Facebook and Fake News Detection: A Survey of Machine Learning, Classification Algorithms, and Dataset Assembly

Laura M. Saxton
31 July, 2023
Assignment 1
RMPP_2023, Unit 7

Table of Contents

1. Introduction

The proliferation of 'fake news' on social media has presented challenges to the integrity and veracity of political information on a scale larger than previous epochs of informational modal shift (Deibert, 2019; Ling & Yao, 2020). It's disruption of the traditional information pipeline has been cited as a direct threat to democratic government systems in the West, and other areas of the world (Deokar et al., 2019; Botha & Pieterse, 2020). Facebook, as one of the most popular social media sites in the world with 3 billion users (Meta, 2023), is at the forefront of companies' attempts to detect and mitigate fake political news on their platforms.

The following literature review aims to discuss what is known about Facebook's role in fake news dissemination, and to explore tools utilized by the platform to detect fake news. The review will then shift focus to Machine Learning (ML) as a tool for fake news detection on social media sites like Facebook. The processes of dataset creation, algorithm deployment, and result evaluation will be critically explored alongside three real-world datasets (BuzzfeedNews, ISOT, and LIAR), three popular classification algorithms (Naive Bayes, Random Forest, and SVM), and the evaluative techniques that are used to verify their effectiveness.

2. Facebook and Fake News

Since the debut of Facebook's *Newsfeed* feature, the platform has been linked to political and electoral tampering through the spread of 'fake news' in multiple countries (Table 1; Ali & Zain-ul-abdin, 2021; Anderson et al., 2021; Andreou et al., 2020; Baptista & Gradim, 2022; Chimuanya & Igwebuike, 2021; Ferrara et al, 2023), the best known of these cases perhaps being the US presidential election of Donald Trump and the UK parliamentary referendum win to exit the European Union ('Brexit'), both in 2016. Though evidence exists of Facebook's news feed fostering the dissemination of news silenced by draconian free speech laws (see examples in Cambodia (Chunly, 2020), the Arab Spring protests (Arafa & Armstrong, 2016), and Latin America (Jackson, 2014)), ex-employees of the company have reported “an unwillingness to recognize and fix [fake news and advertising] problems on a corporate level” (Karppi & Nieborg, 2021: 2640) which contributes to its continued role as a distributor of fake news online.

Though gaps may exist at the corporate level, Facebook has made a public effort to identify political fake news on its platform through manual detection and machine learning (Iosifidis & Nicoli, 2019). Human operators, either expert or crowd-sourced (Asr & Taboada, 2019; Demartini et al., 2022), are employed to manually sift through user posts and news articles, with crowd-sourced success rates ranging between 53% and 78% (Islam et al, 2023). ML algorithms show consistently higher rates of success (Hajli et al., 2022; Lee et al., 2019; Gupta et al., 2022), though Facebook itself has not yet not made public which algorithms are used in its detection tools (Grandinetti, 2023) prohibiting a precise estimation of reliability.

Table 1: Political Processes Influenced by Fake News

Year	**Country**	**Political Action**
2016	United States	Presidential Election
2016	United Kingdom	Referendum
2018	Brazil	Presidential election
2019	Portugal	Presidential election
2019	Nigeria	Presidential election
2023	Russia	Invasion of Ukraine

Although Facebook has not yet released its own ML processes, ML researchers have noted a difficulty in categorizing fake news (Zafarani & Zhou, 2020) and employing uniform algorithms for fake news detection (Benevenuto et al., 2019). Various definitions of fake news can be found in literature often differing not in the breadth of content categorized, rather in the amount of categories and subcategories assigned (Anderson et al., 2021; Carley et al., 2019; Ding et al., 2016; Ghorbani & Zhang, 2020; Gupta & Kaur, 2022; Pennycook et al., 2021; Zafarani & Zhou, 2020). As will be discussed in more depth below, such subtle differences can make uniform quantification among datasets and algorithms difficult (Repede, 2023; Bilal et al., 2022).

