Unraveling the Dynamics of Fake News and Misinformation on Twitter: A Comprehensive Exploration

Amna Hassan
Department of Computer Science
University of Calgary, Canada
amna.hasan1@ucalgary.ca

Kashfia Sailunaz
Department of Computer Science
Dalhousie University, Canada
kashfia.sailunaz@ucalgary.ca

Mehmet Kaya
Department of Computer Engineering
Firat University, Turkey
kaya@firat.edu.tr

Mehmet K. Ozdemir
Department of Computer Engineering
Istanbul Medipol University, Turkey
mkozdemir@medipol.edu.tr

Jon Rokne
Department of Computer Science
University of Calgary, Canada
rokne@ucalgary.ca

Reda Alhajj
Department of Computer Science, University of Calgary, Calgary, Alberta, Canada
Department of Computer Engineering, Istanbul Medipol University, Istanbul, Turkey
Department of Heath Informatics
University of Southern Denmark, Denmark
alhajj@ucalgary.ca

Abstract: *In this digital age, social media plays a pivotal role in instant information propagation across the world. However, the same characteristics also make the social media the major contributor to spreading misinformation. This paper contributes to the ongoing efforts in combating misinformation on social media (specially Twitter) platforms by adopting a holistic approach that combines network analysis, Machine Learning (ML), and psychological insights. It offers a detailed analysis of the Twitter network by exploring the dynamics of misinformation on Twitter. The resharing user and tweet network, as well as tweet and user features like tweet sentiment, user engagements, and their implications on misinformation propagation have been considered and evaluated. Extreme Gradient Boosting (XGBoost) achieved Mean Squared Error (MSE) of 26.81 and R2 score of 0.97 showing high predictive accuracy with relatively small prediction errors for verifying tweet authenticity. This can support the analysis of finding influential users and tweets in order to track and foster and improved understand of misinformation propagation. As the digital landscape continues to evolve, the strategies developed through this research will be instrumental in enhancing the integrity and reliability of information on social media, thereby fostering a healthier digital public discourse.*

Keywords: *Network analysis, machine learning, psychological drivers, combating misinformation, twitter (X).*

Received October 12, 2025; accepted December 10, 2025
https://doi.org/10.34028/iajit/23/4/4

1. Introduction

Social media has transformed information propagation and platforms like Twitter has become a key medium for real-time information. However, the issue of misinformation poses significant challenges to public trust and discourse [35]. Current research works in this domain often lacks accessible, integrated solutions to effectively combat this problem. Most studies about false information on Twitter mostly focus on one aspect at a time such as how misinformation spreads, who shares it, or what the content is [27, 33]. The studies lack the integration of the various dimensions of social networks such as human psychology, and Machine Learning (ML). Due to this limitation, we never fully understand how these different domains influence each other and how they facilitate the spread of misinformation.

Misinformation can lead to severe outcomes by making people believe in false information and forming both truthful and fake social, national, and global issues within a short period of time [18, 31]. For example, fake news propagation can cause political polarization, election interference, hate crimes, mob violence, economic disruptions, conspiracy theories, health misinformation panic in public, and many more fatal outcomes. The objective of this paper is to address the critical challenge of misinformation on Twitter. It focuses on the development of effective strategies for mitigating the impact of misinformation. Through this research, we aim to enhance the integrity and reliability of information on social media platforms.

The research problem addressed in this paper is exacerbated by a noted gap in existing literature. It lacks a holistic approach incorporating the psychological underpinnings of misinformation spread for realizing all of the dynamics of social interaction at play. This research aims to unravel the complex web of misinformation on Twitter by employing a comprehensive mixed-methods approach. Our work blends advanced algorithmic development with in-depth analysis, including psychological drivers. Developing a comprehensive system for influential user identification can also arise few challenges. Some of the challenges are complex algorithmic influences (unraveling Twitter-specific algorithmic biases and their interplay with user psychology in promoting misinformation spread); understanding psychological drivers (quantifying and modeling the psychological factors that contribute to the spread and acceptance of misinformation on Twitter presents a multifaceted challenge, requiring interdisciplinary approaches); and effective visualization of complex data (creating visualizations that accurately represent the psychological factors and complex network interactions involved in misinformation spread, making the data accessible and comprehensible to a diverse audience).

These challenges must be addressed focusing on all of the network and user aspects. A complete network analysis of Twitter is therefore needed to identify patterns in misinformation propagation. Simultaneously, influential nodes of the network should be analyzed considering their psychological factors that influence user behavior and content virality. Additionally, development and application of ML specifically designed for Twitter data, algorithms are required, to predict the authenticity of tweets. The model should incorporate a variety of features, including those reflecting user psychology and network dynamics for structural analysis. Finally, utilization of sophisticated data visualization techniques is essential to elucidate the complex interactions and psychological aspects of misinformation spread on Twitter.

In this paper, we proposed a novel approach combining three domains network analysis, ML, and psychological drivers. The major contributions of this research are as follows:

- Development of Predictive Models: to employ an advanced ML algorithm to Twitter data, aiming to predict the veracity of tweets. The objective is to identify distinguishing features of tweets that could indicate their likelihood of being misinformation, based on aspects such as the content, metadata, and user engagement patterns.
- Comprehensive Network Analysis: to conduct an in-depth examination of Twitter's network to uncover key influencers and the patterns through which misinformation is disseminated.
- Effective Data Visualization: to use sophisticated data visualization techniques to present the findings in an accessible and comprehensible manner. Where the goal is to explain the complex dynamics of misinformation spread on Twitter, thereby making the insights available to a broader audience beyond the academic community.

The rest of the paper is organized as follows: section 2 discusses some recent relevant research works; section 3 explains the system architecture; section 4 shows the experimental results for network analysis and ML; section 5 discusses the psychological driver and their implications for this study; and finally, section 6 concludes the paper with potential future research directions.

2. Related Works

Current research into detecting and mitigating fake news on Twitter often focuses on either network analysis or on ML techniques [3, 22, 28, 44]. They rarely combine these with an understanding of the psychological drivers behind the misinformation [24, 34]. This literature review examines key studies within these domains, highlighting the current advancements. It also focuses on identifying the gaps in addressing the complex dynamics of misinformation preparing the groundwork for this research, its novelty and potential contributions.

