**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**

**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 :**

NLTK (Natural Language Toolkit)