While various semantic differences permeate throughout, generalized categories themselves are fairly stable. *Figure 1* (adapted from Lim et al., 2018; Iosifidis & Nicoli, 2020; & Hajli et al., 2022) presents these categories alongside generalized categories for actors who perpetuate disinformation on social media. In addition to dataset challenges, machine learning algorithms are often found to behave differently depending on which datasets and evaluative measures are employed (Casillo et al., 2020 ; Lee et al., 2019)

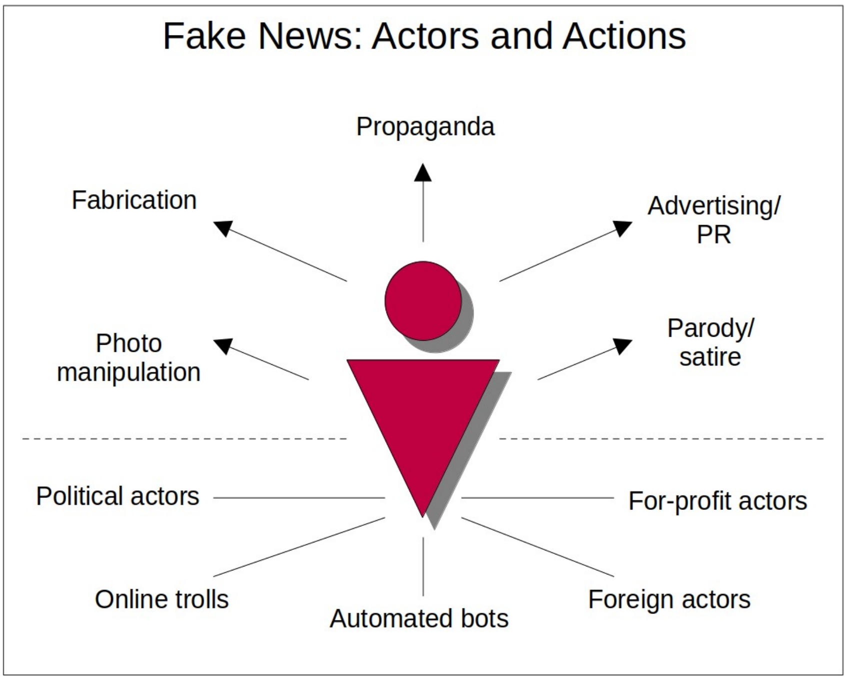

Figure 1: Fake News Categories and Perpetrators

It should be assumed that Facebook faces these obstacles in its own curation of fake news datasets and utilization of ML algorithms. It is thus important to understand common datasets and algorithms utilized for fake news detection, their benefits and limitations, as well as the challenges present in classification evaluation and fake news detection in controlled contexts. To do so, this paper will now look more deeply at the research surrounding fake news datasets and machine learning algorithms to relate these academic studies to possible obstacles faced by Facebook.

3. Machine Learning and Fake News Detection

This review is concerned with traditional ML classification algorithms, as their commonality suggests they could be utilised in some capacity at Facebook. ML for classification algorithms involves four phases, two in the 'dataset' sphere and two in the 'algorithm' sphere (Figure 2; Hirlekar & Kumar, 2020), each of which will be investigated in the sections below.