A. Network Analysis Methods

Srebrenik [41] proposed a pioneering effort in employing network analysis combined with natural language processing to pinpoint the origins of fake news on Twitter. By focusing on ego networks, which center around a primary node (the ego) and its directly connected nodes (alters), they ventured beyond traditional news verification methods. They showed the complex web of interactions among verified Twitter users, leveraging graph theory and ML to sift through patterns indicative of misinformation. They exploited the structural properties of Twitter's social network as it serves as a fertile ground for fake news dissemination. Their study contributed significantly to understand misinformation spread in Twitter. Although the potential to trace misinformation back to its source, offered a novel idea, they had few limitations. Their research struggled with the generalization of findings beyond verified networks and the results were mostly relevant to a dataset where trust was pre-assumed. Additionally, the study's challenge in identifying fake news sources for training data underscored the broader issue of subjectivity and the lack of a universal standard for what constitutes fake news. Their approach could have been improved by integrating unverified user data to enhance the model's comprehensiveness. Finally, incorporating temporal dynamics into the analysis could improve the dynamic property of their approach.

Interestingly, most of the recent Twitter-based fake news propagation research has focused on health misinformation [7]. The misinformation in this domain spread on social networks, particularly during crises like the COVID-19 pandemic was addressed by Duzen *et al.* [11]. Bento *et al.* [5] introduced a novel design and software architecture leveraging Social Network Analysis (SNA) metrics to detect misinformation on Twitter. The proposed prototype encompassed modules for data collection, preprocessing, network creation, centrality calculation, community detection [37], and misinformation analysis. An experimental study on a COVID-19-related Twitter dataset demonstrated the effectiveness of the system. It provided valuable insights into the application of network science metrics for analyzing misinformation spread. For example, their application of centrality calculations and community detection algorithms to uncover patterns of misinformation spread was moved. While the study focused on Twitter, the generalizability of the proposed process and software architecture to other social media platforms remained unclear. Misinformation dynamics can also vary significantly across platforms due to differences in user engagement and content dissemination mechanisms. Additionally, the requirement for significant computational resources might limit its accessibility to practitioners with constrained computational capabilities.

Another COVID-19 misinformation prediction model was proposed by Oroh *et al*. [30] combining both SNA and sentiment analysis. They used centrality measure scores to detect the key actors in the network for the fake news spread. Their sentiment analysis showed a significant relation between the influential actors and their sentiment. These finding led to a proposed approach for finding the relationships between actors, their centrality measures, and their sentiments and their relevance to the potential of misinformation propagation of those actors. Gbaje *et al*. [15] used the Mozdeh software [42] for big data analysis in Nigeria to predict and track the fake news propagation with SNA methods. Their experiments on several major fake news propagation, influential actor identification, and their follower tracking provided some new insights. Their analysis showed that instead of the influencers being the main actors for misinformation propagation, it was the set of followers of the influencers that contributed the most in fake news dissemination.

B. Machine Learning Approaches

The research problem was also analyzed and tested with various ML models in recent years [2, 39]. Soman [40] recently proposed a novel ML-based system for detecting fake news on Twitter using Support Vector Machine (SVM) [21] as the classifier. The feature extraction technique utilized included Term Frequency-Inverse Document Frequency (TF-IDF) applied to bag-of-words and n-grams. The methodology emphasized text pre-processing, encoding, feature extraction, and classification to identify misinformation efficiently. Their integration of TF-IDF with n-grams and SVM achieved high classification accuracy by targeting the rapid dissemination of false information. This proved the model's capability to distinguish between real and fake news effectively. However, the reliance on TF-IDF and n-grams might possibly overlook the potential of other linguistic and semantic features that could improve detection accuracy.

Another SVM-based rumor spread identification model was proposed by Sitio *et al*. [38]. Their analysis of user features, tweet features, and conversation features on a Twitter PHEME dataset showed high classification accuracy with SVM. Their sentiment score distributions for each major event for the tweets and their relevance to the rumors showed high relevance between negative sentiment and rumors. The proposed model was limited to finding a high performance SVM for rumor classification. However, their sentiment analysis provided a potential research scope to explore the relationships between sentiment and rumors.

Sharma and Sharma [36] also experimented on the same PHEME dataset by proposing a weak supervised model for identifying possible rumor spreaders in Twitter. They explored user features, text features, and ego-network features and transformed the tweets dataset into rumor spreader dataset. After applying sentiment analysis on each category of tweets, they defined a score to compute the rumor spreader's intensity. It was calculated by incorporating the frequencies of users' tweet and their total frequency of tweeting. Another score on user importance based on the reply frequencies was computed too. Finally, Graph Convolutional Network (GCN), SVM, Long Short-Term Memory (LSTM) and Random Forest (RF) models were tested for finding possible rumor spreaders. With GCN being the highest performance model, their future plan included exploring multiclass problems with additional datasets. Unlike the other researchers, Tommasel *et al*. [43] proposed a Deep Learning (DL)-based model for tracking the fake news spreaders in social networks. An attention-based architecture was applied to user features for classifying the users most likely to spread fake news. The model achieved high precision showing promising future scopes of DL models for fake news propagators identifications.

C. Psychological Drivers

The psychological perception of misinformation propagation within a network mostly analyze and emphasize human characteristics and their versatility influencing various social behavior [17]. The process includes several factors such as individual characteristics, cognition, ignorance, affection, social process, and their combinations [29]. Ecker *et al*. [12] presented a comprehensive review exploring the multifaceted psychological underpinnings of misinformation belief and its persistence even after corrections. It delved into cognitive, social, and emotional factors that contribute to the formation and endurance of false beliefs. Highlighting the limitations of the information deficit model, it underscored the complexity of misinformation's appeal and the challenges in debunking it effectively. The paper evaluated interventions like prebunking and debunking, examining their effectiveness and the role of social media in spreading and correcting misinformation. The authors provide a holistic approach by offering a thorough examination of the psychological mechanisms behind misinformation belief. They emphasized not just the cognitive but also the social and emotional factors. They provided an insightful analysis of both preemptive and reactive strategies for countering misinformation. Their discussions included the role of social media in misinformation propagation. They integrated findings from psychology, communication studies, and information science, making it valuable across disciplines. While extensive, the review primarily focused on textual misinformation. Their study paid less attention to visual or multimedia forms like deep fakes or manipulated images. The broad scope of their study may gloss over the nuances of specific misinformation contexts or the differential effectiveness of interventions across various domains.

Psychological and motivational factors of fake news spreaders in social media were discussed by Karami *et al*. [23]. A statistical comparative analysis applied with a series of studies to analyze social behaviors highlighted few insightful factors of human characteristics. Motivational factors like uncertainty, emotions, lack of control, relationship enhancement, and rank were analyzed for their study. They studied political news and gossip news for identifying misinformation actor. It showed that social media actors vary in their characteristics compared to the theoretical rumor spreader. Some of the theoretical features can provide suggestions on exploring new features for rumor propagation actors in social media. Metzler and Garcia [26] showed the complexity of the social drivers and algorithmic mechanisms influencing political misinformation. Their in-depth analysis showed the research gaps in the domain and proposed potential future analysis to advance the scopes of such research works.

