Fake News Spread Analysis through Social Media Networks

Yash Dani
York University
Toronto, Canada
yashdani@my.yorku.ca

Eyad Mahmoud
York University
Toronto, Canada
eyad66@my.yorku.ca

Waleed Samouh
York University
Toronto, Canada
waleed.samouh@yorku.ca

Malik Samouh
York University
Toronto, Canada
maliktj@my.yorku.ca

ABSTRACT

The proliferation of fake news on social media platforms has become a significant concern due to its potential to mislead the public, influence political events, and cause social unrest. Fake news spreads rapidly and widely, often outpacing factual information, which can result in widespread misinformation and societal harm. This misinformation can affect elections, health decisions, financial markets, and even social stability. Therefore, understanding the mechanisms of fake news spread and developing effective detection and mitigation strategies is crucial for maintaining the integrity of information in the digital age. This project utilizes advanced machine learning models and network analysis to detect and analyze the spread of fake news, offering insights that could help mitigate its impact and enhance the reliability of online information.

KEYWORDS

fake news, social media, network analysis, misinformation, machine learning, natural language processing, spread modeling

1 INTRODUCTION

The rise of fake news on social media poses significant risks, including misleading the public, influencing political events, and causing social unrest. Fake news spreads rapidly, often outpacing factual information and leading to widespread misinformation. This project investigates the mechanisms of fake news dissemination on social media networks. By analyzing the spread and impact of fake news, we aim to develop strategies for detecting and mitigating its effects, thereby enhancing the reliability of online information.

1.1 Problem Definition

The primary goal of this project is to conduct a comprehensive study on the spread and propagation of news within social media networks and determine if fake news spreads more on Twitter. We focus on addressing the following problems:

1. **Cleaning and Labeling the Dataset**:
   - Categorize news into real, fake, rumor unverified, and rumor verified.
   - Ensure data integrity by identifying and removing duplicates.
2. **Visualizing the Network of News Propagation**:
   - Construct interaction networks of tweets to visualize the spread.
   - Display the initial burst of dissemination versus the formation of longer chains.
3. **Analyzing the Characteristics and Patterns of News Spread**:
   - Compare the speed and reach of fake news versus real news.
   - Investigate the depth of propagation and the cascading effect.
   - Examine attributes like closeness, betweenness, and reaction time.
   - Analyze emotions and sentiments from the tweets and how they affect the spread.

1.2 Key Questions

We aim to answer the following key questions in this project:

(1) **Does fake news spread more than real news on Twitter?**
(2) **Do people with a higher following spread fake news?**
(3) **How do emotions affect reactions towards tweets?**

2 RELATED WORK

Several studies have explored fake news detection and the dynamics of its spread on social media networks.

Papagelis et al. [1] analyzed cascading behavior in the blogosphere, examining 30 million active blogs and 700 million posts. They emphasized the importance of linking behavior and factors influencing information spread in online communities.

Shu et al. [2] provided a comprehensive overview of fake news detection techniques using data mining approaches, highlighting the role of machine learning in identifying deceptive information based on linguistic cues and user interaction patterns.

Vosoughi et al. [3] investigated the spread of true and false news on Twitter, finding that false news spreads faster, farther, and more broadly than true news. This study highlighted the psychological and social factors contributing to the virality of misinformation.

Budak et al. [4] proposed methods to limit the spread of misinformation through intervention strategies in social networks, focusing on identifying key nodes for intervention and simulating the impact of various strategies.

These studies provide a foundation for our work, allowing us to build on their findings. By analyzing our dataset, we aim to uncover key trends and address our research questions, informed by the insights and methodologies of previous research.

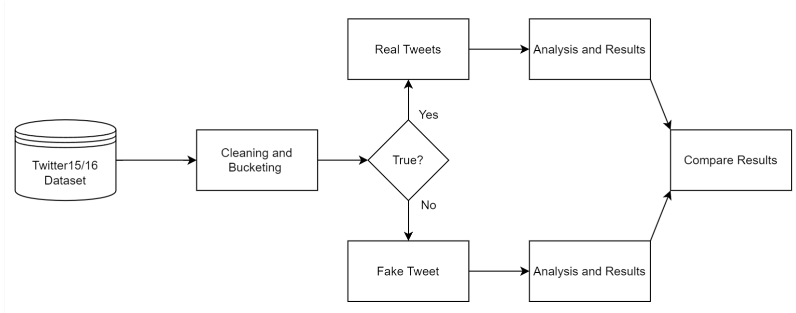

Figure 1: High-level framework of the data collection process. (1) dataset, (2) cleaning and labeling, (3) decide if tweet is fake or real, (4) analysis for each, (5) comparing results

3 METHODOLOGY

3.1 Data Collection

We utilized pre-existing datasets, Twitter15 and Twitter16, for this study. These datasets were identified through the research conducted by Ma et al. [**?**], which can be accessed here. The datasets were collected using a service account on Twitter via the Twitter API. The datasets contain a collection of source tweets along with their corresponding sequences of retweet user IDs. For our study, we focused on tweets labeled as "true" and "fake" to serve as ground truth.