3.1 Machine Learning Datasets

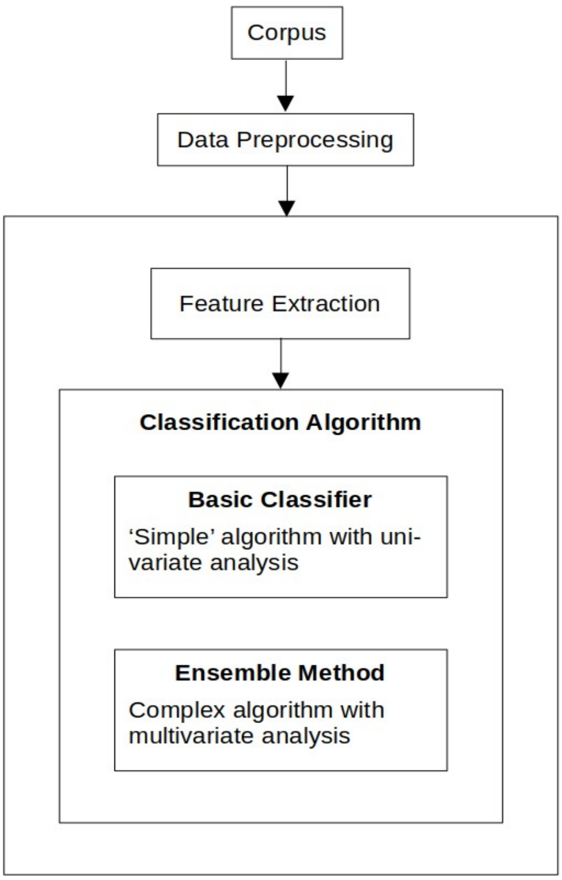

Figure 2: Machine Learning Process

Perhaps the most informative decision of a dataset's characteristics is how a corpus is compiled and annotated. Differences in labeling have been shown to affect feature extraction (Alazab et al., 2021), algorithm selection (Huang, 2020), and nuance when determining fake news (Broniatowski et al., 2019). In this section three datasets (Table 2; Lopez at al., 2021) with multiple appearances in peer-reviewed studies (Alazab et al., 2020; Gantara & Shah, 2022) have been chosen to demonstrate the acuity of these differences when applied to classification algorithm learning.

Annotation of corpus data is often performed by manual annotators (Lopez et al., 2021).

Expert and crowd-sourced annotators can provide reliable multiclass labeling, but, given the scope of fake news across platforms like Facebook, expert manual fact checking can often be time consuming and costly (Caschera et al., 2021; Islam et al., 2023). As a result, the strain of scalability may lead to a decrease in the accuracy of labeling. In addition, crowd-sourced labeling, which can be employed to mitigate scalability issues, has shown disparate rates of success ranging between 53% and 78% (Islam et al, 2023). Because ML algorithms are only as good as the dataset they are trained on (Grover & Misra, 2021), ensuring the accuracy and veracity of datasets is a fundamental requirement.

Table 2: Dataset Examples

Dataset	**Content Type**	**Annotator**	**Labeling**	**Quantity**
ISOT	Entire articles	PolitiFact	Binary	23,481
BuzzfeedNews	Social network posts – Facebook	Journalist team	Multiclass – four levels	2,282
LIAR	Short statements – political	PolitiFact	Multiclass – six levels	~12,800

The datasets in *Table 2* are examples of expert annotated datasets used within multiple research studies as benchmarks for algorithm or feature extraction testing. The BuzzfeedNews and LIAR sets are designed as a multiclass datasets which an be employed with ensemble or binary classifiers depending on the type of feature vectorisation employed (Bilal et al., 2022). This allows more nuance, notably with satire or parody content (Abonizio et al, 2019), but these datasets have been shown to have some weaknesses when deployed. The LIAR dataset “failed to show feature extractions using semantic features” (Alazab et al., 2020: 50) when used with DSSM-LSTM, while the BuzzfeedNews dataset “did not demonstrate the feature extraction using content and context” (Alazab et al., 2020: 50) under matrix-tensor factorisation.

The ISOT dataset is a binary dataset that has shown strong performance when employed among binary algorithms (Gantara & Shah, 2022), often ranging between the mid-80s to 99% accuracy. While this shows positive results, the dataset is limited to "using only text-based features to classify fake news" (Alazab et al., 2020: 50). These limitations demonstrate the attention to detail required when selecting scalable data for feature extraction and algorithm deployment, one required, if not critical, aspect of such detail being data preprocessing.

