FAKE NEWS DETECTION USING MACHINE LEARNING

Project Report

Submitted by:

Name :

Register Number :

Department :

College Name :

Academic Year :

FAKE NEWS DETECTION USING MACHINE LEARNING

Abstract

Fake news has become a serious problem in the digital era, spreading misinformation rapidly through social media and online platforms. This project focuses on detecting fake news using Machine Learning and Natural Language Processing techniques. The system analyzes news content and classifies it as REAL or FAKE using text-based features.

Problem Statement

The goal of this project is to develop an automated system that can accurately identify fake news articles based on their textual content using machine learning classification models.

Objectives

- To study fake news and its impact
- To preprocess and clean news text data
- To extract features using TF-IDF
- To build a machine learning model for classification
- To evaluate the performance of the model

Technologies Used

- Python
- Machine Learning
- Natural Language Processing (NLP)
- Scikit-learn
- Pandas, NumPy
- NLTK

System Architecture

- Data Collection
- Text Preprocessing
- Feature Extraction
- Model Training
- Prediction & Evaluation

Methodology

The dataset is first cleaned by removing unwanted characters, stopwords, and punctuation. TF-IDF vectorization is applied to convert text into numerical format. Machine learning models such as Logistic Regression or Naive Bayes are trained on the dataset and evaluated using accuracy and classification metrics.

Advantages

- Helps prevent the spread of misinformation
- Fast and automated detection
- High accuracy with proper training
- Can be extended to real-time applications

Applications

- Social media platforms
- Online news portals
- Fact-checking organizations
- Educational and research purposes

Conclusion

The Fake News Detection system demonstrates how machine learning and NLP can be effectively used to combat misinformation. With proper datasets and advanced models, the system can achieve high accuracy and can be deployed in real-world applications.