FAKE NEWS DETECTION USING DEEP LEARNING

A PROJECT REPORT

Submitted by

DEEPA ANABATHULA – 00831857

SRUTHI BHONAGIRI– 00828687

Under the guidance of

Professor Mr. Khaled Sayed

For the Course DSCI-6011-02

DEEP LEARNING

UNIVERSITY OF NEW HAVEN

WEST HAVEN, CONNECTICUT

SPRING 2024

TABLE OF CONTENTS

ABSTRACT

Fake news is inaccurate information that is intentionally disseminated for a specific purpose. If allowed to spread, fake news can harm the political and social spheres, so several studies are conducted to detect fake news. Detecting fake news presents a formidable challenge due to the sheer volume of information available online and the rapidity with which it spreads. However, advancements in deep learning offer promising avenues for tackling this problem by leveraging vast datasets and sophisticated algorithms to discern patterns and differentiate between credible and dubious sources.

Manual fake news detection involves techniques that individuals can employ to verify the authenticity of news sources. This may entail visiting fact-checking websites or comparing unverified news with reliable sources.

Automated methods of fake news detection utilize Natural Language Processing (NLP) and Deep Learning (DL) approaches to automate the detection process. This study uses a deep learning method with several architectures such as CNN, LSTM and BERT combined with pre-trained word embedding, trained using dataset. Each data goes through a data augmentation process using the back-translation method to Reduce data imbalances between the classes. The results of the three architectures are outstanding they almost gave similar accuracy scores.

INTRODUCTION

Fake news, the dissemination of false or misleading information disguised as legitimate news, has become a pervasive issue in today's society, particularly with the advent of social media and digital communication channels. Its impact extends beyond mere misinformation, affecting public opinion, electoral processes, and societal discourse.

Manual Fake News Detection

Manual fake news detection involves techniques that individuals can employ to verify the authenticity of news sources. This may entail visiting fact-checking websites or comparing unverified news with reliable sources. However, manual verification is often impractical given the immense volume of information circulating online. Moreover, by the time fake news is identified and addressed, its detrimental effects may have already taken hold.

In an era where information spreads rapidly through social media and other online platforms, the task of manually verifying every piece of news becomes daunting. Individuals are bombarded with a constant stream of articles, posts, and videos, making it challenging to discern fact from fiction (Kong et al., 2020). Even with the availability of fact-checking websites and tools, the sheer volume of information makes it nearly impossible for individuals to thoroughly vet every piece of content they encounter.

Furthermore, fake news often spreads at an alarming rate, fueled by sensationalism and confirmation bias. By the time individuals identify a piece of fake news and attempt to debunk it, it may have already reached a wide audience and influenced public opinion. This rapid dissemination can have serious consequences, leading to misinformation, polarization, and even social unrest.

Moreover, fake news is not always easy to detect, especially for individuals who may lack the media literacy skills necessary to critically evaluate information. False stories are often crafted to mimic legitimate news sources, making them appear credible to the untrained eye. Additionally, the rise of deepfake technology further complicates the issue, as manipulated videos and audio clips can be indistinguishable from authentic content.While manual verification remains an important tool in the fight against fake news, it is clear that relying solely on individual efforts is not sufficient. Addressing the problem requires a multi-faceted approach involving collaboration between technology companies, media organizations, and policymakers (Kong et al., 2020). This may include implementing algorithms to detect and flag suspicious content, promoting media literacy education, and fostering greater transparency in online information ecosystems.

Automated Fake News Detection

Automated methods of fake news detection utilize Natural Language Processing (NLP) and Deep Learning (DL) approaches to automate the detection process. One such study, "Fake News Detection on Social Media: A Data Mining Perspective," highlights the prevalence of fake news on social media platforms and proposes a data mining approach to identify them. Key methodologies include feature engineering, sentiment analysis, and network analysis to detect patterns indicative of fake news propagation. Another significant contribution is a survey on identification and mitigation techniques, which provides an overview of various techniques employed for fake news detection, including content-based, social network-based, and hybrid approaches (Kong et al., 2020). Challenges discussed include the lack of labeled datasets, the evolving strategies of fake news creators, and ethical considerations surrounding censorship and freedom of expression.

Automated methods of fake news detection utilize Natural Language Processing (NLP) and Deep Learning (DL) approaches to automate the detection process.. The study by Shu, Guo, et al. emphasizes the importance of data mining techniques in this endeavor. By analyzing features such as linguistic patterns, sentiment, and network structures, these systems can discern signals indicative of fake news.

