Detection of Fake News on Social Media Using Network Science Approach

Jhonathan Carvalho, Bruno Travençolo and Fabíola S. F. Pereira

[1]Faculty of Computing
Federal University of Uberlândia (UFU) – Uberlândia – MG – Brazil

jhonathanalves_br@hotmail.com,{travencolo,fabiola.pereira}@ufu.br

Abstract. *This paper addresses the detection of fake news on social networks by combining complex networks and artificial intelligence techniques. Recent works have shown progress in solving the problem of detecting fake news using deep learning, which, in general, are penalized by the lack of interpretability and require large amounts of labeled data. In addition, to represent instances, solutions in the literature generally use textual characteristics, social relationships, and information related to engagement on social media. However, there are still gaps to be explored regarding the most relevant features of a fake post taking as a premise the interpretability of the solution. We propose modeling data through ego networks, extracting features from the underlying network, matching with textual features, and using traditional machine learning algorithms to detect and identify fake news on social networks. The experiments, carried out on Twitter data using the popular fake news dataset – FakeNewsNet, show the potential of the proposed approach from the perspectives of interpretability, precision and recall.*

1. Introduction

The concept of fake news refers to false information or reports that are disseminated and presented as true facts [Gelfert 2018]. Social networks have become fertile ground for the spread of false information due to the speed at which information spreads.

Some characteristics within social networks contribute to this misinformation. Platforms like Twitter (now rebranded as X) prioritize content with greater potential for engagement, that is, posts with the highest number of likes, reactions, comments and shares tend to stand out within the platform. Fake news, as it contains aspects and traits of sensationalism, makes attention and interaction with this type of information more attractive to the public, leading to a large number of interactions, shares and comments by the population.

Fake news also tends to have sensationalist titles to get people's attention, with exaggerated or dramatic titles aimed at an immediate emotional reaction. However, it is possible to check the reliability of the information simply with the content of the post, as fake news often contains grammatical and spelling errors [Mourão and Robertson 2019]. These news are generally biased and are usually written in support of a certain opinion or point of view, which often only presents one side of the story.

In this paper we propose a new approach for detection of fake news on social media. We use indicators on social networks, text mining and network analysis for obtaining relevant characteristics for detecting fake news. The proposal validation was done using news detection within the social network Twitter, using the database called FakeNewsNet[1]. The investigation considered fake and true news about politics, assembling a complex network of each entity, retracing the connections of each relationship and building a targeted network around the news.

In this sense, this paper proposes a machine learning model for detection of fake news using network science, which has similar performance and greater interpretability than similar existing models in the literature. For this, artificial intelligence (AI) classifier models such as decision trees and deep neural networks were evaluated.

2. Related work

We listed here some works that are closely related to our proposal. However, the literature in the field is broader and extensive. [Capuano et al. 2023] surveys over several state-of-the-art machine learning and explores several features over multiple datasets/topics. It is worth highlighting some differences between fake news, disinformation and rumors. Fake news is information created with the intention of deceiving and with the aim of appearing to be true news, this information is often politically, financially, sensationalist and maliciously motivated [Godsey et al. 2021]. Misinformation refers to the intentional dissemination of false or misleading information with the express purpose of causing harm or creating a misleading impression. Finally, rumors are unverified information that can be false or true and are spread among people [Wang and Zhuang 2018].

The use of artificial intelligence with deep neural networks has been frequently used in research in the field of data analysis in recent years. In the work by [Krešňáková et al. 2019], deep neural networks and an LSTM (Long Short-Term Memory) model – which is a type of machine learning specialized in recurrent neural networks that helps machines understand long sequences of data, such as natural language sentences – achieved an accuracy of 91.25% from pre-processing data with textual information. The models were trained and evaluated on Fake News dataset obtained from Kaggle data repositories [Lifferth 2018]. The authors reported that the models achieved good performance and consistency. [Jehad and Yousif 2020] proposed the use of decision trees and random decision trees for fake news detection on two types of samples from the Fake News database obtained from Kaggle [Lifferth 2018] – the first in its raw form, and the second on the processed data, applying textual normalization. Using decision trees classifier, they achieved an accuracy of 78.13% for the raw dataset and 89.11% for the processed dataset. They also shown that the results using decision trees were superior than random decision trees.