Additionally, to supplement the dataset, user profile information was crawled using the Twitter API, accounting for 92.76% and 94.58% of user data for Twitter15 and Twitter16, respectively. Missing user information was filled using the mean value of other user features within the same propagation structure. Overall, the datasets provide a comprehensive foundation for analyzing the spread of fake news on social media platforms.

The datasets can be accessed on DropBox.

3.2 Data Labeling and Analysis

The collected data was labeled to categorize tweets into real and fake news. The labeling process involved:

(1) **Data Conversion and Loading**:
- Converted raw tweet data into a structured format suitable for analysis.
- Loaded the data efficiently to ensure comprehensive coverage.

(2) **Data Preprocessing**:
- Cleaned and preprocessed the data, removing duplicates and irrelevant information.
- Categorized tweets into labels such as 'true', 'false', 'unverified', and 'non-rumor', and stored them in respective directories.

(3) **Label Merging**:
- Merged the labeled datasets to form a comprehensive dataset for analysis.

By analyzing these categories, we aimed to uncover key trends and insights that would help answer our research questions and inform our understanding of fake news spread on social media.

3.3 Network Construction

We constructed interaction networks to analyze the spread of news on Twitter. The key elements include:

- **G(N, E): Directed Graph** - This graph represents the direction of information flow.
- **N: Nodes** - Each node represents a user who interacted with a tweet.
- **E: Edges** - Each edge indicates an interaction between users.
- **Source Node** - The source node is the user who posted the original tweet.
- **Propagation Trees** - We constructed separate graphs for each tweet to analyze its propagation.

The network construction process involved extracting edges representing user interactions, constructing directed graphs from these edges, and storing the graphs for further analysis. This approach enabled us to visualize and understand the patterns of news spread on Twitter, providing insights into the dynamics of real and fake news propagation.

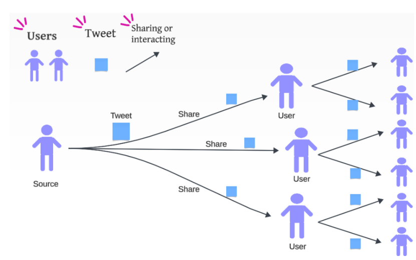

Figure 2: Example of a Network Construction for a Tweet

3.4 Visualization of Graph

The visualization of the network graph for a randomly selected fake news tweet reveals key insights into its propagation dynamics.

3.4.1 ***General Observations.*** The central hub, representing the source user who initially tweeted the fake news, is prominently located at the center with numerous edges radiating outward. This central positioning underscores the significant influence of the source user in disseminating information. The extensive number of edges indicates that the initial tweet was widely shared, reaching a broad audience quickly. The radial pattern exemplifies a cascading effect, where the information spreads outward in successive layers, resembling a ripple effect. This cascading propagation is evident as the information flows through various intermediate users before reaching the outermost nodes. Some nodes have notably higher connectivity, suggesting that certain users played a critical role in further spreading the fake news. Identifying these key nodes is crucial for potential intervention strategies. The overall network density, characterized by numerous interconnected nodes, highlights the extensive reach and potential impact of fake news within the social media network. This visualization provides valuable insights into the structural and behavioral dynamics of fake news dissemination, emphasizing the cascading nature of its spread.

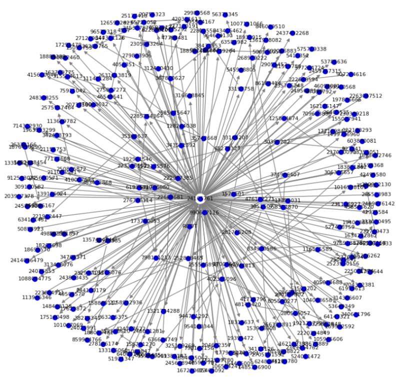

Figure 3: Example of a Network Construction for a Tweet

3.5 Graph Analysis

these analyses were used to study the network

3.5.1 Number of Fake Tweets vs Number of Real Tweets

This metric compares the total number of fake tweets to the total number of real tweets in the dataset. It provides an overview of the distribution of fake and real news within the collected data.

Findings:

- Total unique fake tweets: 20,069
- Total unique real tweets: 16,351

We have observed that there are more fake tweets than real tweets in our sample. This indicates a higher prevalence of fake news on Twitter. Specifically, there are approximately **22.7% more fake tweets than real tweets.** This significant difference gives us early insights on our question.