Similarly, the survey conducted by Zubiaga, Aker, et al. provides valuable insights into the diverse methodologies employed for fake news detection. Content-based approaches focus on analyzing the textual content of articles, whereas social network-based methods examine the propagation patterns of news stories across online platforms. Hybrid approaches combine multiple techniques to enhance detection accuracy. Despite these advancements, several challenges persist in the field.

One major obstacle is the scarcity of labeled datasets necessary for training machine learning models. The dynamic nature of fake news also poses a challenge, as creators continually adapt their strategies to evade detection. Moreover, there are ethical considerations regarding the balance between combating misinformation and preserving freedom of expression. Implementing detection systems may inadvertently lead to censorship or stifling legitimate discourse.

To address these challenges, ongoing research efforts are focused on developing robust detection algorithms capable of adapting to evolving tactics employed by fake news creators. Collaborative initiatives involving interdisciplinary teams are also underway to foster the creation of comprehensive datasets and frameworks for evaluating detection techniques (Rodríguez & Iglesias, 2019). Ultimately, the development of effective fake news detection systems requires a multifaceted approach that considers technological, social, and ethical dimensions.

DATASET

The dataset employed for this project is sourced from the real and fake news dataset, The number of training sentences in the dataset are almost 6,335 sentences.

- A full training dataset with the following attributes:

id: unique id for a news article

title: the title of a news article

author: author of the news article

text: the text of the article; could be incomplete.

label: a label that marks the article as potentially unreliable.

- A testing training dataset with all the same attributes at train.csv without the label.

DEEP LEARNING APPROACHES

The realm of fake news detection has seen significant advancements with the application of deep learning techniques. One notable exploration involves the utilization of Long Short-Term Memory (LSTM) networks, a subtype of recurrent neural networks (RNNs), for this specific task. LSTM networks excel in capturing temporal dependencies within textual data, making them adept at identifying deceptive language patterns commonly found in fake news articles (Rodríguez & Iglesias, 2019). This capability stems from their ability to retain information over long sequences, which is essential when analyzing textual data with intricate linguistic structures.

Building upon this foundation, another significant study delved into the efficacy of BERT (Bidirectional Encoder Representations from Transformers), a cutting-edge transformer-based model, in fake news detection. BERT represents a significant leap forward in natural language processing, leveraging bidirectional attention mechanisms to capture context from both preceding and succeeding words in a sequence. The research findings demonstrated substantial enhancements in classification accuracy compared to conventional machine learning methodologies (Rodríguez & Iglesias, 2019). This improvement underscores the effectiveness of leveraging pre-trained language representations and advanced neural network architectures for discerning the veracity of textual information.

However, despite the evident progress, challenges persist in the field of fake news detection. One prominent concern revolves around the computational resources required for implementing transformer-based models effectively. These models are computationally intensive, demanding substantial hardware resources for training and inference (Rodríguez & Iglesias, 2019). As a result, their widespread adoption may be limited by the availability of high-performance computing infrastructure.

Additionally, there are concerns regarding potential biases inherent in pre-trained language representations, such as BERT. These biases, present in the underlying data used for pre-training, could inadvertently influence the model's decision-making process. Biased representations may lead to skewed predictions, reinforcing existing stereotypes or misconceptions present in the training data (Hiramath & Deshpande, 2019). Addressing these biases is crucial to ensure fair and unbiased decision-making in fake news detection systems.

Despite these challenges, the studies collectively underscore the evolving landscape of fake news detection and the pivotal role that deep learning techniques play in advancing the field. By leveraging sophisticated neural network architectures like LSTM and BERT, researchers continue to push the boundaries of what is achievable in discerning the veracity of textual information. These efforts contribute to the ongoing battle against misinformation and the promotion of information integrity in the digital age.

PROPOSED SYSTEM

The proposed system represents a significant advancement in the field of fake news detection by leveraging deep learning architectures tailored specifically for textual content analysis. Focusing solely on the textual aspects of news articles, the system aims to discern the authenticity of information, disregarding auxiliary factors such as graphs, social network analysis, or images (Khanam et al., 2021). Through the evaluation of three distinct deep learning architectures—LSTM, CNN, and BERT—using two prominent datasets, namely the Fake News Corpus (FNC) and TI-CNN, the system's efficacy is comprehensively examined.

The LSTM-based architecture demonstrates promising results, achieving an accuracy of 91% FNC dataset. Long Short-Term Memory (LSTM) networks are renowned for their ability to capture sequential dependencies within data, making them particularly suitable for analyzing text (Lee et al., 2021). The architecture's respectable performance indicates its proficiency in discerning patterns and nuances within textual content, albeit with slightly lower accuracy on the FNC dataset, possibly due to its specific characteristics or biases.