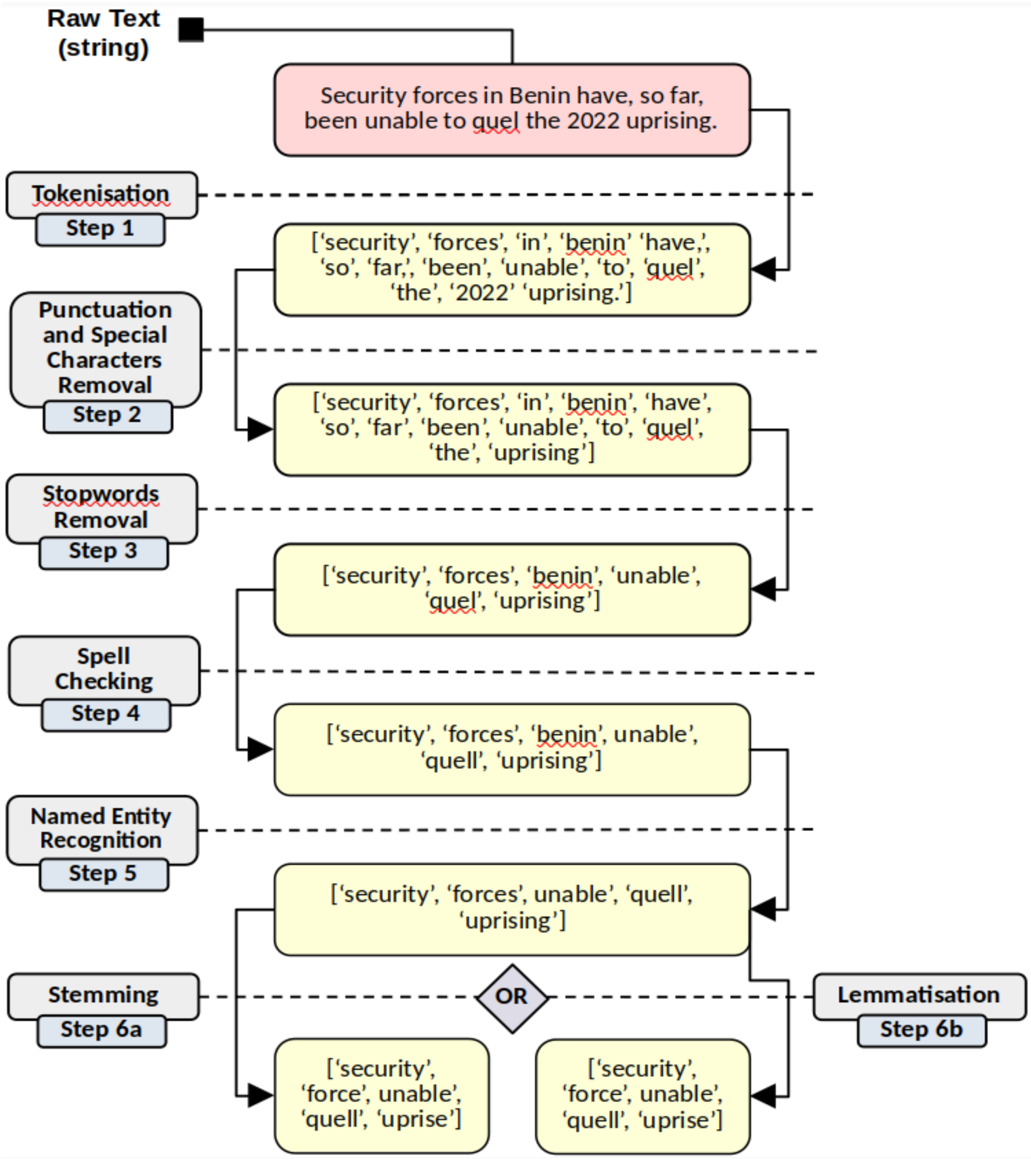

Figure 3: Data Preprocessing

Data preprocessing is a multi-stage filtration which prepares the content of a dataset for feature extraction (Figure 3; Lopez at al, 2021). Datasets without data preprocessing would have too much ‘noise’ in the dataset, burying useful content with filler such as pronouns, prepositions, and word endings. Without this process, the datasets above would be unable to successfully inform an algorithm how to identify fake news.

3.2 Machine Learning Algorithms

Once a corpus has been compiled and preprocessed, feature extraction of the dataset and algorithm learning take place. Each method of feature extraction focuses on unique qualities of the chosen dataset as foci for algorithmic learning and is required for vectorisation of the data into machine-readable values (Bilal et al, 2022). *Table 3* (Patil et al., 2021) has an overview of two popular extraction methods. These extraction methods influence the rate of success of the algorithms for which the data is extracted, which also vary in computational foci and ability (Table 4; Gupta et al., 2019; Kozak et al., 2021; Mathew, 2019). While deep learning algorithms are considered to be the cutting-edge of fake news detection (Hamid et al., 2021), this review is focused on traditional classification algorithms as these are older and more likely to be on Facebook platforms.

Table 3: Feature Extraction Methods

Methodology	**Process**
TF-IDF	Calculates word frequency in a document by comparing the relative frequency of a word in the document with its frequency across the dataset.
N-grams	All combinations of words and letters that can be found in a document. Can be a uni-gram, bi-gram, etc.

