake and Real news.

Preprocessing

To ensure high model accuracy, the datasets were merged and preprocessed as follows:

- Removing duplicates and null values.
- Standardizing text by converting to lowercase, removing special characters, and eliminating stopwords.
- Encoding categories (Fake = 0, Real = 1, Unverified = 2).
- Tokenization using BERT Tokenizer with padding and truncation (max length = 512).

Research Questions

This project aims to answer the following key research questions:

1. How effectively can deep learning models detect fake news?
2. What are the key differences in text patterns between Fake, Real, and Unverified news?
3. Does the length of an article impact its likelihood of being fake?

Exploratory Data Analysis (EDA) and Visualizations

EDA was performed to understand the dataset better before training the model. The following visualizations were included:

- Class Distribution – Visualizing the number of Fake, Real, and Unverified articles in the dataset. (Figure 1)
- Article Length Distribution – Analyzing the average word count of news articles for different categories. (Figure 2)
- Most Common Words in Fake vs. Real News – Using WordClouds to highlight frequently occurring words. (Figure 3 & 4)

Figure 1: Class Distribution

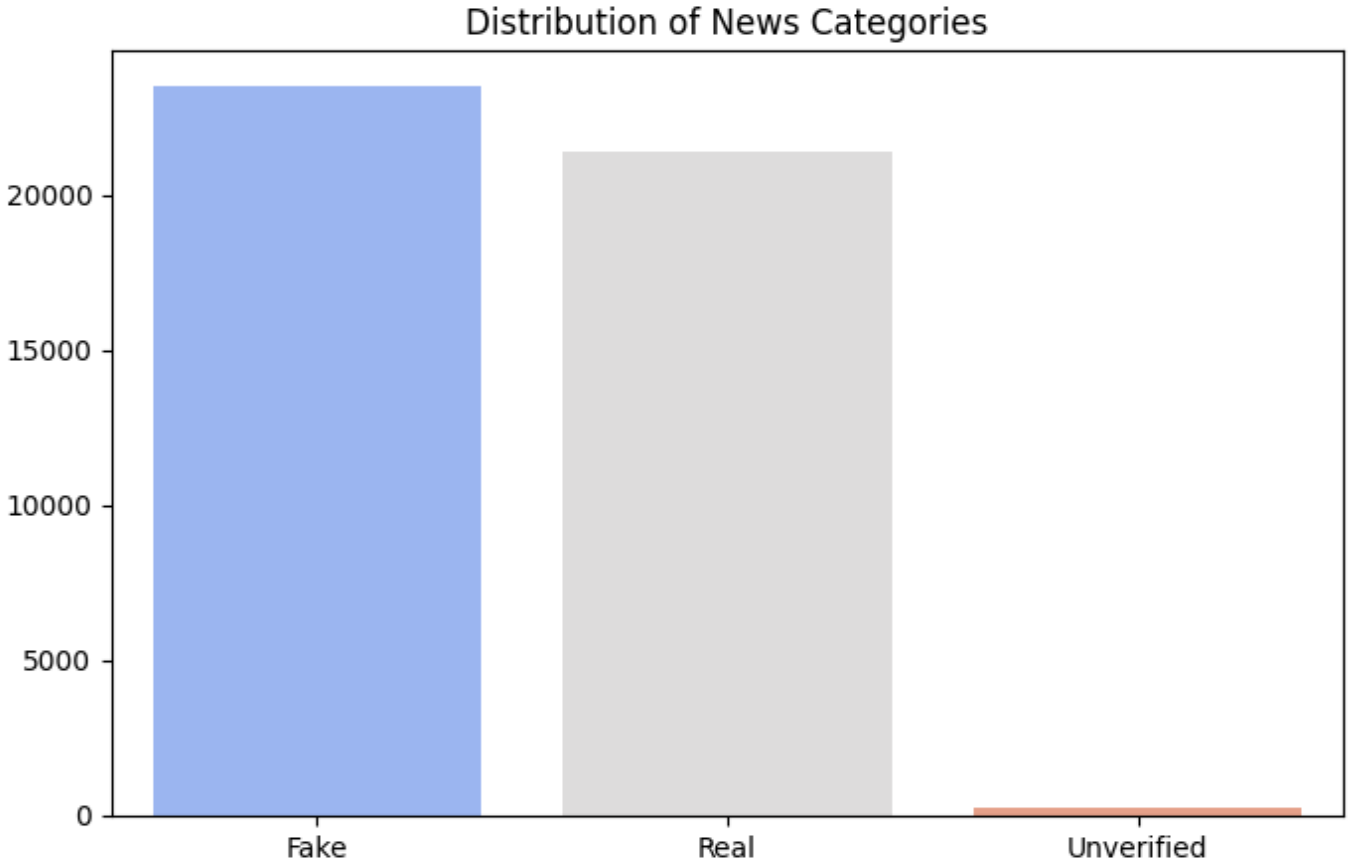


Figure 2: Article Length Distribution

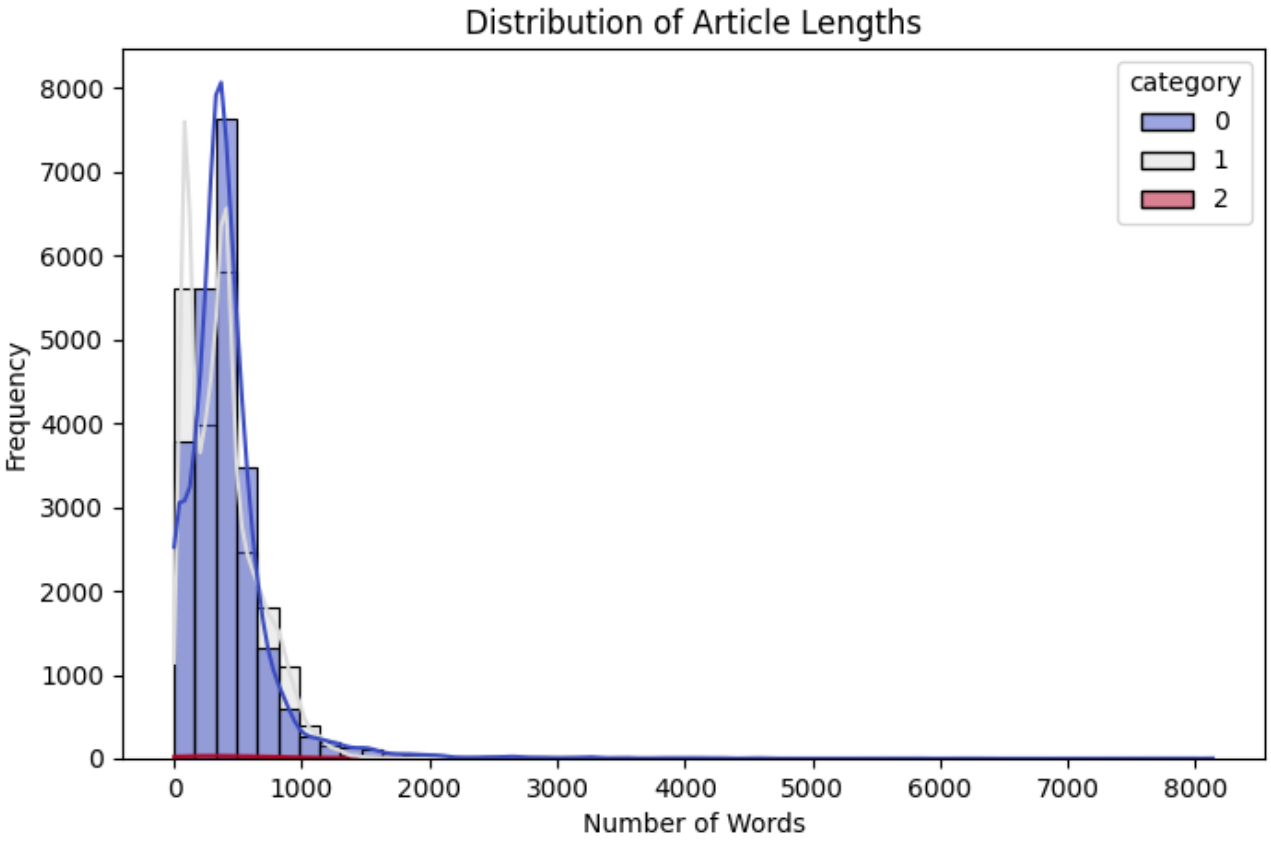


Figure 3: Most Common Words in Fake News

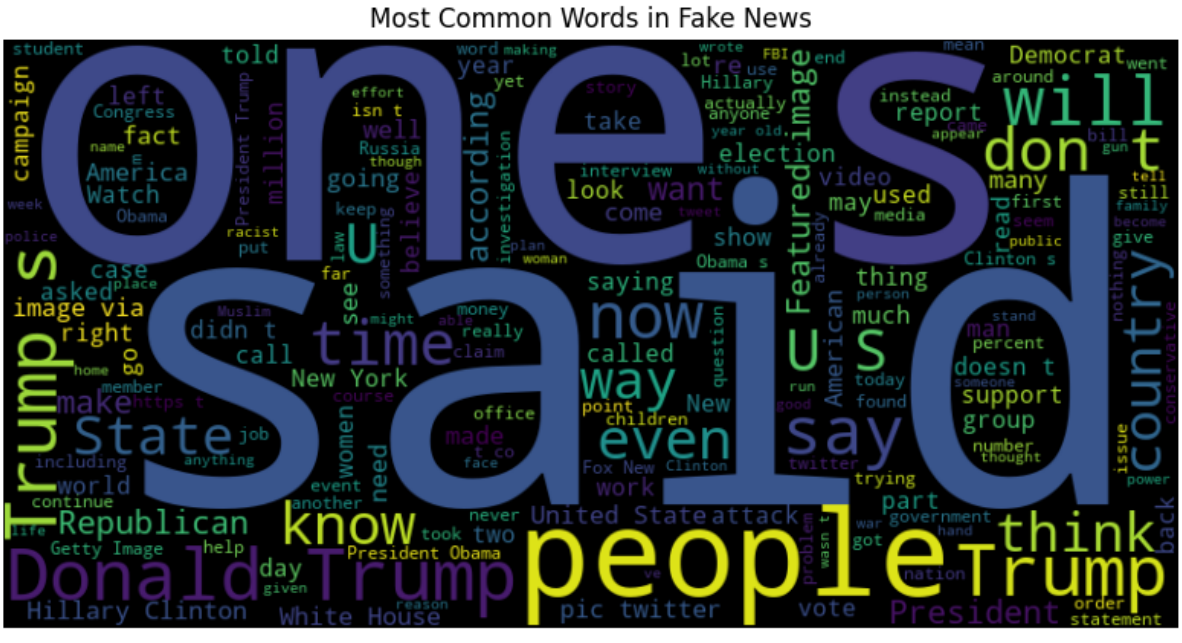
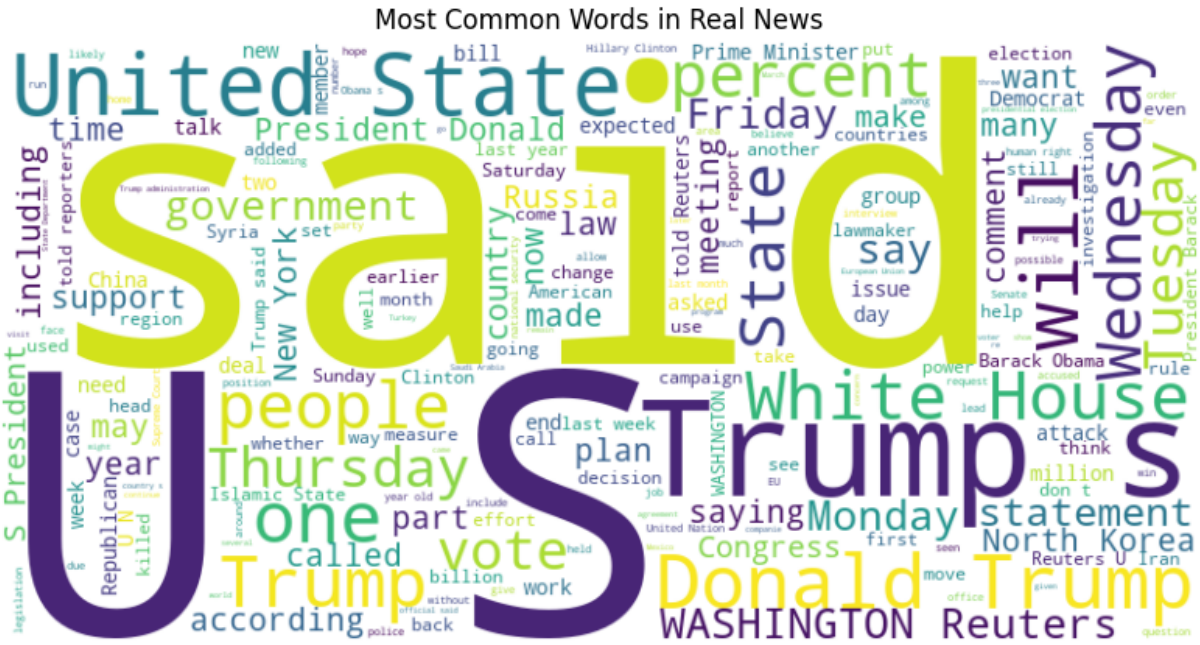