In contrast, the CNN-based architecture excels in accuracy, attaining 97% on the TI-CNN dataset and 82% on the FNC dataset. Convolutional Neural Networks (CNNs) are adept at capturing spatial hierarchies within data, making them well-suited for tasks such as image recognition and, as demonstrated here, textual analysis (Rodríguez & Iglesias, 2019). The architecture's superior performance underscores the effectiveness of CNNs in extracting meaningful features from textual input, resulting in robust discrimination between fake and genuine news articles across both datasets. The BERT-based architecture, leveraging Bidirectional Encoder Representations from Transformers (BERT), achieves a commendable accuracy of 97% on the TI-CNN dataset. However, its performance on the FNC dataset, with an accuracy of 76%, falls short compared to the CNN-based architecture. BERT's strength lies in its contextual understanding of language, enabled by its bidirectional attention mechanism (Mridha et al., 2021). Despite its remarkable performance on the TI-CNN dataset, the architecture's performance on the FNC dataset suggests potential challenges in generalizing its learned representations to diverse textual domains or datasets.

APPROACH

The foundation of any machine learning endeavor lies in the quality and relevance of the data. In the context of combating fake news, this entails collecting a diverse range of news articles from various sources. However, raw data often contains noise, inconsistencies, and biases that can hinder model performance. Hence, a crucial initial step involves meticulous data preprocessing.

This preprocessing stage involves several tasks:

Cleaning

- Removing HTML tags, special characters, and irrelevant metadata to ensure uniformity and consistency across the dataset.

Normalization

- Standardizing text by converting to lowercase, removing punctuation, and handling contractions to streamline the text processing pipeline.

Tokenization

- Segmenting text into individual words or tokens, facilitating subsequent analysis and feature extraction.

Stop word Removal

- Eliminating common words (e.g., "and," "the") that carry little semantic meaning and may introduce noise into the model.

Lemmatization/Stemming

- Reducing words to their base forms to consolidate variations of the same word (e.g., "running" and "ran" both becoming "run").
- Furthermore, to optimize model performance, hyperparameters must be carefully tuned. Bayesian Optimization provides an efficient method for this task, iteratively exploring the hyperparameter space to maximize model performance while minimizing computational resources.

TRAINING DEEP NEURAL NETWORKS

Once the data is cleaned and preprocessed, the next step involves training deep neural networks (DNNs) to categorize news articles. Given the nuanced nature of fake news detection, DNNs offer the capacity to learn intricate patterns and relationships within the data.

Data Training Process

The training process typically involves:

Architecture Selection

- Choosing appropriate neural network architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), or their variants like LSTM (Long Short-Term Memory) networks based on the nature of the data and task requirements.

Feature Extraction

- Transforming text data into high-dimensional feature vectors using techniques like word embeddings (e.g., Word2Vec, GloVe) to capture semantic meaning and context.

Model Optimization:

- Employing optimization algorithms (e.g., stochastic gradient descent, Adam) to iteratively update model parameters and minimize the loss function.

TESTING ON DATASETS

To assess the efficacy of the trained models, rigorous testing on relevant datasets is essential. For categorical classification, the TI-CNN dataset provides a benchmark for evaluating the model's ability to categorize news articles into multiple types. Conversely, for binary classification (i.e., distinguishing between true and fake news), the FNC dataset serves as a standard test bed.

During testing, the following steps are undertaken:

Dataset Splitting

- Dividing the dataset into training, validation, and test sets to evaluate model performance while avoiding overfitting.

Performance Metrics

- Utilizing a range of evaluation metrics including accuracy, precision, recall, F1-score, and area under the ROC curve (AUC) to quantify the model's performance across different dimensions.

Cross-Validation

- Employing cross-validation techniques to ensure the robustness of the model by training and evaluating it multiple times on different subsets of the data.

Evaluation and Analysis

Beyond numerical metrics, a comprehensive evaluation involves qualitative analysis and exploration of the data. This includes:

Data Analysis

- Investigating the characteristics and distribution of the Fake News Corpus, TI-CNN Dataset, and Getting Real About Fake News Dataset to gain insights into the nature of fake news and its dissemination.