Table 1. Summary of few related works from the categories network analysis, ML, and psychological drivers for misinformation propagation detection and analysis.

Ref.	Category	Approaches	Contributions	Limitations
[11]	Network analysis	- SNA	- Prototype development. - COVID-19 vaccination-based Twitter misinformation detection.	- Generalizability among other domains/networks. - Requires additional computational resources.
[30]	Network analysis	- SNA - Sentiment analysis.	- Detected key actors for misinformation spread in Twitter. - Analyzed positive and negative sentiments of users.	- Insufficient exploration on relevance of user sentiment and misinformation propagation. - User influence mostly considered high number of tweets and avoided other tweet and user features.
[15]	Network analysis	- SNA	- Analyzed fake news propagation. - Identified influencers/news creators. - Detected main actors/news propagators (followers of influencers).	- Region specific analysis (Nigeria). - Applied only SNA metrics.
[40]	Machine learning	- SVM	- Classified fake news from Twitter. - Almost perfect performance scores.	- Focused only on TF-IDF and n-grams. - Verification required with other larger datasets, additional features.
[38]	Machine learning	- SVM - Sentiment analysis.	- User, tweet, and conversation features. - High relevance between negative sentiment and rumor.	- Focused mostly on SVM performance improvement
[36]	Machine learning	- SVM - RF - GCN - LSTM	- Used user, text, ego-network and sentiment features - Computed user importance and rumor spreading intensity scores	- Limited to binary classifications. - Focused on only classifying misinformation.
[43]	Machine learning	- Attention-based DL	- Tracked misinformation spreading users in a network.	- Limited to only user features.
[12]	Psychological drivers	- Theoretical analysis	- Discussed contributions of cognitive, social, and emotional factors. - Explored role of social media in misinformation propagation. - Described strategies for countering misinformation.	- Limited to textual misinformation. - Insufficient discussion on multiple domain interventions.
[23]	Psychological drivers	- Statistical analysis	- Compared studies on misinformation propagation in social media. - Analyzed psychological and motivational factors of fake news spreaders.	- Limited to theoretical comparisons. - Focused on political news and gossip news only.
[26]	Psychological drivers	- Complexity analysis	- Discussed social drivers and process of misinformation propagation.	- Limited to political misinformation only.

As mentioned in the previous section, a visible gap in the existing work is the lack of researches combining all these three factors to check how they affect misinformation propagation in any social network. Table 1 shows the summary of the research outcome from three categories and their characteristics. We proposed and discussed a potential approach to combine network analysis, ML, and psychological drivers in the next few sections.

3. System Design

In this section, the proposed system design for analyzing the dynamics of misinformation on Twitter is explained. A detailed description of the dataset utilized for this study is included along with the methodology for network construction and analysis.

A. Dataset

A Twitter dataset named "Tweets and User Engagement" [19] containing basic tweet features and user engagement metrics was utilized to explore the network analysis in this research. The dataset included categorical, numerical, Boolean, date, and text features providing a comprehensive overview of Twitter activities. Unique tweet IDs are associated with weekday, hour, day, language, tweet text, tweet sentiment represented basic tweet characteristic and tweet likes, retweets, reshares provided the influences of the tweet. Additionally, the user ID, and Klout score showed the user influence. These features are crucial for understanding the mechanisms behind spread of misinformation and potential indicators of fake news within Twitter network. Table 2 shows a few key features and their descriptions.

Table 2. Key components of the dataset.

Feature	Data type	Description
TweetID	Text	A unique identifier assigned to each tweet.
Klout score	Numeric	Represents the influence level of a Twitter user, where a higher score signifies a stronger capacity to engage with and influence the user's network. The Klout score is instrumental in pinpointing influencers who may significantly contribute to the spread or containment of misinformation.
Tweet	Text	The actual textual content contained within tweets.
IsReshare	Boolean	A binary indicator signifying whether the tweet is a reshare (retweet).
UserID	Text	A unique identifier assigned to Twitter users.

B. Proposed System Workflow

This research applied network analysis to scrutinize interactions and connections among users, with a special focus on users with high Klout scores. The dataset was processed and a supervised ML algorithm was trained using the features. Features like Klout scores and tweet texts were used to ascertain the authenticity of tweets. Data visualization techniques was employed to underscore the dissemination patterns of misinformation and to identify the networks of influencers on Twitter. Figure 1 shows the basic workflow of the proposed system. By leveraging this meticulously compiled dataset and adopting a mixed-methods approach, this study aspires to shed light on the intricate dynamics of misinformation propagation on Twitter. The ultimate goal is to contribute to the formulation of more efficient strategies for the detection and mitigation of misinformation.

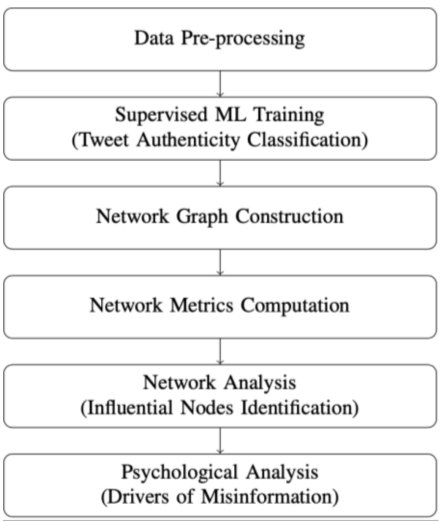


Figure 1. Workflow diagram for the proposed system.

C. Network Analysis and Visualization

1) Network Graph Construction and Visualization: the dataset was used to generate graphs using NetworkX [1], a library designed for the creation, manipulation, and study of complex networks. The edges/connection were created by iterating through the dataset using the UserIDs (i.e., Resharing Users) for tweets to generate the resharing network of users. The generated networks were then visualized as the network graph with matplotlib [4]. This visualization included nodes represented by their UserID or TweetID and edges indicating resharing relationships. The visualization provided the overall structure of the network, identified clusters of activity, and visually confirmed the centrality of the top nodes.
2) Network Metrics: the network centrality measures were considered on the generated graph to detect the misinformation propagation and the 'Degree Centrality' measure was chosen for the task. 'Degree Centrality' computes the number of edges/connection of a node in the network [10]. The 'Degree Centrality' aids in misinformation detection by:

   a) Identifying Influential Spreaders: users with high degree centrality are key players in spreading information. By analyzing these central nodes, researchers can identify accounts that frequently disseminate content, including potentially false or misleading information.
   b) Flagging Viral Content: tweets that achieve high centrality quickly (i.e., are reshared widely in a short period) can be flagged for further review to ensure they do not spread misinformation. Rapid spikes in resharing can sometimes indicate coordinated campaigns to spread false information.
   c) Content Verification Prioritization: given limited resources, content that becomes highly central in the network (widely reshared tweets) can be prioritized for fact-checking and verification. This approach ensures that efforts are focused where

potential misinformation could have the greatest impact.

