Fake News Detection Model Report

By

Muhammad Ahmed Zaheer

Brainwave Matrix Solutions

AI/ML Intern

August, 2024.

Table of Contents

List of Figures

Abstract

This project explores two distinct approaches to fake news detection using machine learning and deep learning techniques. The first approach employs an ensemble voting method, combining four classifiers—Logistic Regression, Decision Tree, Gradient Boosting, and Random Forest—to achieve high accuracy in predicting fake and real news. The second approach utilizes a Long Short-Term Memory (LSTM) neural network, which, although slightly less accurate, effectively captures the sequential nature of text data. Despite the ensemble model producing superior metrics, the LSTM model is considered more suitable for text analysis due to its ability to understand context and word order. Constraints related to time and hardware prevented the combination of LSTM with ensemble voting, as well as the implementation of transformers, which are recognized for their state-of-the-art performance in natural language processing tasks. While the ensemble method was inspired by existing work, the addition of the LSTM model represents a novel contribution to the project. The results underscore the trade-offs between traditional machine learning methods and deep learning in the context of fake news detection.

1 Introduction

Machine learning offers a promising approach to addressing the challenge of fake news detection. By analysing patterns in text data, machine learning models can classify news articles as either genuine or fabricated. In this project, we explore the application of various machine learning classifiers, including logistic regression, decision trees, gradient boosting, and random forests, to the task of fake news detection. Additionally, we investigate the effectiveness of a Long Short-Term Memory (LSTM) neural network, a type of deep learning model particularly suited for sequential data, in improving the accuracy of fake news detection.

The primary goal of this study is to compare the performance of traditional machine learning classifiers with the LSTM model, analysing their ability to accurately distinguish between real and fake news articles. By evaluating these models, we aim to identify the most effective approach for automatic fake news detection and to provide insights into the strengths and limitations of each model. The outcomes of this study are intended to contribute to the development of more robust tools for combating misinformation online.

2 Methodology

The methodology outlines the systematic approach adopted to develop and evaluate the fake news detection model using a Long Short-Term Memory (LSTM) network. This section details the dataset used, data preprocessing steps, model architecture, training procedure, evaluation metrics, and the manual testing process.

2.1 Dataset

The dataset used for this project was acquired from Kaggle and consists of labelled news articles categorized as either "Fake News" or "Not Fake News." The dataset is divided into training and testing sets. The training set was used to train the LSTM model, while the testing set was used for evaluation purposes. The dataset includes the news text and labels, where 0 indicates fake news and 1 indicates genuine news.

2.2 Data Preprocessing

Preprocessing is a critical step in preparing the data for the LSTM model. The following steps were performed:

1. **Text Cleaning:** The text data was cleaned by removing special characters, punctuation, and stopwords. All text was converted to lowercase to maintain uniformity.
2. **Tokenization:** The cleaned text was tokenized using the Tokenizer class from TensorFlow, which converts text into sequences of integers, where each unique word is represented by a unique integer.
3. **Padding:** Since LSTM models require inputs of the same length, the sequences were padded to a uniform length of 300 tokens. Padding ensures that all input sequences have the same dimension, allowing for efficient batch processing during training.

2.3 Model Architecture

The LSTM model was chosen due to its ability to capture long-term dependencies in sequential data, making it well-suited for text classification tasks. The architecture consists of the following layers:

1. **Embedding Layer:** An embedding layer with 5000 input dimensions and 128 output dimensions was used to convert the input sequences into dense vectors of fixed size, which capture semantic information about the words.
2. **SpatialDropout1D Layer:** To prevent overfitting, a SpatialDropout1D layer was added with a dropout rate of 0.2, randomly dropping entire 1D feature maps during training.
3. **LSTM Layer:** The core of the model is an LSTM layer with 100 units, configured with a dropout rate of 0.2 for both standard and recurrent connections. This layer processes the sequence data, learning patterns relevant to distinguishing between fake and real news.

