Universidad Icesi
Facultad de Ingeniería
Departamento de computación y sistemas inteligentes
Computación y estructuras discretas III

Fake news detection using neural networks: a comparative analysis

1. Introduction

The proliferation of fake news in digital media has become a significant challenge in today's information landscape. This project aims to develop and compare different neural network architectures for automatic fake news detection.

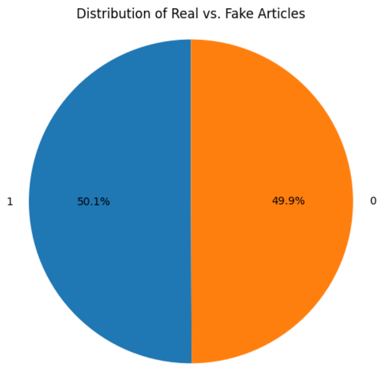
2. Project objectives

- Evaluate the effectiveness of Dense Neural Networks (DNN) in fake news detection Assess the performance of Recurrent Neural Networks (RNN) in fake news detection
- Analyze the capability of Long Short-Term Memory (LSTM) networks in fake news detection
- Compare the performance metrics of these three neural network architectures in the context of fake news detection

3. Data analysis and preprocessing

Dataset overview

The dataset consists of 20,800 news articles, each labeled as either reliable (0) or unreliable (1). The distribution shows a nearly balanced dataset with 50.1% unreliable and 49.9% reliable articles, eliminating any significant class imbalance concerns.



Preprocessing steps

1. Text Tokenization: Implemented using NLTK's word_tokenize function

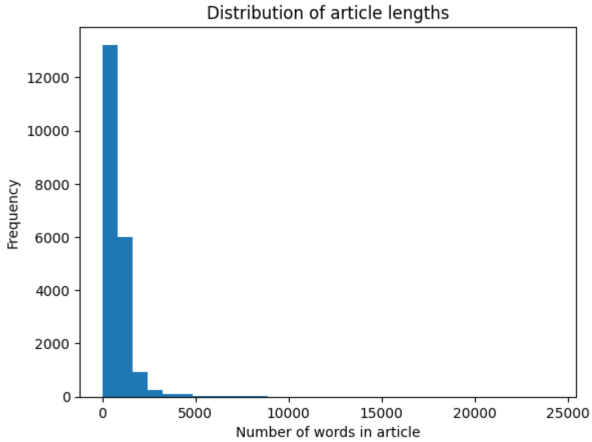
2. Text Normalization:

- Conversion to lowercase
- Removal of stopwords
- Filtering for alphanumeric tokens

3. Sequence Preparation:

- Limited vocabulary to 10,000 most frequent words
- Standardized sequence length to 100 tokens
- Applied padding for consistent input dimensions

The text length distribution analysis reveals that most articles contain fewer than 5,000 words, with a significant concentration below 2,500 words. This influenced our decision to set appropriate sequence lengths for the neural networks.



4. Model architectures and implementation

For this project, we implemented three different types of neural networks, each with specific characteristics suited for text classification:

Dense Neural Network (DNN)

DNNs are feedforward networks where each neuron is connected to every neuron in the subsequent layer. In our fake news detection task, we utilized this architecture for its ability to learn complex patterns in the text data through multiple layers of processing. While DNNs don't inherently handle sequential data, the combination with an embedding layer and global average pooling allowed effective text feature extraction.

Architecture:

- Input Layer: Embedding layer (10,000 words, 128-dimensional)
- Global Average Pooling
- Hidden Layer: 64 neurons with ReLU activation
- Batch Normalization Layer: For normalizing layer inputs, improving training stability
- Dropout Layer (0.5): To prevent overfitting by randomly deactivating 50% of neurons during training
- Output Layer: Single neuron with sigmoid activation for binary classification

Vanilla RNN

Recurrent Neural Networks are designed to work with sequential data by maintaining a hidden state that can capture information from previous inputs. We implemented this architecture to leverage its ability to consider word order and basic contextual relationships in the news articles.

Architecture:

- Input Layer: Embedding layer (10,000 words, 128-dimensional)
- Dropout Layer (0.2): Initial regularization to reduce overfitting in the embedding layer

- RNN Layer: 64 units for sequence processing
- Dropout Layer (0.2): Additional regularization after RNN processing
- Batch Normalization Layer: To normalize outputs and improve model stability
- Output Layer: Single neuron with sigmoid activation for binary classification

LSTM Network

Long Short-Term Memory networks are an advanced form of RNNs designed to better capture long-term dependencies in sequential data. We employed LSTM to potentially capture complex relationships between words and phrases across longer distances in the text, which could be crucial for detecting subtle patterns in fake news.