3.5.2 Average Number of Nodes

The average number of nodes represents the average number of users interacting with the news.

Findings:

- **Twitter15:**
  - Average number of nodes for fake news: 40
  - Average number of nodes for real news: 34
- **Twitter16:**
  - Average number of nodes for fake news: 30
  - Average number of nodes for real news: 21

Analysis: We have observed that the average number of nodes (users interacting with the news) is consistently **higher for fake news** compared to real news in both datasets and over time. This indicates that fake news tends to engage more users compared to real news. More nodes per tweet suggest that fake news reaches and is shared by a larger number of users.

3.5.3 Average Number of Edges

The average number of edges indicates the average number of shares or interactions with the tweet.

Findings:

- **Twitter15:**
  - Average number of edges for fake news: 45
  - Average number of edges for real news: 37
- **Twitter16:**
  - Average number of edges for fake news: 35
  - Average number of edges for real news: 24

Analysis: We have observed that the average number of edges (shares or interactions) is consistently higher for fake news compared to real news in both datasets. This suggests that fake news tends to have more interactions and shares, indicating a wider spread compared to real news.

3.5.4 Average Cascade Size

The average cascade size refers to the number of nodes reached by the news from the initial source. It measures the breadth of the spread. A larger cascade size indicates a wider dissemination of the news.

Findings: The average cascade size for fake news was approximately 25.5% higher than that for real news across both Twitter15 and Twitter16 datasets.

Analysis: We have observed that the average cascade size is consistently higher for fake news compared to real news in both datasets. For Twitter15and Twitter16. On average, fake news had a 25.5% higher cascade size overtime compared to real news, indicating that fake news tends to spread more widely, reaching a larger number of nodes from the initial source.

3.5.5 Average Tree Depth

The average tree depth measures how deeply information penetrates the network. A greater depth means the news spreads through more layers of users before reaching the end nodes, indicating extensive propagation.

Findings: The average tree depth for fake news was very close to that of real news, with fake news being approximately 1.5% higher on average across both Twitter15 and Twitter16 datasets.

Analysis: We have observed that the average tree depth is very similar for fake and real news in both datasets overtime. Overall, the average tree depth for fake news was slightly higher by approximately 1.5%. This suggests that fake news may penetrate the network slightly deeper, spreading through an extra layer of users before reaching the end nodes.

3.5.6 Average Propagation Delay

The average propagation delay is the time taken for the news to spread across the network. A longer delay means the news takes more time to reach the majority of users.

Findings: The average propagation delay for fake news was significantly higher by approximately 96% compared to real news over time.

Analysis: The propagation delay was nearly double for fake news over time. This indicates that fake news takes a longer time to spread across the network, suggesting that it has a more extended propagation period, reaching more users over a more extended period.

3.5.7 Average Reaction Time

The average reaction time measures how quickly users respond to the news after it is posted. A shorter reaction time indicates faster engagement from the audience.

$$\text{Average Reaction Time} = \frac{\sum \text{Reaction Times}}{\text{Number of Interactions}}$$

Findings: The average reaction time for fake news was significantly higher by approximately 121% compared to real news over time.

Analysis: We observed that the average reaction time for fake news is much higher compared to real news. this shows that people look into the news to check weather its true or not and tend to hesitate before sharing and reacting to fake news.

3.5.8 Betweenness Centrality

Betweenness centrality measures the extent to which a node lies on the shortest path between other nodes. High betweenness centrality indicates a user plays a critical role in information dissemination within the network. High betweenness indicates a high following, meaning the person posting or sharing could be an important figure or an average Joe.

Calculation:

For a given graph $G(V, E)$, where V is the set of vertices and E is the set of edges, the betweenness centrality $C_B(v)$ of a vertex v is calculated as:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where: - σ_{st} is the total number of shortest paths from vertex s to vertex t. - $\sigma_{st}(v)$ is the number of those shortest paths that pass through vertex v.

To normalize the betweenness centrality, we divide by the number of pairs of nodes not including the node itself:

$$C'_B(v) = \frac{2C_B(v)}{(n-1)(n-2)}$$

where n is the number of nodes in the graph.

Findings: On average, betweenness centrality for fake news and real news is almost the same.

Analysis: We observed that the average betweenness centrality is almost the same for fake and real news over time. This suggests that users disseminating fake or real news are more often on the critical paths between other users, meaning they are people with high followings and are influential.

3.6 Content Analysis

these analyses were used to study the content of the tweets

3.6.1 Sentiment and Emotion Analysis

Sentiment analysis involves evaluating the sentiment expressed in tweets, categorizing them as positive, negative, or neutral. Emotion detection goes beyond sentiment analysis to identify specific emotions such as joy, anger, sadness, and fear in the tweets. This deeper analysis provides insights into the emotional triggers that might cause users to share or react to tweets. Understanding these emotional responses can help in identifying the factors that contribute to the rapid spread of fake news.