4. **Dense Output Layer:** A dense layer with a sigmoid activation function was used for binary classification, outputting a probability between 0 and 1.

2.4 Model Training

The model was compiled using the Adam optimizer and trained using the binary cross-entropy loss function, which is standard for binary classification tasks. The training process involved five epochs, with a batch size of 64. The model was trained on 75% of the dataset, while 25% was held out for validation.

2.5 Evaluation Metrics

The model's performance was evaluated using precision, recall, F1-score, and accuracy metrics. These metrics provide insight into the model's ability to correctly identify fake and real news articles. The LSTM model achieved an accuracy of 98%, demonstrating its effectiveness.

2.5 Manual Testing

In addition to standard evaluation, manual testing was conducted by inputting various news articles into the trained model to assess its real-world applicability. The manual testing process involved converting the text into sequences using the trained tokenizer, padding the sequences, and predicting the output using the LSTM model. The model's predictions were consistent with expectations, reinforcing its reliability.

This detailed methodology provides a comprehensive understanding of the approach taken to develop and validate the fake news detection model. The choice of an LSTM network and the rigorous preprocessing steps have contributed to the model's high accuracy and robustness.

3 Results

In this study, two approaches were implemented to build a fake news detection model: an ensemble model combining Logistic Regression, Decision Tree, Gradient Boosting, and Random Forest classifiers, and a separate LSTM-based model. Both models were evaluated using the same dataset, with the following performance metrics:

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5844
1	0.99	0.99	0.99	5376
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

Figure 2: Logistic Regression

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5844
1	0.99	0.99	0.99	5376
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

Figure 1: Decision Tree Classifier

	precision	recall	f1-score	support
0	1.00	0.99	1.00	5844
1	0.99	1.00	0.99	5376
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

Figure 4: Gradient Boosting Classifier

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5844
1	0.99	0.99	0.99	5376
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

Figure 3: Random Forest Classifier

	precision	recall	f1-score	support
0	0.98	0.98	0.98	5903
1	0.98	0.98	0.98	5317
accuracy			0.98	11220
macro avg	0.98	0.98	0.98	11220
weighted avg	0.98	0.98	0.98	11220

Figure 5: LSTM

Both models demonstrated high accuracy in detecting fake news, with the ensemble model slightly outperforming the LSTM model in terms of precision, recall, and F1-score. Despite this, the LSTM model's performance remains strong and competitive, particularly considering the advantages of deep learning techniques in handling textual data.

4 Conclusion

In this project, two distinct approaches were used to tackle the problem of fake news detection. The first approach involved ensemble voting with four different classifiers: Logistic Regression, Decision Tree, Gradient Boosting, and Random Forest. The ensemble model delivered highly accurate results, as evidenced by its precision, recall, and F1-score metrics. The second approach utilized an LSTM neural network model, which, while slightly less accurate, demonstrated strong performance in processing and understanding sequential data, which is essential for analyzing textual content like news articles.

Despite the ensemble model yielding marginally better metrics, the LSTM model is considered generally superior for this task due to its ability to capture the context and sequence of words, which are crucial in distinguishing between real and fake news. LSTM models excel at handling sequential data and can learn patterns over time, making them highly effective for text-based applications.

However, due to time and hardware limitations, ensemble voting with the LSTM model and the other classifiers was not pursued. Additionally, the potential use of transformers, which are even more powerful for text analysis, was also constrained by hardware limitations. Transformers offer cutting-edge performance in NLP tasks, but their computational demands exceeded the resources available for this project.

The ensemble approach used in this project was inspired by work found on GitHub. However, I built upon this by integrating the LSTM model, bringing a new dimension to the analysis and demonstrating an alternative method for achieving robust results in fake news detection. This combination of approaches highlights both the strengths of traditional machine learning methods and the advancements possible with deep learning, providing a comprehensive evaluation of fake news detection techniques.