Term Paper I

Critiqued By: Moxaben Bhupatbhai Zalawadia

Net ID: MXZ210014

Title of the Paper: Characterization, Classification and Detection of Fake News in Online Social Media Networks

Author of the Paper: Xavier Jose; S.D Madhu Kumar; Priya Chandran

Publication: IEEE Network

Abstract:
The pervasive spread of fake news in Online Social Media (OSM) networks has become a significant challenge, leading to the spread of misinformation and posing potential harm to individuals and society. Detecting and combating fake news requires a comprehensive understanding of its characteristics, effective classification techniques, and reliable detection mechanisms. This paper focuses on the characterization, classification, and detection of fake news in OSM networks. The paper highlights the impact of fake news in OSM networks on users and society and discusses categories like satire/parody, manipulation, imposter content, sponsored content, false connection, misleading, false context, and fabricated content. It focuses on classification techniques for identifying fake news in OSM networks. Finally, it addresses the detection of fake news in OSM networks. A two-phase solution is proposed for detecting false connection and fabricated content in tweets containing news articles. The paper describes the design of the proposed solution. In conclusion, this paper contributes to the field of fake news detection in online social media networks by proposing an approach that considers both news content and social context features.

Section 1: Introduction

The rise of online social media networks has revolutionized the way information is shared and consumed. However, alongside the benefits of instant connectivity and easy access to information, there is a growing concern about the proliferation of fake news or misinformation in online social media networks. Fake news refers to deliberately fabricated or misleading information presented as factual news. It has the potential to mislead and manipulate users, leading to distorted perceptions and actions. The widespread dissemination of fake news poses significant challenges to individuals, society, and even democratic processes. Fake news creators often utilize this medium to spread misinformation for personal, economic, or political gains. Additionally, malicious user accounts, such as social bots, play a role in publishing and amplifying fake news within online social media networks. While traditional methods of identifying fake news in traditional news media primarily rely on analysing the content of the news itself, the detection of fake news in online social media networks presents a different challenge. In this context, additional factors come into play, specifically the social context features that can complement the analysis of news content. These social context features capture various aspects, including user behaviour, engagement patterns, and the overall social environment surrounding the news article.
In general, various studies and efforts related to the detection of fake news on social media platforms are examined and assessed, while classifying fake news within online social media networks. The discussions focus on solution designs for tackling this issue, which are subsequently implemented and analysed along with the results and further discussion.

Following is a summarized version of the related studies:

Various ML models like Logistic Regression, Naive Bayes, Random Forests, Decision Trees, and Support Vector Machines are used for text classification to detect fake news. SVMs have shown higher accuracy compared to other methods. Fake news titles tend to be longer, while the body contains simpler language compared to real news. Clickbait headlines are commonly found in fake news and can be used for detection. Linear SVM with unigram features performs well. Hybrid deep learning models combining convolutional and recurrent neural networks improve classification. Multi-modal approaches that analyse textual and visual features are effective. Geometric deep learning models using information propagation show promise in detecting fake news.

Section 2: Approach

I. CLASSIFICATION OF FAKE NEWS

The paper addresses various types of fake news in online social media networks. These include satire/parody articles that can mislead readers when taken out of context, image/video manipulation that creates false associations or uses edited visuals to support claims, imposter content that impersonates legitimate news sources, sponsored content disguised as unbiased media, false connection between headlines and content, misleading content that selectively reports information to shape narratives, false context where real facts are shared with incorrect background information, and fabricated content that mimics legitimate articles but contains biased language. The proposed solution aims to detect and address these types of fake news in online social media networks.

II. DESIGN OF SOLUTION

The solution is presented for identifying false connections and fabricated content in news articles shared on Twitter. The system operates in two phases. Initially, it extracts significant words and sentences from the news article using NLP techniques and searches for them in reputable sources. A credibility score is then assigned based on the alignment with credible sources. In the subsequent phase, machine learning models are employed to classify the tweet. This approach combines NLP, source querying, and ML to extract information, evaluate credibility, detect false connections, and identify fabricated content in tweets.