While classification algorithms can be deployed to detect fake news on social media platforms appropriate to scale, they are often limited in their ability to identify multiclass data because “there is not yet a standardized [sic] way to label false averages” (Repede, 2023:79). Thus, utilising a binary algorithm is a common approach, reducing the classification nuance of the dataset while intending to provide more reliable predictive qualities (Caschera et al., 2021; Hang et al, 2023; Repede, 2023). This is demonstrated by the consistently high accuracy scores for binary algorithms with TF-IDF extractions (see Gupta et al., 2022), and the higher rate of accuracy among univariate n-grams when compared to multivariate n-gram alternatives (Ahmed et al, 2017).

Table 4: Fake News Detection Algorithms

Algorithm	Process	Type
Naive Bayes (NB)	Uses Bayes theorem; ‘naive’ because variables are considered independent of each other	Basic - binary
Support Vector Machine (SVM)	Uses hyper-planes (decision boundaries) to classify data; can be used for regression and classification	Basic - binary
Random Forest (RF)	Uses decision trees as variables and requires individual training subsets with a random draw of attributes	Ensemble - multiclass

Table 5: Quantitative Algorithm Evaluation Methods

Measure	Quantification	Equation
Accuracy	Ratio of exactly predicted data to the all observations	$accuracy=\frac{TP+TN}{TP+TN+FP+FN}$
Precision	Ratio of exactly predicted data to the total number of positively predicted data	$precision=\frac{TP}{TP+FP}$
Recall	The number of true positive results	$recall=\frac{TP}{TP+FN}$
F1-Score	Mean between precision and recall; rates overall performance	$f1=\frac{TP+TP}{TP+TP+FP+FN}$

Assessing algorithm success involves a series of statistical equations (Table 5; Ganatra & Shah, 2022; Kozak et al., 2021) designed to measure the following variables (Hliang & Kham, 20XX): True Positive (TP): fake news data annotated as fake; True Negative (TN): true news data annotated as true; False Positive (FP): fake news data annotated as true; False Negative (FN): true news data annotated as fake. What can be concluded from these evaluative measures, when reviewing previous studies, is that the rate of accuracy for any single algorithm changes depending on the dataset and feature extraction method employed.

For example, Gupta et al. (2022) note that Argwal et al.'s 2019 study found an accuracy rate of 61% for SVM with n-gram and TF-IDF feature extractions and the LIAR dataset, while Yazdi et al.'s 2020 study found SVM with K means clustering to have a 95.34% rate with the BuzzfeedNews dataset and a 94.19% rate with the LIAR dataset. Similarly, while Jain & Kasbe's study (2018) found NB paired with a customized dataset and n-grams to have a 93.1% accuracy rate, Bilal et al. (2022) reported 78.17% accuracy with the ISOT dataset and a combined features extraction method. Results for Alazab et al.'s study (2021) found that RF with no feature extraction had a prediction accuracy of 25.92% with the LIAR dataset and 98.45% for the ISOT dataset. Finally, across 20 features and a 294,292 classification model sets, Benevenuto et al. (2019) found that just 2.2% of all algorithm and feature combinations tested had an AUC score of 0.85 or above.

These disparate results show "how hard it is for a single solution to tackle all forms of fake news stories" (Benevenuto et al., 2019: 26). If platforms such as Facebook are to implement fake news detection algorithms, the amount of variance in content across cultures, languages, and media could be cost and time prohibitive. One solution could be to employ binary algorithms to simplify the classification process, but given their lack of nuance awareness such tools may end up violating the ethical principle of free speech (Jacobs, 2022). Another issue may be the scalability of the datasets required for the regulation of 3.3 billion users; automating dataset collection may alleviate some of the burden, but as accuracy of these techniques vary benchmark datasets may necessary to ensure quality (Islam et al., 2023). Facebook would need to be aware of these challenges and continuously innovate their ML algorithm models to mitigate them to stay ahead of fake news dissemination.

4. Conclusion

In this literature review, Facebook's role in fake news dissemination and detection has been discussed as well as the various benefits and challenges of employing machine learning algorithms to fake news detection. While no simple solution to fake news detection exists, machine learning models demonstrate the ability to detect fake news at a rate higher than other scalable detection methods, such as crowd-sourcing. Special attention to dataset requirements and algorithmic feature selection should be paid if platforms such as Facebook wish to reliably detect fake news without losing the nuance of free speech.