Using network science and based on propagation patterns found on the FakeNewsNet database, [Meyers et al. 2020] present in their research that fake news propagates more quickly and deeply than real news. They concluded that the best combination of resources was the use of eight features: average followers and followed, average favorites, average users touched in the last 10 hours, last and first time, number of tweets, number of retweet and percentage of retweets. They used random tree classifier to classify fake and true news, achieving an accuracy of 87%.

In [Han et al. 2020], graph neural networks are used to classify fake news using features such as: whether the user is verified, the timestamp when the user was created,

[1] `https://github.com/KaiDMML/FakeNewsNet`

number of followers, number of friends, number of favorites, status, tweet timestamp, and number of seconds since the first tweet references the publication of the news. The database used was FakeNewsNet and they achieved an accuracy of more than 80%. When considering the characteristics of complex networks, their accuracy rised to 94%. The features used in their model were: in-degree, i.e. the number of users who mentioned user i; out-degree, i.e. the number of users who were mentioned by user i; weighted in-degree, i.e., the number of times user i has been mentioned; weighted out-degree, i.e., the number of times the user mentions others; the number of hop-2 neighbors; the number of external hop-2 neighbors; and the number of tweets collected from timeline. [Chandra et al. 2020] proposed the use of a graph neural network model to make predictions on the FakeNewsNet database. They achieve an F1-Score of 94.69%. The research and analysis takes into account graph convolutional network and graph attention network.

In the work by [Sales Santos and Pardo 2021], complex networks measures were explored combined with artificial intelligence models. Using the Fake.br-Corpus database with omplex networks features as input and the K-Means method, it was possible to obtain an accuracy of 69%.

[Kapadia et al. 2022] proposed a method named FND-NUP (Fake News Detection with News content, User profiles and Propagation networks). They test different models of graph neural networks and found that graph attention networks (GAT) perform well and provide explained results. The model was applied on FakeNewsNet. The news content, the users' profile features and the user's comments were used in the training. They obtained 96% F1-score using GAT on politifact dataset from FakeNewsNet. [Sedik et al. 2022] proposes the use of deep learning for fake news detection. They employed four different models: Convolutional neural network (CNN), Concatenated CNN (C-CNN), Long short-term memory (LSTM) and Gated Recurrent Units. The models were tested on Fake News Challenge from Kaggle repository. C-CNNs algorithm obtained an accuracy of 99.6%. [Xu et al. 2022] proposed a method named GET, an acronym for unified Graph-based sEmantic sTructure mining framework. They model evidences as a graph-structured data in order to capture the long-distance semantic dependency. With their model, they obtained a F1-Score of 69.1% on politifact dataset from FakeNetsNet.

Centrality and community measures in temporal networks have often been used as input for classifiers to train artificial intelligence algorithms to predict fake news. Textual temporal networks, for example, showed excellent results when they take into account the modeling structures for propagating rumors within the social network [Pereira 2021, Linhares et al. 2019].

3. Proposal

Our approach focuses on two main features used to represent news: first, several measurements are extract from complex network derived from tweets; and the second obtained from textual characteristics.

The validation of the proposal was performed using FakeNewsNet [Shu et al. 2017], the same dataset used in part of the related works describe before [Meyers et al. 2020, Han et al. 2020, Chandra et al. 2020, Kapadia et al. 2022, Xu et al. 2022]. FakeNewsNet is a widely used benchmark dataset for studying fake news detection, which combines news content, social context, and user engagement data – ideal to analyze how misinformation spreads.

3.1. Network science features

Graph theory is a branch of mathematics and computer science that studies the properties and behaviors of structures called graphs. Graphs are abstract representations of relationships between objects, and they consist of two main elements known as vertices and edges. Basically vertices are points that represent objects and edges are lines that form relationships between objects. They are typically used to model a wide variety of real-world situations, from computer networks and transportation systems to social networks and chemical structures. Graph theory provides tools and techniques for analyzing and solving problems related to these structures [Barabási and Pósfai 2016].

