**📝 Project Report**

**\*\*Title: \*\*** Fake News Detection using Logistic Regression

**\*\*1. Abstract\*\***

The Fake News Classifier is a machine learning model that distinguishes between real and fake news articles. Using TF-IDF vectorization and logistic regression, it predicts the authenticity of a given news article. This project aims to counter misinformation by offering a lightweight and accurate solution.

**\*\*2. Tools & Libraries Used\*\***

- Python

- pandas

- scikit-learn

- joblib

- streamlit

- TfidfVectorizer

- LogisticRegression

**\*\*3. Dataset\*\***

The dataset used is the “Fake and Real News” dataset from Kaggle, which includes two CSV files: `Fake.csv` and `True.csv`. Each entry includes a news title, full text, subject, and publication date.

**\*\*4. Project Workflow\*\***

1. Load and label data

2. Merge and shuffle

3. Combine title and text

4. Vectorize text using TF-IDF

5. Train logistic regression classifier

6. Save model and deploy using Streamlit

**\*\*5. Results\*\***

The model achieved over \*\*95% accuracy\*\* on the test set, indicating strong performance in classifying fake news.

**\*\*6. Future Enhancements\*\***

- Include BERT or transformer-based models for deeper contextual understanding

- Allow batch classification of multiple articles

- Deploy on web/cloud platforms

- Add explainability (e.g., SHAP values)

Created for **Elevate Labs** Internship by **K.Srinivas.**