5. References

Abonizio, H. Q., Barbon Jr., S., da Fonesca, A. A., de Morais, J. I., & Tavares, G. M. (2019) Deciding Among Fake, Satirical, Objective and Legitimate News: A Multi-label Classification System. In: *SBSI'19, May 20 -24, 2019, Aracaju, Brazil.* ACM: 1 - 8

Ahmed, H., Saad, S., & Traore, I. (2017) Detection of Online Fake News Using N-gram Analysis and Machine Learning Techniques. In: *ISDDC 2017.* Springer: 127 - 138

Alazab, M., Gadekallu, T. R., Hakak, S., Khan, S., Khan, W. Z., & Maddikunta, P. K. R. (2020) An Ensemble Machine Learning Approach Through Effective Feature Extraction to Classify Fake News. *Future Generation Computer Systems*, 117: 47 - 58

Ali, K. & Zain-ul-abdin, K. (2021) Post-truth Propaganda: Heuristic Processing of Political Fake News on Facebook during the 2016 US. Presidential Election. *Journal of Applied Communication Research*, 49 (1): 109 – 128

Anderson, A., Huhe, N., Preston, S., Robertson, D. J., & Shephard, M. P. (2021) Detecting Fake News on Facebook: the role of Emotional Intelligence. *PloS ONE*, 16 (3): 1 -13

Andreou, A., Benevenuto, F., Goga, O., Santos de Oliveria, L., Silva, M., Vaz de Melo, P. O. (2020) Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook. In: *WWW '20: The Web Conference 2020, April 20 – 24, 2020, Taipei, Taiwan*: 224 – 234

Arafa, M. & Armstrong, C. (2016) "Facebook to Mobilize, Twitter to Coordinate Protests, and Youtube to Tell the World": New Media, Cyberactivism, and the Arab Spring. *Journal of Global Initiatives: Policy, Pedagogy, Perspective*, 10 (1): 73 – 102

Asr, F. T. & Taboada, M (2019) Big Data and Quality Data for Fake News and Misinformation Detection. *Big Data & Society*: 1 -14

Baptista, J. P. & Gradim, A. (2022) Online Disinformation on Facebook: the Spread of Fake News During the Portuguese 2019 Election. *Journal of Contemporary European Studies*, 30 (2): 297 – 312

Benevenuto, F., Correia, A., Murai, F., Reis, J. C. S., & Veloso, A. (2019) Explainable Machine Learning for Fake News Detection. In: *WebSci '19, 30 June – 3 July, 2019, Boston, MA, USA.* ACM: 17 – 26

Bilal, M., Fayaz, M., Khan, A., & Khan, S. U. (2022) Machine Learning for Fake News Classification with Optimal Feature Selection. *Soft Computing*, 12: 7763 - 7771

Broniatowski, D. A., Diab, M., Hossieni, P., & Levi, O. (2019) Identifying Nuances in Fake News vs. Satire: Using Semantic and Linguistic Cues. *ArXiv*. Cornell: 1 – 5

Botha, J. & Pieterse, H. (2020) Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Study. In: *ICCWS 2020 15th International Conference on Cyber Warfare and Security*. ACPI: 57 - 66

Carley, K. M., Liu, H., Morstatter, F., Wu, L. (2019) Misinformation in Social Media: Definition, Manipulation, and Detection. *SIGKDD Explorations*, 21 (1): 80 - 90

Caschera, M. C., D'Ulizia, A., Ferri, F., & Grifoni, P. (2021) Fake News Detection: a Survey of Evaluation Datasets. *PeerJ Computer Science*: 1 - 34

Chimuanya, L. & Igwebuike, E. E. (2021) Legitimating Falsehood in Social Media: A Discourse Analysis of Political Fake News. *Discourse & Communication*, 15 (1): 42 – 58

Casillo, M., Colace, F., Conte, D., De Santo, M., Lombardi, M., Mottola, S., Santaniello, D. (2020) A Multi-feature Bayesian Approach for Fake News Detection. In: *International Computational Data and Social* Networks*:* 333 -344

