**Fake News Detection System Report**

**1. Introduction**

The **Fake News Detection System** is designed to classify news articles as **real** or **fake** based on their textual content. It utilizes **Natural Language Processing (NLP)** and **machine learning algorithms** to analyze and predict the authenticity of news articles.

**2. Approach Used**

**Data Preprocessing**

* Loaded the dataset containing news articles.
* Handled missing values by replacing them with empty strings.
* Created a new feature content by combining title and author columns.
* Cleaned the text by:
  + Removing special characters.
  + Converting text to lowercase.
  + Eliminating stopwords.
  + Applying stemming using **PorterStemmer** to reduce words to their root forms.

**Feature Extraction**

* Used **TF-IDF (Term Frequency-Inverse Document Frequency)** vectorization to convert textual data into numerical form.

**Model Training & Evaluation**

* Split the dataset into **training and testing** sets.
* Trained a **Logistic Regression** model for classification.
* Evaluated the model using the **accuracy score**.

**3. Challenges Faced**

During development, the following challenges were encountered:

* **Data Imbalance:** Unequal distribution of fake and real news impacted model performance.
* **Feature Engineering:** Careful selection and handling of features were required.
* **Computational Complexity:** Processing large text data increased computation time.
* **Overfitting Risk:** Needed to balance model performance without overfitting.

**4. Model Performance & Future Improvements**

The **Logistic Regression** model performed well in detecting fake news. However, further improvements can be made:

* **Deep Learning Models:** Implementing **LSTMs** or **Transformers** for better accuracy.
* **Ensemble Methods:** Combining multiple models to improve classification.
* **Real-time Data Integration:** Using APIs to detect fake news dynamically.

**5. Uploading the Project to GitHub**

Follow these steps to upload the project:

**Initialize a Git Repository**

git init

**Add Files to Repository**

git add .

**Commit the Changes**

git commit -m "Initial commit - Fake News Detection System"

**Create a New Repository on GitHub**

1. Go to [GitHub](https://github.com/).
2. Click on **"New Repository"** and provide a name.
3. Copy the repository **URL**.

**Push the Code to GitHub**

git remote add origin <repository\_url>

git branch -M main

git push -u origin main

Now, the **Fake News Detection System** is successfully uploaded to GitHub.

**6. Conclusion**

The **Fake News Detection System** demonstrates the power of **NLP** and **machine learning** in combating misinformation. While the current model provides **good accuracy**, integrating **advanced techniques** and **real-time sources** will enhance its reliability.