D. Machine Learning Algorithms

The primary goal of this research was to develop a predictive model capable of forecasting the reach of tweets within the Twitter network. To achieve this, the data was processed for feature engineering and model selection, culminating in the successful prediction of tweet reach.

1. Feature Engineering and Data Preprocessing: the TF-IDF [8] vectorization technique was applied to transform the textual content of the tweets into a numerical format facilitating an ML analysis. This technique evaluates the importance of a word in a document within a collection of documents by considering both its frequency in the document and its inverse frequency across the entire document corpus. TF-IDF is widely used in information retrieval and text mining to reflect the significance of terms in documents. This advanced text processing method enabled the extraction of features such as sentiment scores, topic relevance, and keyword presence from the tweet text. It enhanced the model's ability to understand and predict based on the content's nature. Additional features such as 'EngagementScore' and 'KloutEngagementInteraction' were derived by combining the RetweetCount, Likes, and Klout scores. These engineered features aimed to capture the influence and engagement dynamics behind tweet dissemination, serving as potent indicators of a tweet's potential reach.
2. ML Model Selection: Extreme Gradient Boosting (XGBoost) regressor [6] was selected for this research due to its superior performance in handling sparse data, efficiency in computational speed, and predictive accuracy. It was chosen after experimenting with several ML models with the tweet and user feature sets. XGBoost is an advanced implementation of gradient boosting algorithms known for its speed and performance. It has been widely adopted in ML competitions and practical applications due to its efficiency in handling various types of data and its capacity for parallel processing. XGBoost stood out for its ability to manage high-dimensional data and its robustness against overfitting, making it particularly suited for our predictive modeling task.

4. Experimental Results and Discussions

The experiments and their results when the proposed algorithm was applied to Twitter dataset are discussed in this section. It discusses the network generation, visualization, and centrality measurements. This section also mentions their implications for misinformation propagation along with the ML outcomes.

A. Network Analysis

For our graph generation, we initialized a directed graph 'G' and then added edges based on resharing actions. The edges were added by iterating through a subset of the dataset. Specifically, for each row where IsReshare was true, an edge was created from the UserID (the resharing user) to the TweetID (the original tweet), effectively mapping the resharing relationship. The degree centrality was then calculated for all nodes in the graph serving as an indicator of the importance or influence of each node (i.e., UserID) within the network. By sorting the nodes based on their degree centrality, top 5 nodes were identified. They represented the users or tweets with the highest resharing activity or popularity in the network. Table 3 shows the top 5 node IDs and their corresponding degree centrality measures [9]. Figure 2 shows the network visualization for the resharing network. Labels were added to the top 5 nodes only to enhance the clarity of the visualization.

Table 3. Top 5 nodes (i.e., nodes with highest popularity) from the generated network.

Popularity rank	Node ID	Degree centrality
1	'tw-3179389829'	0.017730496453900707
2	'tw-3316009302'	0.007978723404255319
3	'tw-105163529'	0.007978723404255319
4	'tw-82221860'	0.0070921985815602835
5	'tw-15812482'	0.0062056737588652485

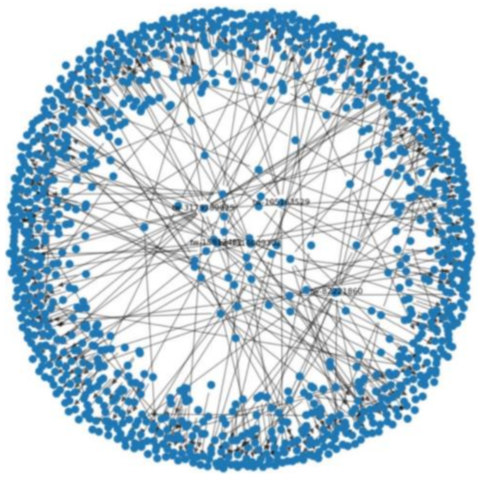


Figure 2. Network visualization for the resharing network.

B. Machine Learning Model

The ML model was trained and validated with the dataset and the performances were evaluated by the Mean Squared Error (MSE) [14] and R2 score [13]. The dataset was meticulously split into training and validation sets to ascertain the model's generalizability. The training-validation division was applied to ensure that the model could perform reliably on unseen data. This is a critical step in validating the models' applicability to real-world scenarios.

The model exhibited outstanding performance, achieving a MSE of 26.81 and a R2 score of 0.97. These metrics indicated a high predictive accuracy, with the R2 score reflecting the model's ability to explain approximately 97% of the variance in tweet reach. Such results underscored the effectiveness of the approach in identifying significant predictors of tweet reach, including specific keywords, user influence scores, and optimal posting times. Figure 3 shows the MSE and R2 scores of the model.

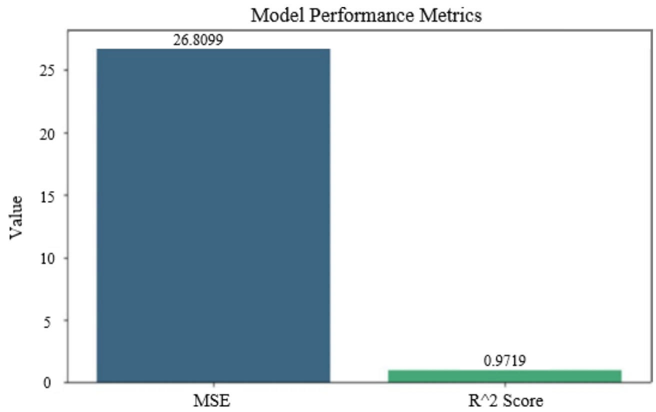


Figure 3. ML model performance scores.

The integration of TF-IDF vectorization and XGBoost helped the model to predict tweet reach with high accuracy. Additionally, it contributed to the broader understanding of misinformation dynamics on social media platforms. The findings from this study are poised to aid stakeholders in strategically enhancing the visibility and impact of their content on Twitter. This marks a significant stride towards mitigating the spread of misinformation and fostering a healthier digital public discourse.

C. Data Feature Visualization

The dataset features were further explored for providing insights on user engagements, tweet sentiments, and retweets.

1. User Engagement: the Klout score provides a measure of user influence on Twitter that can help with identifying highly influential users. These users usually generate higher engagement rates with their tweets by further exploring interactions such as likes, retweets, and replies between them and other users mentioned in their tweets. As depicted in Figure 4, a clear correlation between higher Klout scores and increased reach was observed, indicating that users with greater engagement on Twitter ten to have a wider audience. The Klout score serves as a measure of the user's influence, with a higher score suggesting more significant influence and a broader reach. Understanding this relationship is crucial, especially when addressing misinformation, as it under-scores the tendency for users to trust information from credible sources. On social media platforms, credibility often aligns with higher user engagement, highlighting the importance of assessing influence metrics in evaluating the spread of information.