Mathematical Explanation: The sentiment and emotion analysis process can be described mathematically as follows:

1. **Tweet Representation:** Each tweet T_i is represented as a string of text. Let $T = \{T_1, T_2, \ldots, T_n\}$ be the set of all tweets in the dataset.

2. **Sentiment and Emotion Detection Models:** Pre-trained sentiment analysis model S and emotion detection model E are used to evaluate the emotional content of each tweet. The models S and E are functions that take a tweet T_i as input and output sentiment and emotion labels along with their confidence scores:

$$S(T_i) = (\text{label}(T_i), \text{score}(T_i))$$

$$E(T_i) = (\text{emotion}(T_i), \text{emotion\_score}(T_i))$$

where $\text{label}(T_i)$ is the sentiment label (e.g., positive, negative, neutral), $\text{score}(T_i)$ is the sentiment confidence score, $\text{emotion}(T_i)$ is the emotion label (e.g., joy, anger, sadness, fear), and $\text{emotion\_score}(T_i)$ is the emotion confidence score.

3. **Result Compilation:** For each tweet T_i, the results from the sentiment and emotion analysis are compiled into a structured format. This includes the tweet ID, content, sentiment label, sentiment score, emotion label, and emotion score. The compiled results are stored in a JSON file for further analysis.

The overall sentiment and emotion analysis process for a dataset of tweets can be summarized as:

$$\text{Results} = \{(T_i, S(T_i), E(T_i)) \mid T_i \in T\}$$

3.6.2 Emotion analysis

Emotion detection goes beyond sentiment analysis to identify specific emotions such as joy, anger, sadness, and fear in the tweets. This deeper analysis provides insights into the emotional triggers that might cause users to share or react to tweets.

Findings: We observed the following from our emotion detection analysis:

- For real news, the top emotions identified were **disgust, sadness, and anger.**
- For fake news, the top emotions identified were **disgust, joy, and fear.**

Analysis: The charts show the average cascade size by emotion label for both real and fake news in the Twitter15 dataset. We observed that:

- **Disgust** was the top emotion for both real and fake news, indicating that most tweets on twitter evoke a reaction of disgust as most news would.
- **Joy** was also a significant emotion for fake news, ranking second. This suggests that tweets evoking joy are more likely to be shared, contributing to the viral nature of fake news.
- **Anger** was also ranked high in fake news meaning people use fake news for propaganda to to push some agenda to make people fear them or others.
- **Fear** and **sadness** were more prominent in real news compared to fake news, indicating that these emotions often trigger a response as real news is reality and straight to the point.

Overall, emotions play a crucial role in the propagation of tweets, with some emotions helping news spread more.

3.6.3 Topic Modeling

Topic modeling was performed to identify the main topics discussed in the tweets using Latent Dirichlet Allocation (LDA). This technique categorizes the content of tweets, enabling us to understand the themes associated with fake or real news. By analyzing prevalent topics, we gained insights into the subjects that capture user interest and contribute to the spread of information. The topic modeling process involved preprocessing the tweets, creating a dictionary and corpus, and extracting topics. This analysis is crucial for uncovering the key themes driving the dissemination of information on social media.

3.7 Cascade Triggering Analysis

Cascade triggering analysis aims to predict the spread of information on social media by examining the features of the network and content. We employed two models: a Random Forest Model and an Advanced Model.

3.7.1 Random Forest Model

The Random Forest Model utilizes an ensemble of decision trees to predict cascade size and depth. The methodology for implementing this model includes several key steps:

Data Collection and Preprocessing:

- **Data Sources:** Data were collected from Twitter datasets (Twitter15 and Twitter16), including labels such as 'true', 'false', 'unverified', and 'non-rumor'.
- **Graph Features:** For each tweet, we constructed a graph where nodes represent users and edges represent interactions (e.g., retweets). Key graph features extracted include the number of nodes, number of edges, mean in-degree, mean out-degree, degree centrality, and clustering coefficient.
- **Content Features:** Sentiment scores, sentiment labels, and emotion labels were extracted from the content analysis.
- **Combining Features:** Graph and content features were combined to form a comprehensive feature set for each tweet.
- **Encoding Categorical Variables:** Sentiment and emotion labels were encoded using the LabelEncoder from scikit-learn.

Model Training and Evaluation:

- **Training and Test Split:** The dataset was split into training (80%) and testing (20%) sets.
- **Model Training:** A Random Forest Regressor was trained on the training set using the combined features.
- **Evaluation Metrics:** The model was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), R-squared (R^2), and Cross-Validation MSE.