In this way, complex networks use graph theory to study systems composed of many interconnected elements that exhibit behaviors and properties that cannot be easily predicted from the individual characteristics of their components.

In complex networks, degree is a basic measure that indicates the number of connections that a node has with other nodes in the network. It is defined as the number of edges connected to a given node in the network [Barabási and Pósfai 2016]. For example, in a social network, a user's classification can be defined as the number of friends or number of followers, that is, the degree is an important measure to understand the structure and topology of a network and to identify central nodes or influential.

A scale-free network is a type of complex network in which the degree of nodes follow a power-law distribution, meaning that some nodes, known as hubs, have a very large number of connections, while most nodes have just a few connections [Barabási and Pósfai 2016]. Twitter cannot be considered a random network. Although the process of adding new followers can be considered random, the resulting network has a non-uniform degree distribution, which makes it a scale-free network.

By understanding how the web works and how information about it spreads, it is possible to develop strategies to limit the spread of false information and encourage the spread of accurate and useful information. Some users may actively post content, while others may simply consume content. Furthermore, there may be fluctuations in the number of followers, retweets and likes received by each user, which generates heterogeneity in the distribution of these values across the network.

In this sense, connection analysis in complex networks is crucial to understanding the structure and dynamics of the network, allowing the identification of patterns and properties of the network. Using the Louvain algorithm it is possible to identify communities. Communities are used to identify groups of people who connect more strongly with each other, that is, in social networks, for example, they are groups of people who are more connected to each other than to the rest of the network [Barabási and Pósfai 2016]. These groups can represent social circles, groups of friends, colleagues, interest groups, and other types of communities. Communities are excellent metrics to help identify the dissemination of information and opinions within social networks. In a given group, it is possible to understand the dynamics of influence and their power within the network.

In social networks, centrality refers to the level of importance of a node in relation to other nodes in the network. Centrality is an important measure in the theory of complex networks because it allows us to understand the position of a node in the network and the role it plays in the transmission and propagation of information within the network [Barabási and Pósfai 2016].

The degree of centrality refers to the number of connections that a node has to the other nodes in the network [Barabási and Pósfai 2016]. Nodes with higher centrality are considered more important in the network, as they have more connections and, therefore, greater capacity to influence other nodes. Centrality analysis within social networks is useful for identifying the most important and influential nodes in a network and for understanding the structure and dynamics of the network [Barabási and Pósfai 2016]. Identifying nodes with high centrality can be useful, for example, to understand the spread of information or opinions in a network, or to identify individuals with greater influence or leadership power.

The associativity coefficient is a measure of the correlation between the degrees of a node's neighbors in a complex network [Barabási and Pósfai 2016]. In the context of social networks, the associativity coefficient measures the propensity of a user's friends to have the same number of friends. Finally, the associativity coefficient is an important indicator in social network analysis, as it helps to understand the network structure and identify interest groups. In a social network, a high associativity coefficient may indicate the existence of well-defined groups or communities, while a low associativity coefficient may indicate a more random or less organized structure. For example, in a social network, the associativity factor can be used to understand how people group around common interests. When the membership score is high for a group of users who share a certain interest, it may indicate the existence of a well-defined community around that interest. On the other hand, a low associativity coefficient may indicate that these users are dispersed across the network and do not have a well-defined community structure [Barabási and Pósfai 2016].

In our analysis using network science, the network topology was analyzed in the perspective of egocentric networks, where the ego has the highest degree within the network, being responsible for exercising dominance over relationships [Morselli 2009].

An egocentric network is a subnetwork of a larger network centered on a single node called the ego. In social network analysis, the egocentric network is used to study an individual's connections with their immediate contacts, that is, their closest friends and their followers on a social network [Zhou and Zafarani 2019]. In short, news on social networks can be represented at some network levels, the first of which is the ego level, followed by the triad, community and finally the general network level [Zhou and Zafarani 2019].

In this sense, the network topology was designed as an egocentric network, taking into account the entity's circle of friends, such as its friends and friends of friends who have relationships in common, where the ego is the central node, thus the network analysis occurs from this node [Morselli 2009]. Figure 1 illustrates an ego network – the node with highest radius is the ego node and the direction of the edges indicates who is a follower and who the entity follows.