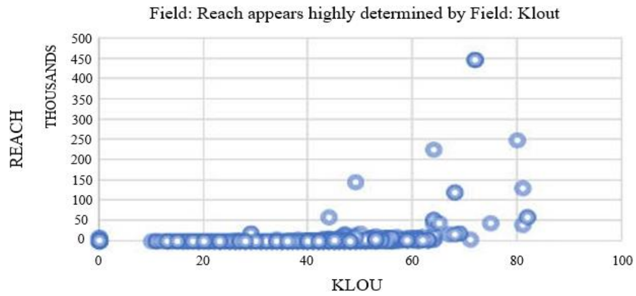


Figure 4. User engagement graph showing user reach vs Klout score.

2. Overall Sentiment: analyzing the distribution of tweets across different days of the week can offer valuable insights into the time or day users tend to engage or post more. Additionally, sentiment analysis enables us to gauge the emotional tone of digital text, distinguishing between positive, negative, or neutral sentiments. Using the tweet sentiment data from our dataset, Figure 5 illustrates the relationship between weekdays and tweet sentiments, revealing that Sundays and Fridays exhibited a higher sentiment, indicating predominantly positive emotions. This observation on the relevance between sentiment and weekday aligns with the typical weekend time frame from Friday to Sunday, during which users tend to express more positive sentiments. Understanding this pattern provides insight into the days when users are more inclined to share positive sentiments, offering valuable context for interpreting social media activity.

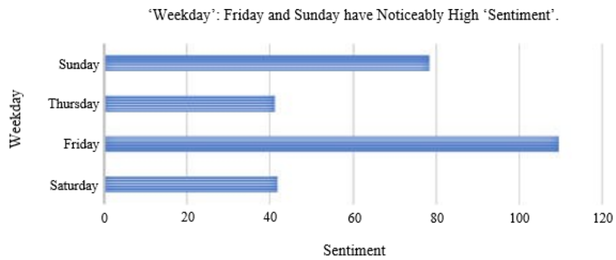


Figure 5. Relationship between tweet sentiment and the day of the tweet posted.

3. Retweet Activity: exploring retweet activity during the identified time bracket could yield intriguing patterns of user behaviors. Since positive sentiment tends to be heightened during the weekend, it would be worthwhile to examine whether this correlates with an increase in retweets. This analysis sheds light on whether users are more likely to share content during periods of elevated positivity, potentially indicating a higher level of engagement or endorsement. As represented in Figure 6, there was a noticeable increase in retweet activity during the identified time bracket, reinforcing the notion that this period is significant for the spread of information, including misinformation. The surge in retweets during this timeframe suggests heightened engagement and amplification of content, which can play a pivotal role in disseminating both accurate and misleading information.

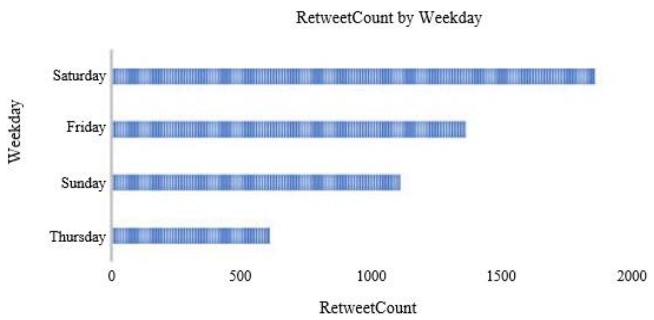


Figure 6. Relationship between retweets and weekday.

5. Psychological Drivers of Misinformation

This section delves into the psychological aspects that influence the spread of misinformation. It also discusses their integration with the findings from network analysis and the ML model developed in this study.

1. Understanding Psychological Drivers: in the intricate landscape of misinformation, the psychological underpinnings play a pivotal role in shaping individual and collective responses to fake news and misleading content on social media platforms like Twitter. As Pennycook and Rand [32] mentioned, the tendency of individuals to prefer information that confirms their pre-existing beliefs, known as confirmation bias, and the influence of social identity on information consumption create fertile ground for misinformation to flourish. Despite political affiliations often being cited as a primary motivator for susceptibility to fake news, the same body of research revealed that individuals are actually more adept at distinguishing truth from falsehood when the information aligns with their political beliefs. Instead, the challenges in discerning truth stem from a lack of critical reasoning, relevant knowledge, and the reliance on heuristics such as familiarity.

Moreover, a significant disconnection exists between what individuals believe and what they choose to share on social media, with this gap primarily attributed to inattention rather than deliberate dissemination of false information. Psychological drivers of misinformation are multifaceted, involving cognitive biases, emotional reactions, and social influences. Key factors include:

a) Confirmation Bias: individuals tend to favor information that confirms their preexisting beliefs and views, making them more susceptible to misinformation aligning with these beliefs.
b) Social Identity Theory [25]: people are more likely to trust and share information within their in-group, which can lead to echo chambers where misinformation circulates unchallenged.
c) Emotional Engagement: misinformation often evokes strong emotional responses, such as fear or anger, which can increase its spread as individuals react and share impulsively.

These insights underscore the necessity of developing interventions that encourage users to prioritize accuracy and engage in reflective thinking. By integrating psychological understanding with network analysis and ML algorithms, as demonstrated in this study, we can devise more effective strategies to mitigate the spread of misinformation. This holistic approach not only addresses the content and dissemination patterns but also considers the psychological factors driving the appeal and spread of misinformation, thereby offering a comprehensive solution to enhance the digital information ecosystem's integrity.

2. Integration with Network Analysis: network analysis offers a macroscopic view of how misinformation travels through social networks on Twitter. By identifying key influencers (nodes) seen in Figure 2 and the structure of information dissemination, we can apply psychological insights to understand the dynamics at play:
   a) Influencer Dynamics: influencers with high Klout scores can significantly impact what is perceived as credible information. Understanding the psychological influence, they wield can help in developing strategies to engage these influencers in spreading accurate information.
   b) Echo Chambers [16]: network analysis reveals the clustering of users around certain beliefs or misinformation. Recognizing these patterns enables us to target interventions more effectively, breaking through echo chambers with diverse, accurate information.

3. Connecting to ML: the ML model developed in this study that predicted tweet reach and identified potential viral content, can be enhanced by incorporating psychological insights for
   a) Feature Engineering [20]: adding features that capture emotional tone or alignment with common biases can improve the model's ability to predict the spread of misinformation.
   b) Sentiment Analysis [45]: integrating sentiment analysis results with ML models offers a deeper understanding of how emotional engagement drives tweet virality, aligning with psychological insights on emotional responses to misinformation.