Implementation Details:

- **Random Forest Parameters:** The Random Forest Regressor was configured with default parameters and a random state of 42 to ensure reproducibility.
- **Feature Importance:** The importance of each feature in predicting the cascade was determined by the Random Forest model.

3.7.2 Advanced Model

The Advanced Model incorporates additional machine learning algorithms to improve the prediction of cascade sizes and depths. This section describes the methodology for implementing the advanced models:

Data Collection and Preprocessing:

- **Data Sources and Features:** Similar to the Random Forest Model, we used Twitter datasets (Twitter15 and Twitter16) and extracted graph and content features.
- **Feature Engineering:** The same set of combined features was used, including graph features (e.g., degree centrality, clustering coefficient) and content features (e.g., sentiment scores).
- **Encoding Categorical Variables:** Sentiment and emotion labels were encoded using the LabelEncoder.

Model Training and Evaluation:

- **Training and Test Split:** The dataset was split into training (80%) and testing (20%) sets.
- **Advanced Models:** Two additional models were used:
  - **Ridge Regression:** This linear model was used as a baseline for comparison.
  - **MLP Regressor:** A Multi-Layer Perceptron (neural network) was employed for more complex non-linear predictions.
- **Evaluation Metrics:** The models were evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), R-squared (R^2), and Cross-Validation MSE.

Implementation Details:

- **Ridge Regression Parameters:** The Ridge Regressor was configured with a random state of 42.
- **MLP Regressor Parameters:** The MLP Regressor was configured with a hidden layer size of (100,), activation function 'relu', and solver 'adam'.
- **Comparison and Analysis:** The performance of the advanced models was compared to the Random Forest model to determine improvements in prediction accuracy.

4 EXPERIMENTS AND RESULTS

4.1 Longest Chain Experiment

In this experiment, we identified the top 3 files in the fake news dataset by depth and extracted the single longest chain. This analysis helps us understand the propagation dynamics and cascading effects of fake news. The selected viral tweet, which serves as propaganda towards Vladimir Putin, showcases the extensive depth and complex propagation patterns. we used the Depth-First Search (DFS) to find the longest-chain per dataset.

Algorithm 1 Depth-First Search (DFS) for Longest Chain

Input: Graph $G(N, E)$ with nodes N and edges E
Output: Depth of each tree in the graph
function DFS(v, $depth$)
    $visited[v] \leftarrow$ True
    $max\_depth \leftarrow depth$
    for all $u \in neighbors(v)$ **do**
        if $visited[u] = False$ **then**
            $current\_depth \leftarrow$ DFS$(u, depth + 1)$
            $max\_depth \leftarrow \max(max\_depth, current\_depth)$
        end if
    end for
    return max_depth
end function

Initialize $visited[N] \leftarrow$ False for all N
$longest\_chain \leftarrow 0$
for all $v \in N$ **do**
    if $visited[v] = False$ **then**
        $chain\_depth \leftarrow$ DFS$(v, 0)$
        $longest\_chain \leftarrow \max(longest\_chain, chain\_depth)$
    end if
end for
return $longest\_chain$

4.1.1 Viral Tweet Analysis

The visualized graph of the viral tweet demonstrates the extensive reach and depth of the tweet within the network. The network exhibits a pronounced cascading effect, where information spreads through multiple layers of users, indicating a significant impact on the network's structure and behavior.

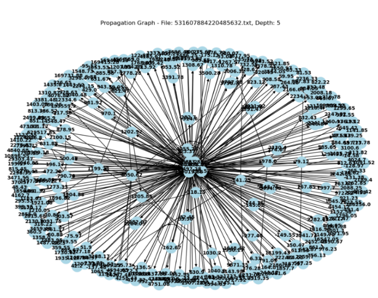

Figure 4: Propagation Graph of a Viral Tweet

Analysis: Figure 4 illustrates the propagation graph of a viral tweet. The extensive reach and multiple layers of users show a significant cascading effect, where the information spreads outward through successive interactions. This indicates a broad dissemination of the tweet, engaging a large number of users and facilitating the spread of misinformation.

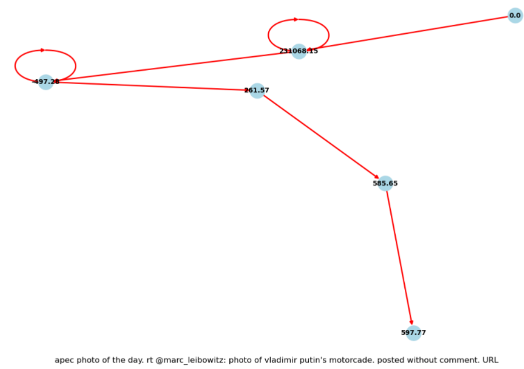

Figure 5: Analysis of the Longest Chain in the Propagation Graph

Analysis: Figure 5 shows the analysis of the longest chain within the propagation graph. The longest chain highlights the depth of the tweet's penetration into the network, reaching multiple layers of users. This deep penetration is indicative of the tweet's ability to engage users and prompt further dissemination, thereby amplifying its impact. The presence of loops in the graph either indicates users replying in a chain of the same tweet or retweeting. Retweeting, in particular, amplifies the reach of the tweet to the user's own audience, further spreading the information. this shows us how deep news spreads over a chain of users, visualizing how deep a single fake tweet can go.