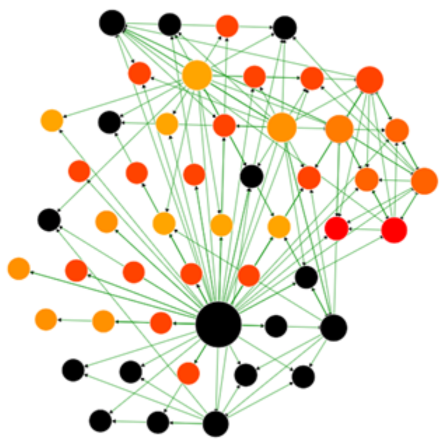

Figure 1. Example of an egocentric network created from data collected from the FakeNewsNet database. The ego node (the largest node) is the user that tweeted a news.

3.2. Textual features

Data preprocessing is a fundamental phase in data analysis, which aims to prepare initial data for more accurate and efficient analysis. This step is crucial, since the raw data collected is not ready for analysis, as it often contains errors, inconsistencies, missing values and other problems that can reduce the quality of the results obtained [Farhangfar et al. 2007]. Therefore, it is important to carry out this step carefully in order to ensure that the data used in the analysis is reliable and relevant to the assessment method.

After preprocessing, data transformation was performed in order to convert raw data into an analytical format. This step includes data normalization, standardization and categorical coding of data, preparing the information for input into AI models.

Next, the ten most frequent words from different parts of the tweets were extracted: news title (headline); news text; profiles text; locations words; text extracted from the timeline of each profile; text extracted from the description of each profile. Figure 2 shows the extracted words for each part of the tweet and the respective frequency of the word in the dataset.

In this sense, the word extraction part was synthesized from the entire dataset, where for each news the maximum number of repetitions of each word is extracted.

Figure 3 summarizes all features proposed in this paper that represent news to be used in machine learning algorithms.

4. Experiments

In this section we present the experimental setup to evaluate our proposal. The data extracted from the database FakeNewsNet were classified using decision trees, random decision trees and deep neural networks classifiers.

A total of 787 data were extracted, with a balance of 401 true news and 386 fake news.

News Title	
Word	Freq.
trump	91
obama	73
breaking	54
us	50
news	49
president	49
transcript	41
says	34
new	34
clinton	29

News text	
Word	Freq.
people	4315
going	4157
think	3844
president	3367
would	3182
said	3046
know	2993
us	2984
thats	2972
one	2783

Profile text	
Word	Freq.
news	13482
love	6402
us	4746
life	4586
world	4402
conservative	4396
im	3960
trump	3869
politics	3702
follow	3491

Locations	
Word	Freq.
usa	40073
united	14559
new	14407
states	12760
ca	11441
york	9268
washington	8780
ny	8297
dc	8116
california	7065

Timelines	
Word	Freq.
rt	2243076
new	240193
via	237662
trump	235458
us	201058
people	191489
one	187306
amp	182775
like	160482
dont	139364

Profile description	
Word	Freq.
news	25848
love	19997
life	12890
im	12457
us	11451
politics	10934
world	10760
retired	10469
trump	10365
proud	10148

Figure 2. Ten most frequent (freq.) words from different parts of the tweets: news title (headline); news text; profile text; locations; timelines; and profile descriptions.

social features	# tweets
	# retweets
textual features	title-trump, title-obama, ...
	body-people, body-going, ...
	profile-news, profile-love,
	location-usa, location-united, ...
	timeline-rt, timeline-new, ...
	description-news, description-love, ...
ego-networks features	# nodes
	# edges
	inDegree
	outDegree
	degree
	assortativity coefficient
	maxDegree
	avgDegree
	density

Figure 3. Features set used to represent an instance of news. The number of textual features varies with the size of the Vocabulary.

In all experiments, the dataset was divided into training and test samples according to the 70/30 split, widely adopted in the literature, as suggested in [Brown and Mues 2012]. The selection of the number of words was based on the work by [Carvalho et al. 2023].

4.1. Feature set

The following features were considered in all experiments.

