**Fake News Detection Project Report**

**Introduction**

The rise of misinformation has made fake news detection a critical task in the modern digital landscape. This project aims to develop a machine learning-based system to classify news articles as either **Real** or **Fake** based on textual analysis. The application is built using **Flask** for deployment and leverages a trained model to predict the authenticity of news content.

**Dataset**

The dataset used in this project consists of labeled news articles with two categories:

* **Fake News**: Articles that contain misleading or false information.
* **Real News**: Authentic and verified news content.

The dataset was preprocessed and split into training and testing sets to train the machine learning model.

**Technology Stack**

* **Programming Language**: Python
* **Framework**: Flask (for web application)
* **Machine Learning Model**: Trained using Scikit-learn
* **Libraries**: Pandas, NumPy, Scikit-learn, NLTK, Flask
* **Model Storage**: model.pkl (pre-trained model)

**Data Preprocessing**

1. **Text Cleaning**: Removed stopwords, punctuation, and special characters.
2. **Tokenization**: Split text into meaningful words.
3. **Vectorization**: Converted text data into numerical format using TF-IDF.
4. **Model Training**: Applied algorithms such as Logistic Regression and Naïve Bayes.

**Model Performance**

The trained model was evaluated on accuracy, precision, recall, and F1-score. The best-performing model was selected based on test set performance.

**Application Implementation**

The Flask web application includes:

* A user interface (UI) with input for entering news text.
* Backend processing to predict whether the input news is Fake or Real.
* Displaying results with confidence scores.

**Files Included**

* app.py - Main Flask application.
* model.pkl - Trained machine learning model.
* templates/ - HTML files for the web interface.
* requirements.txt - Dependencies for running the project.
* report.pdf - Detailed documentation of the project.

**Future Enhancements**

* **Deep Learning Integration**: Implementing LSTM or BERT models for improved accuracy.
* **API Development**: Creating an API for real-time fake news detection.
* **Dataset Expansion**: Including multilingual datasets to handle a broader range of news articles.

**Conclusion**

This project successfully detects fake news using machine learning techniques and provides a simple web interface for users. It serves as a foundation for combating misinformation in digital media.