An explainable ensemble of multi-view deep learning model for fake review detection

Rami Mohawesh
a,b,
⇑
, Shuxiang Xu
a
, Matthew Springer
a
, Yaser Jararweh
c
, Muna Al-Hawawreh
d
,

Sumbal Maqsood
a

Introduction

The proliferation of online reviews has led to an increase in fake reviews, posing a significant challenge for consumers and businesses. Detecting fake reviews is crucial for maintaining the integrity of online platforms and ensuring informed consumer decision-making. This literature review aims to provide an overview of existing fake review detection models, their limitations, and the proposed new model's innovative features and capabilities.

Existing Fake Review Detection Models

This section provides an analysis of various existing fake review detection models, including TensorD, SWNN, ABNN, AEDA, and SAE. The review discusses the features, algorithms, and performance of these models, highlighting their limitations in accurately detecting fake reviews. The section emphasizes the need for new detection models that efficiently identify fake reviews by combining features from different perspectives, such as review content, reviewers, and products.

Proposed Multi-View Deep Learning Model

The proposed multi-view deep learning model is an ensemble model that integrates features from review content, reviewers, and products. It consists of parallel word and sentence Bi-LSTM architectures, convolutional neural network (CNN), and deep neural network (DNN). The section outlines the methodology, feature engineering, and experimental results of the proposed model, demonstrating its effectiveness in detecting fake reviews on real-world datasets from Yelp.com.

Explainability in Fake Review Detection Models

This section discusses the importance of explainability in machine learning-based detection models and its impact on the reliability and trustworthiness of the models. It highlights the use of the SHAP method and attention mechanism to interpret the proposed model and understand its

behavior. The section emphasizes the significance of explainability in enhancing the decision-making process for customers and businesses, particularly in sensitive research areas.

Future Directions and Conclusion
The literature review concludes by highlighting the potential for improvement in fake review detection models and the need for responsible approaches to machine and deep learning-based models. It outlines future research directions, including the evaluation of the proposed method's effectiveness on cross-domain datasets and the integration of multiple classifiers and feature extraction methods to improve accuracy on larger datasets.

LSTMCNN: A hybrid machine learning model to unmask fake news Deepali Goyal Dev a , Vishal Bhatnagar b , Bhoopesh Singh Bhati c , Manoj Gupta d,**, Aziz Nanthaamornphong e,*

Based on the content provided, the literature review focuses on identifying and detecting fake news using machine learning and data mining tools. The paper by Ahmad et al. proposes the use of a machine learning ensemble strategy for automated categorization of news articles. Kaliyar et al. conducted comprehensive training of various machine learning algorithms through collaborative approaches and assessed their performance across four real-world datasets using linguistic features to discern between genuine and fraudulent content. The results of their novel approach demonstrated the superiority of their collaborative ensemble strategy over individual learners.The paper also discusses the challenges of identifying fake news, including the rapid dissemination of false information on social media platforms and the difficulty in authenticating misleading information. The establishment of a comprehensive system designed to provide a dependable solution to address this issue is proposed.The literature survey provides a clear and organized summary of comprehensive comparative analysis with baseline studies. The hybrid approach of CNN and LSTM outperforms others with an accuracy rate of 0.98. The study uses voice and text features and analytical representations to spot fake news. The primary goal in training is to identify textual characteristics that distinguish fake articles from genuine ones.Overall, the literature review provides valuable insights into the challenges of identifying and detecting fake news and proposes machine learning and data mining tools as a potential solution.

Boosting Fake News Detection Performance with Proper Word Embeddings

Introduction

Fake news has become a significant problem in recent years, and various methods have been explored to detect it. This literature review focuses on the use of proper word embeddings to boost fake news detection performance.

Related Works

Several studies have been conducted to detect fake news using different methods, such as N-gram and TF-IDF methods for feature extraction, and various classifiers such as Stochastic Gradient Descent (SGD), Support Vector Machine (SVM), and Multilayer Perceptron (MLP). Other studies have explored the use of deep learning methods such as Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Bidirectional LSTM. Some studies have also tested the impact of secondary features such as news domains, news writers, and headlines on fake news detection performance.

Research Method

The research method used in this study consists of two main phases: the training phase and the testing phase. The training phase involves retrieving and pre-processing the data, augmenting the data using the back-translation method, and training the model using deep learning methods. The testing phase involves evaluating the model's performance using various metrics such as accuracy, precision, recall, and F1-score.

Data and Devices Used

The data used in this study consists of several datasets, such as the ISOT Fake News Dataset, Fake News Dataset, and Fake News Detection Dataset. The devices used in this study include a computer with an Intel Core i7 processor, 16GB RAM, and an NVIDIA GeForce GTX 1080 Ti graphics card.

Experimental Results and Discussions

The experimental results show that the use of proper word embeddings can significantly improve fake news detection performance. The study shows that data augmentation has a positive effect, especially in improving model performance consistency. Several deep learning methods such as CNN, Bidirectional LSTM, and LSTM have been tested, and the results obtained were varied. Some studies have reported high accuracy values, such as 98.75% accuracy using Bidirectional LSTM, while others have reported lower accuracy values, such as 91.48% accuracy using Unidirectional LSTM.

Conclusions and Future Work

In conclusion, the use of proper word embeddings can significantly improve fake news detection performance. However, there are still some challenges that need to be addressed, such as the lack of proper Indonesian text processing methods in the current Indonesian fake news detection system. Further data collection and adjustments to Indonesian text processing methods are needed to address these challenges. Future work can also explore the use of other data augmentation methods and deep learning methods to improve fake news detection performance.