AI Model to Detect Fake News

Khuram Iqbal

October 13, 2024

1. Project Title

AI Model to Detect Fake News

2. Team Members

- Khuram Iqbal (2022-CS-48, khuramiqbalofficial@gmail.com)

3. Objective

The objective of this project is to develop a machine learning model that can accurately identify fake news articles by analyzing text patterns, checking sources, and cross-referencing information using Natural Language Processing (NLP) techniques.

4. Background and Motivation

With the rise of digital media, fake news has become a significant issue, misleading readers and causing societal harm. The rapid spread of misinformation across platforms poses a threat to public trust and safety. Detecting fake news using AI offers a scalable solution that can sift through massive volumes of data and provide real-time verification, aiding both readers and fact-checkers.

5. Proposed Methodology

We will approach the project with the following steps:

Data Collection

We will use publicly available datasets like the FakeNewsNet and the LIAR dataset. Additional news articles will be scraped from trusted sources for real vs. fake news classification.

Data Preprocessing

The text data will be cleaned by removing stopwords, punctuation, and special characters. Tokenization and lemmatization will be applied to normalize the text.

Algorithms/Models

We plan to implement traditional machine learning models like Logistic Regression, Random Forests, and more advanced models like BERT and GPT-3 for text classification. The performance of these models will be evaluated based on accuracy, precision, recall.

Cross-Referencing Sources

We will integrate a feature to cross-reference the news articles with credible sources using APIs from fact-checking websites like PolitiFact or Snopes, further improving the model's reliability.

6. Expected Outcomes

The project aims to build a robust system capable of detecting fake news with high accuracy. The final model will offer an accessible tool that can provide real-time news validation, minimizing the spread of misinformation.

7. Resources Required

- Python (Pandas, scikit-learn, PyTorch)
- NLP Libraries (spaCy, NLTK, HuggingFace Transformers)
- Datasets: FakeNewsNet, LIAR, or scraped news data

8. Conclusion

This project will contribute towards fighting the proliferation of fake news using AI, leveraging cutting-edge NLP and machine learning techniques to ensure news accuracy and prevent misinformation.