4.2 Reaction Time and Emotion Analysis

In this experiment, we used sentiment analysis and emotion detection to categorize the emotions expressed in tweets. We then compared the reaction times for tweets that expressed fear against those that expressed other emotions. The analysis helps us understand how different emotions influence the speed of user interactions and shares.

Algorithm 2 Computing Average Reaction Times

function COMPUTEREACTIONTIMES(S, E_t)
  Initialize sets T_{E_t} and T_o to store reaction times for target and other emotions respectively
  Initialize sets I_{E_t} and I_o to store tweet IDs for target and other emotions respectively
  for all $e \in S$ **do**
    Extract tweet ID i and emotion label L_i
    if $L_i = E_t$ **then**
      Add i to I_{E_t}
    else
      Add i to I_o
    end if
  end for
  for all reaction time data files **do**
    for all lines in the file **do**
      Parse the parent and child tweet IDs and reaction time
      if parent ID is in I_{E_t} **then**
        Append reaction time to T_{E_t}
      else if parent ID is in I_o **then**
        Append reaction time to T_o
      end if
    end for
  end for
  Calculate the average reaction time for target emotion:

$$T_{E_t} = \frac{1}{|T_{E_t}|} \sum_{t \in T_{E_t}} t$$

  Calculate the average reaction time for all other emotions:

$$T_o = \frac{1}{|T_o|} \sum_{t \in T_o} t$$

  return T_{E_t} and T_o
end function

Mathematical Explanation: We measured the average reaction time for tweets categorized by different emotions using the following relationships:

- T_f: Average reaction time for the "fear" emotion
- T_o: Average reaction time for all other emotions

Given the data, we can express the relationships as:

$$T_f = \frac{1}{N_f} \sum_{i=1}^{N_f} t_{f,i}$$

$$T_o = \frac{1}{N_o} \sum_{i=1}^{N_o} t_{o,i}$$

Where:

- N_f is the number of reaction times for the "fear" emotion
- $t_{f,i}$ is the i-th reaction time for the "fear" emotion
- N_o is the number of reaction times for all other emotions
- $t_{o,i}$ is the i-th reaction time for all other emotions

The difference between the average reaction time for "fear" and the average reaction time for all other emotions can be calculated as:

$$\Delta T = T_f - T_o$$

Where ΔT represents the difference in average reaction times between the "fear" emotion and all other emotions.

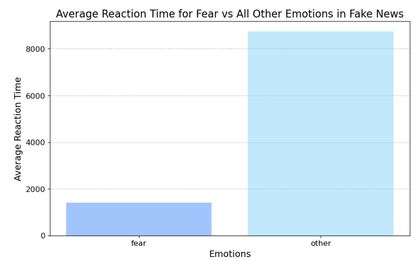

Figure 6: Average Reaction Time for Fear vs All Other Emotions in Fake News

Findings: The average reaction time for fake news expressing fear was significantly higher by approximately 121% compared to other emotions.

Analysis: We observed that tweets expressing fear had a significantly shorter reaction time compared to those expressing other emotions. This indicates that people react and share tweets more quickly when they are afraid or faced with fear. This finding suggests that when malicious propaganda is pushed on Twitter, it may cause fear and prompt quicker reactions from users. This rapid response to fear-driven content highlights the importance of addressing and mitigating the spread of such emotionally charged fake news.

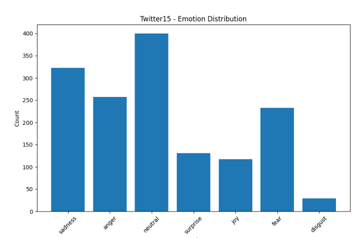

Figure 7: Twitter 15 emotion distribution

4.3 Topic Modeling

In this section, we discuss the results of topic modeling conducted on the Twitter15 and Twitter16 datasets. We used Latent Dirichlet Allocation (LDA) to identify key topics within the tweets. Additionally, word clouds were generated to visualize the most frequent words.