4. Combining Network Analysis, ML, and Psychological Drivers: the intersection of psychological drivers with network analysis and ML provides a holistic approach to tackling misinformation propagation:
   a) Tailored Interventions: by understanding the psychological factors at play, interventions can be tailored to address specific biases or emotional triggers, making them more effective.
   b) Influencer Engagement: identifying influencers through network analysis and understanding their

psychological impact allows for targeted engagement strategies, using influential voices to promote accurate information.

c) Adaptive Models: incorporating psychological insights into ML models makes these tools more adaptive to the nuanced ways misinformation spreads, enhancing their predictive accuracy and utility in real-time misinformation detection.

The aim of this paper was show that the integration of psychological insights with network analysis and ML offers a powerful framework for understanding and combating misinformation on Twitter. By recognizing the complex interplay of psychological factors with the structural and content-related aspects of information spread, this research moves closer to developing effective strategies for mitigating the impact of misinformation, ultimately contributing to a more informed and resilient digital public sphere.

6. Conclusions

As social media keeps evolving to be one of the most influential channels for propagating both true and fake information, a multi-dimensional approach to address the spread of misinformation becomes more evident day-by-day. A novel multi-domain approach is proposed in this paper to address the misinformation propagation in Twitter. The integration of predictive modeling, network analysis, and psychological insights represents a significant contribution to the field. This contribution positions this research as a considerable step forward in the battle against misinformation on social media platforms.

This research successfully demonstrated the potency of combining network analysis with ML models to predict tweet reach, and identify potential viral content and influential users. The employment of the TF-IDF vectorization technique, along with the construction of new features like the EngagementScore and KloutEngagementInteraction, showed significant improvements in the predictive capabilities of our ML model. The use of the XGBoost algorithm further solidified the model's accuracy, as evidenced by the high R2 score and low MSE in predicting tweet reach. Furthermore, the study has underscored the critical role of psychological drivers in the spread of misinformation. By integrating psychological insights into the analysis providing a more nuanced understanding of why certain pieces of information go viral and how misinformation can be more effectively countered. The exploration of sentiment analysis and user engagement patterns has yielded valuable insights into the dynamics of information dissemination on Twitter. This highlighted the importance of considering psychological factors in combating misinformation. The visualization techniques employed have made the findings accessible to a broader audience. They have also offered compelling visual evidence of the patterns and trends in misinformation spread on Twitter. These visualizations serve as powerful tools for illustrating the complex relationships and interactions that drive the dissemination of false information on the platform.

By employing advanced ML algorithms tailored to Twitter's unique dataset, this research predicts the veracity of tweets. It also delves into the psychological mechanisms at play, addressing a critical gap in the literature. Furthermore, through comprehensive network analysis and sophisticated data visualization, it transcends academic boundaries, making the intricate dynamics of misinformation spread accessible to a broader audience. This endeavor seeks to identify and counteract misinformation on Twitter, while enhancing the digital public sphere's integrity by understanding the psychological components driving misinformation's appeal and dissemination. Additionally, this research lays the ground work for future research directions for applying similar methods on other social networks and online groups where users share their thoughts and opinions. Another potential extension would be including more features of both users and their posts and exploring their relevance, dependencies, correlations to find patterns that affects misinformation generation and propagation. Exploring latent psychological drivers other than sentiments by analyzing historical data of users such as their connections, interests, online activities, changes in their interests and connections over time can provide additional features to identify and minimize misinformation in any online user network.

References

[1] Aggarwal T., NetworkX: A Comprehensive Guide to Mastering Network Analysis with Python, https://medium.com/@tushar_aggarwal/networkx-a-comprehensive-guide-to-mastering-network-analysis-with-python-fd7e5195f6a0, Last Visited, 2025.

[2] Aimeur E., Amri S., and Brassard G., "Fake News, Disinformation and Misinformation in Social Media: A Review," *Social Network Analysis and Mining*, vol. 13, pp. 1-36, 2023. https://doi.org/10.1007/s13278-023-01028-5

[3] Arranz-Escudero O., Quijano-Sanchez L., and Liberatore F., "Enhancing Misinformation Countermeasures: A Multimodal Approach to Twitter Bot Detection," *Social Network Analysis and Mining*, vol. 15, pp. 1-22, 2025. https://link.springer.com/article/10.1007/s13278-025-01435-w

[4] Barrett P., Hunter J., Miller J., Hsu J., and Greenfield P., "Matplotlib a Portable Python Plotting Package," *in Proceedings of the 14th Astronomical Data Analysis Software and Systems Conference*, Pasadena, pp. 91-95, 2004. file:///C:/Users/user/Downloads/matplotlib_--_A_Portable_Python_Plotting_Package%20(1).pdf

[5] Bento A., Cruz C., Fernandes G., and Ferreira L.,

"Social Network Analysis: Applications and New Metrics for Supply Chain Management a Literature Review," *Logistics*, vol. 8, no. 1, pp. 1-20, 2024. https://doi.org/10.3390/logistics8010015

[6] Brownlee J., XGBoost for Regression, https://machinelearningmastery.com/xgboost-for-regression/, Last Visited, 2025.

[7] Cano-Marin E., Mora-Cantallops M., and Sanchez-Alonso S., "The Power of Big Data Analytics over Fake News: A Scientometric Review of Twitter as a Predictive System in Healthcare," *Technological Forecasting and Social Change*, vol. 190, pp. 1-14, 2023. https://doi.org/10.1016/j.techfore.2023.122386

[8] Chaudhary M., F-IDF Vectorizer Scikit-Learn, https://medium.com/@cmukesh8688/tf-idf-vectorizer-scikit-learn-dbc0244a911a, Last Visited, 2025.

[9] Derr A., Network Centrality Infographic-Visible Network Labs, https://visiblenetworklabs.com/2021/04/16/understanding-network-centrality/, Last Visited, 2025.

[10] Dey P., Bhattacharya S., and Roy S., "A Survey on the Role of Centrality as Seed Nodes for Information Propagation in Large Scale Network," *ACM/IMS Transactions on Data Science*, vol. 2, no. 3, pp. 1-25, 2021. https://doi.org/10.1145/3465374

[11] Duzen Z., Riveni M., and Aktas M., "Analyzing the Spread of Misinformation on Social Networks: A Process and Software Architecture for Detection and Analysis," *Computers*, vol. 12, no. 11, pp. 1-17, 2023. https://doi.org/10.3390/computers12110232

[12] Ecker U., Lewandowsky S., Cook J., Schmid P., and et al., "The Psychological Drivers of Misinformation Belief and its Resistance to Correction," *Nature Reviews Psychology*, vol. 1, pp. 13-29, 2022. https://doi.org/10.1038/s44159-021-00006-y

[13] Frost J., How to Interpret R-Squared in Regression Analysis, https://statisticsbyjim.com/regression/interpret-r-squared-regression/, Last Visited, 2025.