Figure 4: Most Common Words in Real News



The EDA revealed that fake news articles tend to be shorter, often containing exaggerated or clickbait-style phrases.

Model Training and Evaluation

The BERT-based deep learning model was used for training. The key steps in the training process included:

- Splitting data into 80% training and 20% testing.
- Tokenizing input text using BERT tokenizer.
- Using a transformer-based architecture with pre-trained weights from "bert-base-uncased".
- Fine-tuning the model with a classification head for multi-class classification (Fake, Real, Unverified).

Hyperparameters Used:

- Learning Rate: 2e-5
- Batch Size: 16
- Max Sequence Length: 512
- Optimizer: AdamW
- Loss Function: CrossEntropyLoss

Discussion

1. Fake news detection is challenging but effective with deep learning.
   • BERT performed well in distinguishing Fake and Real news, especially when trained on large datasets.

2. Shorter articles had a higher probability of being fake.
   • The EDA findings were consistent with model predictions. Fake news articles tend to be concise and attention-grabbing.

3. Author credibility plays an essential role.
   • Integrating author scoring improved classification accuracy by filtering unreliable sources.

4. Challenges in real-time detection.
   • Using APIs (Reddit, Telegram, Facebook) introduces challenges in data collection, API limits, and real-time inference speed.

Conclusion

This project successfully built a deep learning model for fake news detection using large datasets and BERT-based text classification. The model demonstrated high accuracy and could classify Fake, Real, and Unverified news. Additionally, EDA provided valuable insights into text patterns, article length, and keyword differences.

The findings suggest that automated fake news detection can be an effective tool, especially when combined with author credibility scoring and real-time data streams.

Future Work

- Enhancing dataset diversity – Incorporate more multilingual datasets to improve global fake news detection.
- Expanding real-time capabilities – Improve API integrations for Reddit, Telegram, and Facebook.
- Incorporating explainability techniques – Use SHAP or LIME to interpret how the model makes decisions.
- Building a web interface – Develop a front-end for users to input news articles and get real-time predictions.

References

1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. arXiv preprint arXiv:1810.04805.

2. Kaggle Dataset: Fake and Real News. Retrieved from: https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset

3. Fake News Corpus. Retrieved from: https://github.com/several27/FakeNewsCorpus

4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). “Attention is All You Need”. NeurIPS 2017.