- **Vocabularies**: consists of the top 10 words from each part of the news: title, news (body), profile text, location, timeline and profile description text. In total, 60 words are used. In this sense, classifying news using words consists of knowing whether there is a frequency of words based on the characteristics of the entry.
- **Complex Networks**: consists of extracting measurements from networks, such as number of nodes, number of links, communities, input centrality, output centrality, degree of centrality, maximum degree, average degree, density and coefficient of associativity.
- **Vocabularies and Complex Networks**: union of sets of characteristics of words and complex networks.

4.2. Machine Learning models and parameters

The machine learning models and the parameters used to identify fake news were:

- **Decision trees.**
  The decision tree was built using the following parameters: criterion as entropy, splitter as best, max_depth as None, min_samples_split as 2, min_samples_leaf as 1, min_weight_fraction_leaf as 2, max_features as None, random_state as None, max_leaf_nodes as None, class_weight as None, min_impurity_decrease to 0.0 and ccp_alpha to 0.0.
- **Random decision trees.**
  The decision tree was built using the following parameters: criterion like entropy, n_estimators as 100, max_depth to None, min_samples_split to 2, min_samples_leaf to 1, min_weight_fraction_leaf to 0.0, bootstrap as True, oob_score to False, n_jobs as None, random_state to None, max_features like sqrt, warm_start to False, class_weight to None, verbose as 0, max_samples to None, max_leaf_nodes to None, class_weight to None, min_impurity_decrease to 0.0, and ccp_alpha to 0.0.
- **Deep neural networks.**
  The neural network model used was a deep sequential machine learning using ReLu and sigmoid activation function, Adam optimizer, twenty percent dropout and a hidden layer with 121 neurons.

4.3. Additional Experiments

One additional experiment was also executed:

- **Varying the number of words.**
  We tested the use 10, 50, 100 and 200 words as Vocabulary features.

5. Results

The three models were applied to a dataset with the three groups of characteristics described, resulting in a total of nine main experiments. The results are described in what follows.

5.1. Decision trees

Using a decision tree classifier with Vocabulary and Complex Networks, iterating over a repetition of 100 times, it is possible to obtain results with an average F1-Score of 76.70%, best result of 78.97%, and worst evaluation of 73.64%, as shown in Table 1.

The next experiment only takes into account the use of Vocabulary features, that is, when Complex Networks are removed from the data characteristics, the evaluation drops significantly, proving that when Vocabulary and Complex Networks are used together there is a gain of approximately 10% as shown in the Table 1.

Finally, the last experiment using Decision Trees considered only the characteristics of Complex Networks. Table 1 demonstrates that when removing Vocabulary and keeping only network measurements data, the evaluation drops by around 16%.

Table 1. F1-Score using decision trees considering the three sets of evaluated characteristics.

Feature set	Best	Worst	Average
Complex Vocabularies and Networks	82.13%	84.25%	86.05%
Vocabularies	78.97%	73.64%	76.70%
Complex Networks	72.57%	67.82%	70.57%

5.2. Random decision trees

For the first prediction using random decision trees with Vocabulary characteristics and Complex Networks, iterating over a repetition of 100 times, it is possible to obtain results with an average F1-Score of 90.99%, best result of 92.94%, and worst rating of 89.34%, as shown in Table 3.

The next experiment only takes into account the characteristics of Vocabulary, removing Complex Networks features, the evaluation drops by around 11%, as shown in Table 3.

Finally, when Complex Networks are removed from the data, that is, using only Vocabulary, the evaluation drops by 6%, as shown in Table 3. Therefore, the use of Complex Networks on the set is important to obtain better results.

Table 2. F1-Score using random decision trees considering the three sets of evaluated characteristics.

Feature set	Better	Worse	Average
Vocabularies and Complex Networks	92,94%	89,34%	90,99%
Vocabularies	87,64%	82,78%	85,75%
Complex Networks	82,78%	77,44%	80,31%

5.3. Deep neural networks

Using deep neural networks and F1-Score as a performance evaluation method, Table 3 demonstrates that the model's best evaluation was 91.66% in an execution of 100 times, with an average of 89.63 % and worst rating if 86.06% with a total of 10 words.