[14] Frost J., Mean Squared Error (MSE), https://statisticsbyjim.com/regression/mean-squared-error-mse/, Last Visited, 2025.

[15] Gbaje E., Agwu C., Odigie I., and Yani S., "Curtailing Fake News Creation and Dissemination in Nigeria: Twitter Social Network and Sentiment Analysis Approaches," *Journal of Information Science*, vol. 51, no. 4, pp. 987-999, 2025. https://doi.org/10.1177/01655515231160029

[16] GCFLearnFree.org, Digital Media Literacy, What is an Echo Chamber?, https://www.youtube.com/playlist?list=PLpQQipWcxwt9NUsBX4KpO4PwHMilgzEh1, Last Visited, 2025.

[17] Gwiazdzinski P., Gundersen A., Piksa M., Krysinska I., and et al., "Psychological Interventions Countering Misinformation in Social Media: A Scoping Review," *Frontiers in Psychiatry*, vol. 13, pp. 1-12, 2023. DOI: 3389/fpsyt.2022.974782

[18] Jahng M., "Is Fake News the New Social Media Crisis? Examining the Public Evaluation of Crisis Management for Corporate Organizations Targeted in Fake News," *Journal of Strategic Communication*, vol. 15, no. 1, pp. 18-36, 2021. https://doi.org/10.1080/1553118X.2020.1848842

[19] Jensen K., Tweets and User Engagement, https://www.kaggle.com/datasets/thedevastator/tweets-and-user-engagement/data, Last Visited, 2025.

[20] Johary, Liora Formerly DataScientist, Feature Engineering: Importance for Machine Learning, https://liora.io/en/feature-engineering-importance-for-machine-learning, Last Visited, 2025.

[21] Kanade V., What is a Support Vector Machine? Working, Types, and Examples, https://www.spiceworks.com/tech/big-data/articles/what-is-support-vector-machine/, Last Visited, 2025.

[22] Kansara P. and Adhvaryu K., "A Survey of Fake Data or Misinformation Detection Techniques Using Big Data and Sentiment Analysis," *SN Computer Science*, vol. 5, pp. 955, 2024. https://doi.org/10.1007/s42979-024-03297-z

[23] Karami M., Nazer T., and Liu H., "Profiling Fake News Spreaders on Social Media Through Psychological and Motivational Factors," *in Proceedings of the 32nd ACM Conference on Hypertext and Social Media*, London, pp. 225-230, 2021. https://doi.org/10.1145/3465336.3475097

[24] Kwao L., Yang Y., Zou J., and Ma J., "A Survey of Approaches to Early Rumor Detection on Microblogging Platforms: Computational and Socio-Psychological Insights," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 15, no. 1, pp. e70001, 2025. https://doi.org/10.1002/widm.70001

[25] Mcleod S., Social Identity Theory in Psychology, https://www.simplypsychology.org/social-identity-theory.html, Last Visited, 2025.

[26] Metzler H. and Garcia D., "Social Drivers and Algorithmic Mechanisms on Digital Media," *Perspectives on Psychological Science*, vol. 19, no. 5, pp. 735-748, 2024. https://doi.org/10.1177/17456916231185057

[27] Muhammed S. and Mathew S., "The Disaster of Misinformation: A Review of Research in Social Media," *International Journal of Data Science and Analytics*, vol. 13, no. 4, pp. 271-285, 2022.

https://link.springer.com/article/10.1007/s41060-022-00311-6

[28] Munoz P., Diez F., and Bellogin A., "Modeling Disinformation Networks on Twitter: Structure, Behavior, and Impact," *Applied Network Science*, vol. 9, no. 4, pp. 1-35, 2024. https://link.springer.com/article/10.1007/s41109-024-00610-w

[29] Munusamy S., Syasyila K., Shaari A., Pitchan M., and et al., "Psychological Factors Contributing to the Creation and Dissemination of Fake News Among Social Media Users: A Systematic Review," *BMC Psychology*, vol. 12, pp. 1-14, 2024. https://doi.org/10.1186/s40359-024-02129-2

[30] Oroh A., Bandung Y., and Zagi L., "Detection of the Key Actor of Issues Spreading Based on Social Network Analysis in Twitter Social Media," *in Proceedings of the IEEE Asia Pacific Conference on Wireless and Mobile*, Bandung, pp. 206-212, 2021. DOI:10.1109/APWiMob51111.2021.9435268

[31] Park C., "Why People Rely on Fact-Checkers? Testing Theses of "Perceived Severity of Fake News" and "Disappointment in News Media"," *Journalism Studies*, vol. 25, no. 1, pp. 1-18, 2024. https://doi.org/10.1080/1461670X.2023.2289878

[32] Pennycook G. and Rand D., "The Psychology of Fake News," *Trends in Cognitive Sciences*, vol. 25, no. 5, pp. 388-402, 2021. https://doi.org/10.1016/j.tics.2021.02.007

[33] Raponi S., Khalifa Z., Oligeri G., and Di Pietro R., "Fake News Propagation: A Review of Epidemic Models, Datasets, and Insights," *ACM Transactions on the Web*, vol. 16, no. 3, pp. 1-34, 2022. https://dl.acm.org/doi/10.1145/3522756

[34] Rathje S. and Van Bavel J., "The Psychology of Virality," *Trends in Cognitive Sciences*, vol. 29, no. 10, pp. 914-927, 2025. https://doi.org/10.1016/j.tics.2025.06.014

[35] Saxena A., Saxena P., and Reddy H., Fake News Propagation: A Review of Epidemic Models, Springer, 2022. https://link.springer.com/chapter/10.1007/978-981-16-3398-0_16

[36] Sharma S. and Sharma R., "Identifying Possible Rumor Spreaders on Twitter: A Weak Supervised Learning Approach," *in Proceedings of the International Joint Conference on Neural Networks*, Shenzhen, pp. 1-8, 2021. https://doi.org/10.1109/IJCNN52387.2021.9534185

[37] Singh D. and Choudhury P., "Community Detection in Large-Scale Real-World Networks," *Advances in Computers*, vol. 128, pp. 329-352, 2023. https://doi.org/10.1016/bs.adcom.2021.10.007

[38] Sitio C., Sibaroni Y., and Prasetiyowati S., "Identifying Possible Rumor Spreaders on Twitter Using the SVM and Feature Level Extraction," *Jurnal Teknik Informatika*, vol. 4, no. 3, pp. 611-618, 2023. https://doi.org/10.52436/1.jutif.2023.4.3.868

[39] Smith S., Kao E., Mackin E., Shah D., and et al., "Automatic Detection of Influential Actors in Disinformation Networks," *Proceedings of the National Academy of Sciences*, vol. 118, no. 4, pp. 1-10, 2021. https://doi.org/10.1073/pnas.2011216118

[40] Soman A., "Fake News Detection Using Machine Learning Method," *International Journal of Advance Research and Innovative Ideas in Education*, vol. 8, no. 3, pp. 4044-4049, 2022. https://ijariie.com/manuscript/17300

[41] Srebrenik B., Fake News Detection Through Ego Network Analysis, https://github.com/briansrebrenik/Final_Project, Last Visited, 2025.