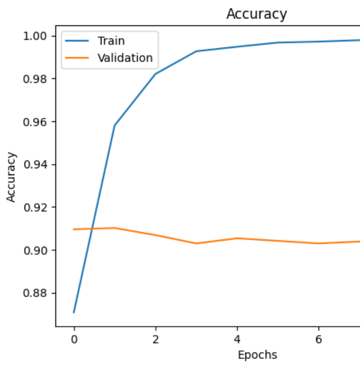
Architecture:

- Input Layer: Embedding layer (10,000 words, 128-dimensional)
- LSTM Layer: 64 units
- Output Layer: Single neuron with sigmoid activation

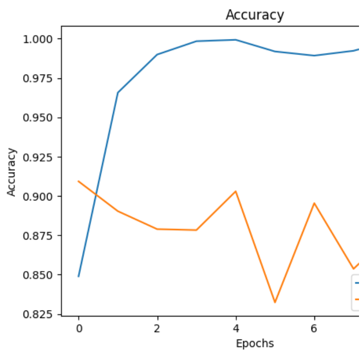
All models were compiled using binary cross-entropy loss and Adam optimizer, chosen for their effectiveness in binary classification tasks and optimization capabilities, respectively.

5. Performance evaluation

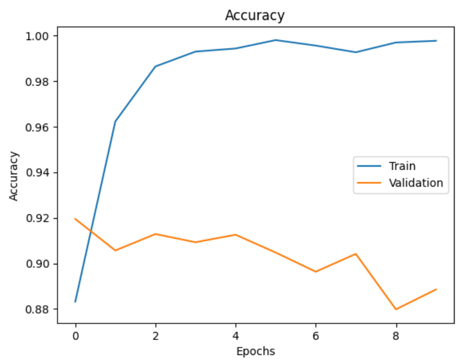
DNN:



RNN:



LSTM:



Model performance metrics

Model	Accuracy	Precision	Recall	F1-Score	Kappa
Dummy	0.4875	0.4875	0.50	0.6554	--------
DNN	0.9100	0.9049	0.9112	0.9081	0.8201
RNN	0.8819	0.8787	0.8791	0.8789	0.7637
LSTM	0.8963	0.8674	0.9294	0.8974	0.7930

6. Comparative analysis

The DNN model demonstrated superior performance across most metrics, achieving the highest accuracy (91.00%) and precision (0.9049). This strong performance suggests that even a relatively simple neural architecture can effectively capture the relevant features for fake news detection.

The LSTM model showed competitive performance with the highest recall (0.9294) and a strong F1-score (0.8974), though slightly lower accuracy than the DNN. Its strength in recall indicates better performance at identifying fake news articles, which could be attributed to its ability to maintain long-term dependencies in text data.

The vanilla RNN, while still performing significantly better than the baseline dummy classifier, showed the lowest performance among the three models with an accuracy of 0.8819 and F1-score of 0.8789. This aligns with the known limitations of vanilla RNNs in handling long sequences due to the vanishing gradient problem.

All three models significantly outperformed the dummy classifier baseline (accuracy: 0.4875), demonstrating the effectiveness of neural network approaches for this classification task.

7. Conclusion and future work

The study demonstrates that deep learning models, particularly Dense Neural Networks, can effectively detect fake news with high accuracy. The DNN model's superior performance, achieving 91% accuracy, suggests that the structural patterns in news articles can be effectively captured even without complex sequential processing. While LSTM showed stronger recall performance, indicating its value in reducing false negatives, the overall effectiveness of the simpler DNN architecture is particularly noteworthy from a practical implementation perspective.

Future work could explore:

Implementation of attention mechanisms to potentially enhance the performance of LSTM and RNN models

Integration of metadata features such as article source, publication date, and author reputation

Exploration of transformer-based architectures like BERT or GPT to potentially combine the benefits of both dense and sequential processing

Investigation of model interpretability to understand which features contribute most to classification decisions

Analysis of model performance on different types of fake news to identify potential weaknesses and biases

Hybrid architectures that combine the strengths of DNNs (overall accuracy) and LSTMs (high recall)

8. References

1. Natural Language Toolkit (NLTK). Bird, Steven, Edward Loper and Ewan Klein (2009).

2. M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar and M. S. Rahman, "A Comprehensive Review on Fake News Detection With Deep Learning," in IEEE Access, vol. 9, pp. 156151-156170, 2021, doi: 10.1109/ACCESS.2021.3129329.

3. Z Khanam *et al* 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1099** 012040.