Human Evaluation

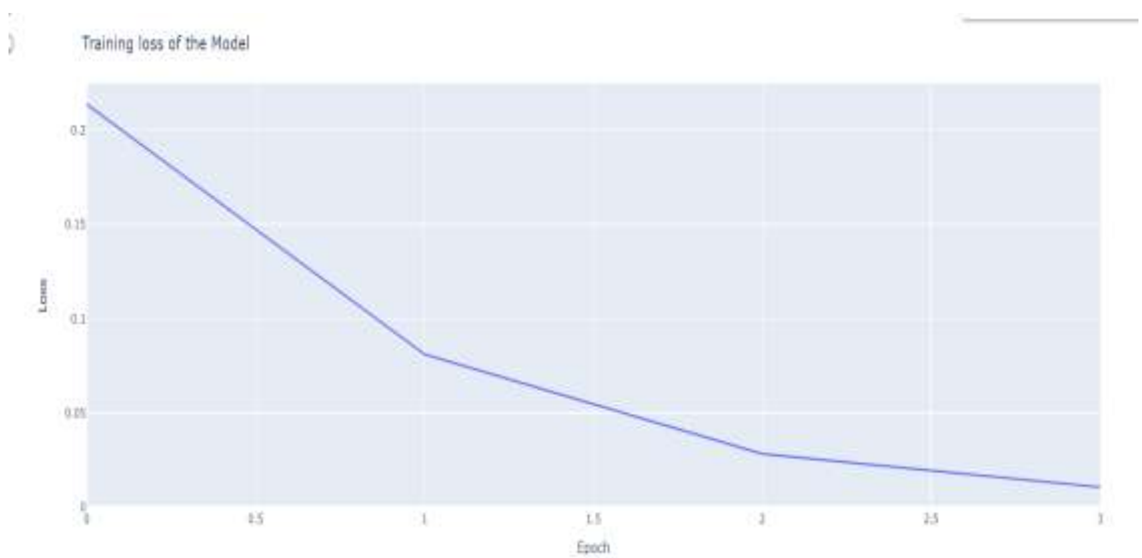
- Augmenting quantitative metrics with human evaluation to assess the quality of model predictions and identify potential misclassifications or biases.

Visualization

- Visualizing model predictions, decision boundaries, and feature importance to gain interpretability and insights into the model's decision-making process.

RESULTS

- Combining LSTM, BERT, and CNN can offer a more comprehensive understanding and detection of fake news by leveraging their unique strengths.
- These advanced models can be optimized for real-time analysis, allowing for timely interventions against the spread of misinformation.
- Despite challenges, the versatility of LSTM, BERT, and CNN architectures suggests potential applicability across various domains beyond fake news detection
- Accuracy Score by CNN-87.87%
- Accuracy Score by LSTM-86.66%
- Here, we plotted training loss curve of the model



CONCLUSION

The study concludes by applying a data augmentation process to each dataset using the back-translation method to alleviate class imbalance. Tests demonstrate the positive impact of data augmentation on model performance consistency. Various deep learning methods, including CNN, Bidirectional LSTM, and ResNet, are evaluated using different datasets. Bidirectional LSTM emerges as the top performer across all test datasets, aided by pre-trained word embeddings such as GloVe and fastText (Kong et al., 2020). The combination of deep learning methods with popular word embeddings is thoroughly examined, with publicly accessible datasets undergoing cleansing, augmentation, and preprocessing processes. Deep learning models like LSTM, BERT, and CNN can be scaled to handle large datasets, making them adaptable to the growing volume of online content. While technology plays a pivotal role, it's essential to balance innovation with ethical considerations, ensuring fairness, transparency, and respect for user privacy.Continued advancements in deep learning and NLP techniques promise even more sophisticated and accurate fake news detection systems in the future.

REFERENCES

1. Hiramath, C. K., & Deshpande, G. C. (2019, July). Fake news detection using deep learning techniques. In 2019 1st International Conference on Advances in Information Technology (ICAIT) (pp. 411-415). IEEE.
2. Khanam, Z., Alwasel, B. N., Sirafi, H., & Rashid, M. (2021, March). Fake news detection using machine learning approaches. In IOP conference series: materials science and engineering (Vol. 1099, No. 1, p. 012040). IOP Publishing.
3. Kong, S. H., Tan, L. M., Gan, K. H., & Samsudin, N. H. (2020, April). Fake news detection using deep learning. In 2020 IEEE 10th symposium on computer applications & industrial electronics (ISCAIE) (pp. 102-107). IEEE.
4. Lee, D. H., Kim, Y. R., Kim, H. J., Park, S. M., & Yang, Y. J. (2019). Fake news detection using deep learning. Journal of Information Processing Systems, 15(5), 1119-1130.
5. Mridha, M. F., Keya, A. J., Hamid, M. A., Monowar, M. M., & Rahman, M. S. (2021). A comprehensive review on fake news detection with deep learning. IEEE access, 9, 156151-156170.
6. Rodríguez, Á. I., & Iglesias, L. L. (2019). Fake news detection using deep learning. arXiv preprint arXiv:1910.03496.