False News:

- **Twitter15:** Key topics included misinformation about events such as plane crashes and false alerts. Example topics:
  - *Topic 1:* "url", "talk", "Angela", "person", "photo"

– *Topic 2:* "url", "fukushima", "water", "ocean", "leak"
- **Twitter16:** Topics centered around sensationalist news, such as mass shootings and political misinformation. Example topics:
  – *Topic 1:* "url", "watch", "anonym", "list", "employe"
  – *Topic 2:* "url", "shoot", "mass", "syrian", "biolog"

True News:

- **Twitter15:** Topics often reflected real events and memorials, such as those for Paul Walker and soldiers. Example topics:
  – *Topic 1:* "url", "Paul", "walker", "di", "spider"
  – *Topic 2:* "url", "shot", "memori", "ottawa", "war"
- **Twitter16:** Real news focused on significant incidents and official statements. Example topics:
  – *Topic 1:* "url", "rainbow", "white", "hous", "color"
  – *Topic 2:* "url", "soldier", "ottawa", "break", "hostag"

Insights:

- False news topics tend to focus on sensationalism and misinformation, often with high emotional content.
- True news topics are more centered around factual events and memorials.
- Word clouds further highlighted the prominence of certain keywords, providing visual affirmation of the key topics identified.

4.4 Cascade Triggering Analysis

Data Preprocessing:

- Graph features such as the number of nodes, edges, degree centrality, and clustering coefficient were extracted.
- Content features included sentiment scores, sentiment labels, and emotion labels.
- Features were combined and encoded using LabelEncoder.

Model Training and Evaluation:

- The data was split into training (80%) and testing (20%) sets.
- A Random Forest Regressor was trained and evaluated using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), R-squared (R^2), and Cross-Validation MSE.

Results:

- **Twitter15:** MSE = 10.70, MAE = 1.38, R^2 = 0.64, Cross-Validation MSE = 22.88.
- **Twitter16:** MSE = 19.30, MAE = 1.86, R^2 = 0.58, Cross-Validation MSE = 81.06.
- Feature importances highlighted that node and edge counts, as well as sentiment scores, were significant predictors.

4.4.1 Advanced Model

The advanced model included Ridge Regression and MLP Regressor for improved prediction accuracy.

Model Training and Evaluation:

- Ridge Regression and MLP Regressor were used alongside Random Forest for a comprehensive analysis.
- Models were evaluated using the same metrics: MSE, MAE, R^2, and Cross-Validation MSE.

Results:

- **Twitter15:** MSE = 2.13, MAE = 0.90, R^2 = 0.61, Cross-Validation MSE = 2.45.
- **Twitter16:** MSE = 2.19, MAE = 0.92, R^2 = 0.57, Cross-Validation MSE = 2.60.
- The advanced models showed improved performance, particularly in reducing the error metrics compared to the Random Forest model.
- Feature importances from Ridge Regression indicated consistent relevance of graph features and sentiment scores.

Comparison and Insights:

- The advanced models generally outperformed the Random Forest model, indicating the value of more complex algorithms for this task.
- Insights from feature importance analysis suggest that both network structure and sentiment analysis are crucial for predicting cascade dynamics.
- Notably, the advanced models had significantly lower MSE and MAE values, highlighting their effectiveness.

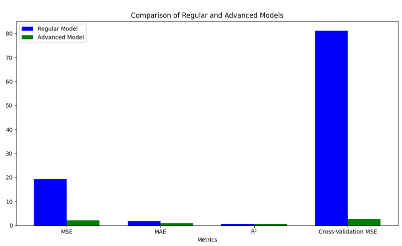

Figure 8: Result Comparison of both Models

5 DISCUSSION

5.1 Insights

Our analysis of fake news propagation on social media networks reveals several key insights:

- **Prevalence of Fake News:** Fake news is more prevalent on Twitter, comprising approximately 22.7% more tweets than real news.
- **User Engagement:** Fake news engages more users, consistently having a higher average number of nodes compared to real news.
- **Propagation Patterns:** Fake news spreads more widely, with an average cascade size 25.5% higher than real news.
- **Depth of Spread:** Fake news penetrates slightly deeper into the network, with an average tree depth 1.5% higher than real news.
- **Propagation Delay:** Fake news has nearly double the propagation delay of real news, indicating a more extended spread period.
- **Reaction Times:** Users take 121% more time to react to fake news, likely due to skepticism and the need for verification.
- **Emotional Engagement:** Emotions like disgust, joy, and fear significantly impact the spread of fake news. Tweets

expressing fear prompt faster reactions and wider dissemination.
- **Centrality Measures:** Influential users, with high betweenness and closeness centrality, play a crucial role in spreading both fake and real news.

These insights highlight the dynamics of fake news spread on social media and underscore the importance of effective detection and mitigation strategies, particularly for emotionally charged and influential content.

5.2 Possible Mitigations

Our analysis revealed that users typically take longer to react to fake news compared to real news, suggesting a period of hesitation and verification. This latency presents a window for potential intervention. To mitigate the spread of fake news, we propose the implementation of a fact-checking mechanism on platforms like Twitter. Upon detecting potentially fake news, the platform could label the content accordingly and restrict its sharing and retweeting capabilities.