When the characteristics of Complex Networks are removed from the data and using only the Vocabulary to classify the news, there is a 10% reduction in the worst evaluation, as shown in Table 3.

And when the characteristics of Vocabulary are removed, leaving only the features from Complex Networks, there is a reduction of almost 6% in the worst evaluation, as shown in Table 3.

Table 3. F1-Score using deep neural networks considering the three sets of characteristics evaluated.

Feature set	Better	Worse	Average
Vocabularies and Complex Networks	91,66%	86,06%	89,63%
Vocabularies	80,97%	75,86%	78,90%
Complex Networks	88,88%	80,00%	86,24%

5.4. Model comparision

Considering decision trees, random decision trees and deep neural networks, Table 4 demonstrates the comparison of experiments using Vocabularies and Complex Networks, and demonstrates that the results of random decision trees were superior in all scenarios.

Table 4. Comparison between decision trees, random decision trees and deep neural networks using F1-Score.

Model	Better	Worse	Average
Decision Trees	88,13%	84,25%	86,05%
Random Decision Trees	92,94%	89,34%	90,99%
Deep Neural Networks	91,66%	86,06%	89,63%

5.5. Varying the number of words

Additionally, experiments were carried out with different numbers of words in the Vocabulary. Using decision trees with different numbers of words, on an average of 100 executions, selecting the best evaluation among them, the evaluation tends to drop significantly, that is, with a sample of 70% for training and 30% for testing, the the results drops 1.63% when using Vocabulary of 200 words instead of 10 as shown in Table 5.

When using random decision trees, the performance also worsened, that is, the greater the number of words extracted from the dataset using random decision trees as a classification algorithm, in an average of 100 executions, selecting the best evaluation among them, the evaluation tends to fall, that is, with a sample of 70% for training and 30% for testing, the evaluation fell by 0.14% when a 200-word word is used instead of 10 as shown in Table 5.

When using deep neural networks as a classifier, in an average of 100 executions, selecting the best evaluation among them, the evaluation increased by around 0.84% when using 200 words instead of 10, that is, in a sample 70% for training and 30% for testing as shown in Table 5.

Table 5. Comparison using F1-Score increasing the frequency of words.

Model	10	50	100	200
Decision Trees	88.13%	88.00%	85.25%	86.50%
Random Decision Trees	92.94%	92.98%	92.68%	92.80%
Deep Neural Networks	91.66%	91.98%	92.43%	92.50%

The results demonstrate that depending on the artificial intelligence algorithm used to carry out the classification, it is more advantageous to select a smaller number of words than a greater diversity of words.

6. Conclusion

In this work we proposed a method to classify fake news using artificial intelligence models. Three classifiers was evaluated: decision trees, random decision trees and deep neural networks. The approach included the modeling of a complex networks, that is, the extraction of features based on ego networks. Also, the frequency of words extracted from the textual content is also considered in the classifications. As a result of these strategies, it was possible to obtain accuracy and precision above 90%, with an interpretative solution.

Considering the constructed complex networks, characteristics extracted from the networks were used, such as nodes, links, communities, centrality, assortativity, density and degree. The experiments demonstrate that the use of complex networks has a significant impact on the quality of the classification, and that the use of news related words is also important to achieve good performance. Our approach managed to identify relevant characteristics regarding the fake news detection task.

Finally, based on the discoveries and knowledge acquired during the research, it was possible to classify fake news using textual features combined with features extracted from ego networks.

Acknowledgements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (402543/2021-1;306436/2022-1).

References

Barabási, A. and Pósfai, M. (2016). *Network Science*. Cambridge University Press.

Brown, I. and Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, 39(3):3446–3453.

Capuano, N., Fenza, G., Loia, V., and Nota, F. D. (2023). Content-based fake news detection with machine and deep learning: a systematic review. *Neurocomputing*, 530:91–103.

Carvalho, H. C. F. B., Pitangui, C. G., Dorca, F. A., Oliveira, C. S., Trindade, E. A. C., Andrade, A. V., and Assis, L. P. (2023). Probabilistic classification of educational videos considering comments: an experimental analysis on youtube. In *XII Congresso Brasileiro de Informática na Educação*, Passo Fundo. Brasil/Inglês. Uma escola para o futuro: tecnologia e conectividade a serviço da educação.

