Fake News Detection Using Machine Learning Algorithms

By student
Maen Mohammad Alsarayrah

Supervisor
Prof. Qeethara Al-Shayea

Al-Zaytoonah University of Jordan , 2024

Abstract

This thesis investigates the application of advanced machine learning (ML) techniques in detecting fake news, a growing issue in digital media that erodes public trust and distorts democratic discourse. Utilizing a comprehensive dataset of news articles verified by credibility assessment platforms, several ML strategies were implemented and compared, including Naive Bayes, Support Vector Machines (SVM), Decision Trees, Logistic Regression, Random Forests, and Neural Networks, as well as unsupervised methods like K-means clustering and Latent Semantic Indexing. The research demonstrates that ensemble methods, particularly Random Forests, exhibit superior accuracy and reliability. Incorporating Natural Language Processing (NLP) techniques such as sentiment analysis and word vector analysis (word2vec/GloVe), along with text representation methods like TF-IDF, further enhances detection precision by evaluating the emotional tone and semantic content of the news text. Additionally, the thesis examines the impact of article metadata, including publication patterns and topic frequency, on the likelihood of misinformation. The findings indicate that while ML techniques are powerful tools for identifying fake news, challenges such as data diversity and algorithmic biases persist. A hybrid approach combining multiple ML techniques and robust data preprocessing methods is proposed to improve detection rates. The implications of these technologies in real-world applications are discussed, with suggestions for future research emphasizing the need for interdisciplinary approaches to refine the effectiveness of fake news detection systems. This study contributes to ongoing efforts to create a reliable digital information ecosystem and underscores the critical role of ML in combating misinformation on a large scale.