Furthermore, our findings on cascade size and propagation patterns, illustrated in Figures 3 and 4, indicate that the largest burst of dissemination occurs shortly after the initial posting. By targeting this critical phase, the platform can significantly curb the rapid spread of misinformation. This approach not only leverages the observed reaction delay but also aims to contain the spread before it reaches a wider audience.

Additionally, enhancing user education on recognizing fake news and promoting critical evaluation of information sources can further strengthen the impact of these technical measures. By combining algorithmic interventions with user awareness initiatives, social media platforms can create a more informed and resilient user base, effectively reducing the prevalence and impact of fake news.

6 CONCLUSION

This study aimed to address three key questions regarding the spread of fake news on Twitter:

(1) **Does fake news spread more than real news on Twitter?** Yes, our analysis reveals that fake news is more prevalent and spreads more widely than real news. The average cascade size for fake news was significantly higher, indicating it reaches a larger audience and has more extensive propagation patterns.
(2) **Do people with a higher following spread fake news?** Our findings suggest that influential users, who often serve as critical connectors between other users, play a significant role in spreading both fake and real news. Measures such as betweenness and closeness centrality indicate that users with high followings are crucial in disseminating information, including fake news.
(3) **How do emotions affect reactions towards tweets?** Emotions significantly impact the spread of fake news. Tweets expressing emotions such as disgust, joy, and fear are more likely to be shared. Notably, tweets expressing fear had significantly shorter reaction times, indicating quicker engagement and wider dissemination. This underscores the need to address emotionally charged content to curb misinformation.

Overall, our study provides valuable insights into the dynamics of fake news propagation on social media networks. It highlights the importance of effective detection and mitigation strategies, particularly focusing on influential users and emotionally charged content. Implementing fact-checking mechanisms and restricting the spread of identified fake news are crucial steps towards maintaining the integrity of information on social media platforms. Additionally, educating users to critically evaluate information sources can further strengthen these efforts.

6.1 Supplementary Materials

The workspace, including the code, figures, results and datasets, used in this study can be accessed through the following link: Project Workspace.

REFERENCES

[1] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.

[2] Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.

[3] Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., ... & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554-559.

[4] Budak, C., Agrawal, D., & El Abbadi, A. (2011). Limiting the spread of misinformation in social networks. In *Proceedings of the 20th international conference on World wide web* (pp. 665-674).

[5] Srivastava, G., Kumar, P., Ojha, R. P., Srivastava, P. K., Mohan, S., & Srivastava, G. (2020). Defensive Modeling of Fake News Through Online Social Networks. *IEEE Transactions on Computational Social Systems*, 7(5), 1159-1167. https://ieeexplore.ieee.org/abstract/document/9167402

[6] Schroeder, D. T. (2022). Explaining News Spreading Phenomena in Social Networks: From Data Acquisition and Processing to Network Analysis and Modelling (Order No. 29780950). Available from *ProQuest Dissertations & Theses Global*; ProQuest Dissertations & Theses Global Closed Collection. (2734703082). https://www.proquest.com/docview/2734703082?pq-origsite=gscholar

[7] Caponi, S., Khalifa, Z., Oligeri, G., & Di Pietro, R. (2022). Fake News Propagation: A Review of Epidemic Models, Datasets, and Insights. *ACM Transactions on the Web*, 16(3), 12. https://dl.acm.org/doi/full/10.1145/3522756

[8] Wang, Y., Park, C.-H., & Seo, Y.-S. (2019). Fake News Analysis Modeling Using Quote Retweet. *Electronics*, 8(12), 1377. https://www.mdpi.com/2079-9292/8/12/1377

[9] Li, S., Li, J., & Kao, H.-Y. (2021). MVAN: Multi-View Attention Networks for Fake News Detection on Social Media. *IEEE Access*, 9, 106907-106917. https://ieeexplore.ieee.org/abstract/document/9497048

[10] Shrestha, A., & Spezzano, F. (2022). Characterizing and Predicting Fake News Spreaders in Social Networks. *International Journal of Data Science and Analytics*, 13, 385-398. https://link.springer.com/article/10.1007/s41060-021-00291-z

[11] Rodaghi, A., & Oliveira, J. (2022). The theater of fake news spreading, who plays which role? A study on real graphs of spreading on Twitter. *Expert Systems with Applications*, 189, 116110. https://www.sciencedirect.com/science/article/pii/S095741742101441X

[12] Aramaki, E., Maskawa, S., & Morita, M. (2011). Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 1568-1576). https://aclanthology.org/D11-1145/

[13] Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013). Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 729-736). https://dl.acm.org/doi/10.1145/2487788.2488033