[42] Thelwall M., Mozdeh, GitHub, https://github.com/MikeThelwall/Mozdeh, Last Visited, 2025.

[43] Tommasel A., Rodriguez J., and Menczer F., "Following the Trail of Fake News Spreaders in Social Media: A Deep Learning Model," *in Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, Barcelona, pp. 29-34, 2022. https://dl.acm.org/doi/10.1145/3511047.3536410

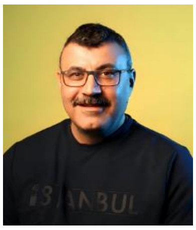
[44] Ul Hussna A., Alam M., Islam R., Alkhamees B., and et al., "Dissecting the Infodemic: An In-Depth Analysis of COVID-19 Misinformation Detection on X (Formerly Twitter) Utilizing Machine Learning and Deep Learning Techniques," *Heliyon*, vol. 10, no. 18, pp. 1-22, 2024. https://doi.org/10.1016/j.heliyon.2024.e37760

[45] Wankhade M., Rao A., and Kulkarni C., "A Survey on Sentiment Analysis Methods, Applications, and Challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731-5780, 2022. https://doi.org/10.1007/s10462-022-10144-1

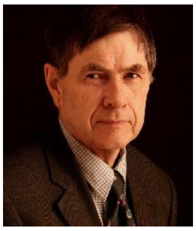
Amna Hassan is a Strategic Resource Manager @ HUB International and Founder @ Clientera. She completed her BSc degree in Computer Science from the University of Calgary. She served as Vice President of Strategy for her AI, a university club aimed at promoting women's advancement in AI and ML. She focuses in transforming complex theoretical concepts into practical outcomes. She holds the "IBM Professional Data Science Certification". Her research interests include Data Analysis, Machine Learning, Deep Learning, Social Media Analysis, Misinformation, Fake News Detection.

Kashfia Sailunaz is a Computer Science and Engineering (CSE) graduate from Bangladesh. She received her MSc and PhD degrees in Computer Science from the University of Calgary. Kashfia worked as a lecturer in CSE at Bangladesh University of Business and Technology (BUBT) for 4 years and later at Military Institute of Science and Technology (MIST) for 1.5 years. She also has teaching experience at UofC where she worked as a sessional instructor in Computer Science and in Electrical and Software Engineering. Her research covers various aspects of Data Science, including Data Collection, Cleaning, Storing, Processing, and Analysis using Machine Learning, Deep Learning, Data Mining and Network Analysis Techniques. She has contributed to the foundations and applications by developing and tuning-up some existing techniques to successfully serve a variety of applications including bioinformatics and healthcare with emphasis on MRI and CT image analysis for tumor/cancer detection.

Mehmet Kaya received his BSc, MSc and PhD degrees in Electrical-Electronics Engineering and Computer Engineering from Firat University, Turkey, in 1996, 1998 and 2003, respectively. Currently, he is Professor in the Department of Computer Engineering, Firat University. He spent 2002 as a Visiting Scholar at the ADSA Laboratory, Department of Computer Science, University of Calgary, Canada. He has published over 120 papers in refereed international journals and conferences. His primary work and research interests are in the areas of Data Mining, Social Network Analysis, Natural Language Processing and Machine Learning.



Mehmet K. Ozdemir [IEEE Senior Member] completed his BSc and MSc in Electrical Engineering at METU, Ankara Türkiye in ’96 and ’98 in, respectively. He received his PhD from Syracuse University, Syracuse, USA in ’05 from Electrical Engineering. He obtained his two-year Business Management from University of Toronto in ’10. Between ’99 and ’13 he has worked in industry in the area of Broadband Communication with the focus on CATV Systems, 4G Wireless Systems, and Cloud Solutions. Currently, Dr. Ozdemir works at Istanbul Medipol University as a full professor. He is the Chair for the artificial Intelligence Engineering Department and holds the Position of Associate Dean. Dr. Ozdemir’s research interests are Secure Wireless Link Designs, ISAC, AI Applications for Communication Systems and most recently Odor Sensing.



Jon Rokne is a Professor in the Department of Computer Science at the University of Calgary where he served as department Chair from 1989 to 1996. His research interests include Interval Analysis, Global Optimization, Computer Graphics and Social Networks. He is a co-editor of the prestigious Encyclopedia of Social Network Analysis and Mining. Dr. Rokne organized the installation of the First University FDDI network in Canada at the University of Calgary, and in 1990 he actively worked towards connecting the University to the World Wide Web (WWW). He spearheaded the acquisition and installation of a top 500 supercomputing facility in 1992. An IEEE member since 1970, Dr. Rokne has volunteered in a number of capacities. In the Computer Society, he completed two terms as Vice-President, Publications and three terms on the Board of Governors. He has served as Vice-President, IEEE Publication Services and Products Board and as a member of the Board of Directors of IEEE for two terms.

Reda Alhajj is a professor of Computer Science affiliated with the University of Calgary, Istanbul Medipol University, and the University of Southern Denmark. He served on the program committee of several international conferences. He is founding editor in chief of the Springer premier journal “Social Networks Analysis and Mining,” founding editor-in-chief of Springer Series “Lecture Notes on Social Networks”, founding editor-in-chief of Springer journal “Network Modeling Analysis in Health Informatics and Bioinformatics,” founding co-editor-in-chief of Springer “Encyclopedia on Social Networks Analysis and Mining,” founding steering chair of the flagship conference “The International Conference on Advances in Social Network Analysis and Mining”, and three accompanying symposiums FAB, FOSINT-SI and HI-BI-BI. He served as guest editor of a number of special issues and edited a number of conference proceedings. Dr. Alhajj’s research concentrates primarily on Data and Network Science from Management to Integration and Analysis. Current research efforts include: Data Management and Mining, Social Network Analysis with Applications in Sociology, Computational Biology and Bioinformatics, Homeland Security, etc., Sequence Analysis with Emphasis on Domains like Financial, Weather, Traffic, Energy, etc.