A stance detection model is used to determine the stance of a text towards a target text. It classifies the stance of the news body text into four categories: Agree, Disagree, Discusses, and Unrelated. The model takes the news headline and body text as input and analyses their relationship. The goal is to assess whether the news body text agrees, disagrees, discusses the same topic, or is unrelated to the news headline.

Also, a fabricated content classifier is used to distinguish between legitimate news articles and fabricated content. Fabricated content mimics real news but differs in writing style, structure, and language. By utilizing linguistic features and analysing writing patterns and sensational language, the classifier categorizes the news text as either fabricated content (class 1) or not fabricated content (class 0). Its purpose is to identify deceptive news stories and combat the spread of misinformation.

III. IMPLEMENTATION

During the implementation phase, Stance Detection Model is trained using the Fake News Challenge dataset. Preprocessing includes text conversion, punctuation and stop word removal, and TF-IDF feature extraction. Multiple classifiers are trained, such as Logistic Regression, Decision Tree, Random Forest, Multinomial Naive Bayes, and Support Vector Machine.

The Fabricated Content Classifier is also trained using the same Fake News dataset. Preprocessing involves text conversion, punctuation removal, and word lemmatization. Word embedding is applied, and the model architecture includes an Embedding layer, LSTM layer, and dense output layer. Training involves binary classification using binary cross entropy and the ADAM optimization algorithm.

Overall, the approach includes preprocessing, feature extraction, and training classifiers to detect stance and identify fabricated content in news articles.

Section 3: Pros and Cons of the Paper discussed in section 2

Pros:

1. Comprehensive understanding: The approach considers various aspects of fake news, leading to a more holistic understanding of its characteristics and detection mechanisms.
2. Social context consideration: Social context features are incorporated, enhancing the accuracy of fake news detection.
3. Integration of techniques: Multiple techniques such as NLP, source querying, ML, stance detection, and fabricated content classification are combined, increasing the robustness of the approach.
4. High accuracy: The proposed solution achieved high accuracies in detecting fake news, indicating its effectiveness.

Cons:

1. Limited focus: The approach primarily focuses on detecting false connections and fabricated content, potentially overlooking other types of fake news like image/video manipulation.
2. Dataset limitations: Reliance on existing datasets may be a constraint, especially for regional languages, affecting the generalizability of the approach.
3. Lack of comparative analysis: The paper does not provide a direct comparison with alternative approaches, making it difficult to evaluate its relative strengths and weaknesses.
4. Implementation challenges: The paper does not explicitly address potential challenges in deploying and scaling the proposed system.

Section 4:

Summary:

The spread of fake news on social media platforms has significantly increased, leading to various negative consequences. To address this issue, this paper explores the characteristics and types of fake news in online social media networks and proposes an effective solution for detecting fake news, specifically focusing on identifying false connections and fabricated content. The proposed solution consists of two main components: a stance detection model and a fabricated content classifier. The stance detection model achieved an impressive accuracy of 90.37% using Logistic Regression, while the fabricated content classifier attained an accuracy of 93.46% with Bi-directional LSTM.

Direction: How can we extend the paper discussed in section 2

To extend the paper on fake news detection in online social media networks, consider incorporating multimodal features, exploring user-centric approaches, addressing evolving fake news techniques, adapting the research to regional languages and social contexts, developing real-time detection mechanisms, ensuring robust evaluation, and addressing ethical considerations.

References:

1. Xavier Jose; S.D Madhu Kumar; Priya Chandran. Characterization, Classification and Detection of Fake News in Online Social Media Networks.
2. Jamal Abdul Nasir, Osama Subhani Khan, and Iraklis Varlamis. Fake news detection: A hybrid cnn-rnn based deep learning approach. International Journal of Information Management Data Insights, 1(1):100007, 2021.