

Fake News Detection Using Machine Learning

Name: Bhargavi

College Name: Sankethika Vidya Parishad of Technology and Management

Email: bhargavibharu25122000@gmail.com

Abstract

In today's digital world, the spread of fake news has become a serious issue. This project presents a machine learning-based solution that classifies news articles as real or fake using Natural Language Processing (NLP) techniques. The system is trained on a labeled dataset and uses algorithms like Logistic Regression to make predictions.

Problem Statement

Fake news spreads quickly and can influence public perception, politics, and social behavior. Manual detection is not scalable, hence an automated method using ML is needed.

Methodology

We used a dataset containing labeled news articles. The data was cleaned, preprocessed using NLP, vectorized using TF-IDF, and trained on multiple models. Logistic Regression performed best in our case.

Technologies Used

Python, scikit-learn, pandas, NumPy, NLTK, TF-IDF

Algorithm

Logistic Regression was used for binary classification.

Results

Achieved an accuracy of ~94% on the test data.

Conclusion

The model successfully detects fake news with high accuracy. The project demonstrates how machine learning can help tackle misinformation online.

Future Scope

Future improvements include deploying a web app, using deep learning models like BERT, and extending support for multiple languages.

References

1. <https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset>
2. scikit-learn documentation
3. NLP research papers