KINGS ENGINEERING COLLEGE

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

COLLEGE CODE : 2108

PROJECT TITLE : Fake News Detection Using NLP

MENTOR : Mr.V.Sundarraj

Team Members :

Vignesh C

Kishore R

Manoj Kumar L

Omaravind L

INTRODUCTION:

In an era dominated by digital information, the unchecked spread of misinformation poses a formidable challenge to public discourse and decision-making. The rise of fake news has underscored the critical need for robust mechanisms to distinguish between reliable information and deceptive content. Leveraging the advancements in Natural Language Processing (NLP) and machine learning, this project aims to develop a sophisticated system for the automated detection of fake news.

The proliferation of online news sources, social media platforms, and the ease of content sharing have created an environment where misinformation can rapidly gain traction. Recognizing the urgency of addressing this issue, our project adopts a multi-faceted approach that combines NLP techniques, machine learning models, and web-scraping tools.

ABSTRACT:

The proliferation of misinformation in online news articles poses a significant threat to public discourse and decision-making. This project employs Natural Language Processing (NLP) techniques to develop an automated system for fake news detection. The methodology involves text preprocessing, feature extraction, semantic analysis using

advanced NLP models, and source credibility assessment. Machine learning models, trained on labeled datasets, are employed for classification, with ensemble methods enhancing overall accuracy. The system also incorporates continuous learning, fact-checking, and user feedback mechanisms to adapt to evolving misinformation tactics. The proposed approach demonstrates promising results in discerning between reliable and deceptive news, contributing to the ongoing efforts to mitigate the impact of fake news in the digital information landscape.

OBJECTIVES :

Text Preprocessing:

- Clean and preprocess the text data to remove noise, such as HTML tags, special characters, and stopwords.
- Tokenize the text into words or subwords for analysis.

Feature Extraction:

- Extract relevant features from the text, such as word frequency, n-grams, and sentiment analysis.
- Word embeddings (like Word2Vec or GloVe) can capture semantic relationships between words.

Semantic Analysis:

- Use NLP models to understand the semantics of the text. Bidirectional Encoder Representations from Transformers (BERT) is a powerful pre-trained model for this purpose.
- Analyze the context and meaning of words to identify subtle nuances that might indicate misinformation.

Source Credibility:

- Analyze the credibility of the news source. Check for a history of misinformation or biased reporting.
- Cross-reference information with reliable sources to verify claims.

Contextual Analysis :

- Understand the context in which certain statements are made. Misleading information often relies on taking statements out of context.
- Analyze the overall tone and sentiment of the article.

Fact-Checking:

- Use fact-checking databases to verify specific claims and statements made in the news.
- Automated fact-checking tools can assist in quickly flagging potentially false information.

Machine Learning Models:

- Train machine learning models, such as classification algorithms, on labeled datasets of real and fake news.
- Features extracted from the text and other relevant metadata can be used as inputs to these models.

Ensemble Methods:

- Combine the predictions of multiple models using ensemble methods. This can improve overall accuracy and robustness.

TOOLS USES :

- NLTK (Natural Language Toolkit).
- Spacy.
- Git.
- Scikit-learn.
- TensorFlow and PyTorch.
- Fact-checking APIs.
- Jupyter Notebooks.

Results and Evaluation:

- Present the findings from the application of your model to real-world data.
- Utilize metrics such as accuracy, precision, recall, and F1-score to quantify the performance.
- Visualize results with charts or graphs if applicable.

Future Work:

- Discuss potential enhancements and future directions for the research.
- Address any limitations of the current approach and propose solutions for improvement.

Conclusion:

- Summarize the key findings, contributions, and implications of the research.