Chandra, S., Mishra, P., Yannakoudakis, H., Nimishakavi, M., Saeidi, M., and Shutova, E. (2020). Graph-based modeling of online communities for fake news detection.

Farhangfar, A., Kurgan, L. A., and Pedrycz, W. (2007). A novel framework for imputation of missing values in databases. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 37(5):692–709.

Gelfert, A. (2018). Fake news: A definition. *Informal Logic*, 38(1):84–117.

Godsey, D. D., Hu, Y.-H., and Hoppa, M. A. (2021). *A Multi-layered Approach to Fake News Identification, Measurement and Mitigation*, pages 624–642. Springer International Publishing.

Han, Y., Karunasekera, S., and Leckie, C. (2020). *Graph Neural Networks with Continual Learning for Fake News Detection from Social Media*. arXiv– https://arxiv.org/abs/2007.03316.

Jehad, R. and Yousif, S. A. (2020). Fake news classification using random forest and decision tree (j48). *Al-Nahrain Journal of Science*, 23(4):49–55.

Kapadia, P., Saxena, A., Das, B., Pei, Y., and Pechenizkiy, M. (2022). Co-attention based multi-contextual fake news detection. In Pacheco, D., Teixeira, A. S., Barbosa, H., Menezes, R., and Mangioni, G., editors, *Complex Networks XIII*, pages 83–95, Cham. Springer International Publishing.

Krešňáková, V. M., Sarnovský, M., and Butka, P. (2019). Deep learning methods for fake news detection. In *2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo)*, pages 000143–000148.

Lifferth, W. (2018). *Fake News*. Kaggle – https://kaggle.com/competitions/fake-news.

Linhares, C. D. G., Ponciano, J. R., Paiva, J. G. S., Travençolo, B. A. N., and Rocha, L. E. C. (2019). Visualisation of structure and processes on temporal networks. In Holme, P. and Saramäki, J., editors, *Computational Social Sciences*, pages 83–105. Springer International Publishing, Cham.

Meyers, M., Weiss, G., and Spanakis, G. (2020). Fake news detection on twitter using propagation structures. In van Duijn, M., Preuss, M., Spaiser, V., Takes, F., and Verberne, S., editors, *Disinformation in Open Online Media*, pages 138–158, Cham. Springer International Publishing.

Morselli, C. (2009). *Inside Criminal Networks*. Springer New York.

Mourão, R. R. and Robertson, C. T. (2019). Fake news as discursive integration: An analysis of sites that publish false, misleading, hyperpartisan and sensational information. *Journalism Studies*, 20(14):2077–2095.

Pereira, F. (2021). Caracterização da propagação de rumores no twitter utilizando redes textuais temporais. In *Anais do X Brazilian Workshop on Social Network Analysis and Mining*, pages 25–31, Porto Alegre, RS, Brasil. SBC.

Sales Santos, R. L. and Pardo, T. A. S. (2021). Structural characterization and graph-based detection of fake news in portuguese. In *Anais do XIII Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*, pages 199–208. SBC.

Sedik, A., Abohany, A. A., Sallam, K. M., Munasinghe, K., and Medhat, T. (2022). Deep fake news detection system based on concatenated and recurrent modalities. *Expert Systems with Applications*, 208:117953.

Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017). Fake news detection on social media: A data mining perspective. *SIGKDD Explor. Newsl.*, 19(1):22–36.

Wang, B. and Zhuang, J. (2018). Rumor response, debunking response, and decision makings of misinformed twitter users during disasters. *Natural Hazards*, 93(3):1145–1162.

Xu, W., Wu, J., Liu, Q., Wu, S., and Wang, L. (2022). Evidence-aware fake news detection with graph neural networks. In *Proceedings of the ACM Web Conference 2022*, WWW '22, page 2501–2510, New York, NY, USA. Association for Computing Machinery.

Zhou, X. and Zafarani, R. (2019). Network-based fake news detection: A pattern-driven approach. *SIGKDD Explor. Newsl.*, 21(2):48–60.