

Fake News Detection System

Abstract

The Fake News Detection System is a machine learning-based project designed to identify whether a given news article is true or fake. The project utilizes natural language processing (NLP) techniques to analyze the textual content of news articles and classify them using trained models. This system helps in combating misinformation by providing quick and reliable predictions about the authenticity of news content.

Introduction

With the rise of online news and social media, the spread of fake news has become a major concern worldwide. Detecting fake news manually is time-consuming and unreliable. Therefore, this project implements an automated approach using NLP and machine learning to distinguish between true and fake news articles. The system is trained using a Kaggle dataset containing both true and fake news samples.

Tools Used

Python 3	Streamlit	Scikit-learn
NLTK	Pandas & NumPy	Joblib

Steps Involved in Building the Project

- Data Collection – Downloaded True and Fake news datasets from Kaggle.
- Data Preprocessing – Cleaned and tokenized text data using NLTK.
- Feature Extraction – Converted text to numerical form using TF-IDF Vectorization.
- Model Training – Trained machine learning models such as Logistic Regression and Passive Aggressive Classifier.
- Model Evaluation – Evaluated model accuracy and selected the best-performing model.
- Deployment – Built a Streamlit web application for real-time prediction.

Conclusion

The Fake News Detection System effectively classifies news articles as fake or true based on their content. It demonstrates the power of NLP and machine learning in tackling misinformation. The system can be further enhanced by integrating deep learning models and expanding the dataset for better generalization.