Chunly, S. (2020) Social Media and Counterpublic Sphere in an Authoritarian State: Exploring Online Political Discussions Among Cambodian Facebook Users. *Discourse*, *Context & Media*, 34: 1 - 9

Deibert, R. J. (2019) The Road to Digital Unfreedom: Three Painful Truths about Social Media. *Journal of Democracy*, 30 (1): 25 – 39

Deokar, A. V., Gupta, A., Kauten, C., Qin, X., & Zhang, C. (2019) Detecting Fake News for Reducing Misinformation Risks Using Analytics Approaches. *European Journal of Operational Research*, 279: 1036 - 1052

Demartini, G., Nicolas, M., Papotti, P., Saeed, M., & Traub, N. (2022) Crowdsourced Fact-Checking at Twitter: How Does the Crowd Compare with Experts? In: *CIKM '22*, *17 – 21 October*, *2022*, *Atlanta*, *GA*, *USA*. *ACM*: 1736 - 1746

Ding, Y., Guo, B., Liang, Y., Yao, L., & Yu, Z. (2016) The Future of Misinformation Detection: New Perspectives and Trends. *Journal of the ACM*, 1 (1): 1 – 23

Ferrara, E., Jindal, N., Luceri, L., & Pierri, F. (2023) Propaganda and Misinformation on Facebook and Twitter During the Russian Invasion of Ukraine. In: *WebSci '23*, *30 April – 1 May*, *2023*, *Austin*, *TX*, *USA*. ACM: 65 – 74

Gantara, A. & Shah, M. N. (2022) A Systematic Literature Review and Existing Challenges Toward Fake News Detection Models. *Social Network Analysis and Mining*, 12: 1 - 21

Ghorbani, A. A. & Zhang, X. (2019) An Overview of Online Fake News: Characterization, Detection, and Discussion. *Information Processing and Management*, 57: 102025

Grandinetti, J. (2023) Examinging Embedded Apparatuses of AI in Facebook and Tiktok. *AI & Society*, 38: 1273 - 1286

Grover, *J. and Misra*, *R. (2021) Sculpting data for ML: The first act of Machine Learning*. *United States: Jigyasa Grover & Rishabh Misra*.

Gupta, A. K., Jain, A.; Khatter, H., & Shakya, A. (2019) A Smart System for Fake News Detection Using Machine Learning. In: *2nd International Conference on Issues and Chellendes in Intelligent Computing Techniques*. IEEE: 1 - 4

Gupta, R., Hong, W-C., Patel, K., Rohera, D., Sharma, R., Shethna, H., Tanwar, S., & Thakkar, U. (2022) A Taxonomy of Fake News Classification Techniques: Survey and Implementation Aspects. *IEEE Access*, 10: 30367 - 30394

Gupta, S. & Kaur, K. (2022) Towards Dissemination, Detection and Combating Misinformation on Social Media: a Literature Review. *Journal of Business & Industrial Marketing*: 1 -19

Hajli, N., Saeed, U., Shirazi, F., & Tajvidi, M. (2022) Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence. *British Journal of Management*, 33: 1238 – 1253

Hamid, A., Keya, A. J., Monowar, M. M., Mridha, M. F., & Rahman, S. (2021) A Comprehensive Review on Fake News Detection with Deep Learning. *IEEE Access*, 9: 156151 - 156170

Hang, J-Y., Jia, B-B., Liu, J-Y., & Zhang, M-L. (2023) Learning Label-Specific Features for Decomposition-based Multi-class Classification. *Frontiers Computer Science*, 17 (6): 1 -10

Hirlekar, V. V. & Kumar, A. (2020) Natural Language Processing Based Online Fake News Detection Challenges – A Detailed Review. In: *Fifth International Conference on Communication and Electronics Systems*. IEEE: 748 - 754

Hliang, M. M. M. & Kham, N. S. M. (2020) Defining News Authenticity on Social Media Using Machine Learning Approach. In: *2020 IEEE Conference on Computer Applications*: 1 - 6

Huang, J. (2020) Detecting Fake News with Machine Learning. *Journal of Physics: Conference Series*: 1 – 6

Iosifidis, P. & Nicoli, N. (2020) The Battle to End Fake News: A Qualitative Analysis of *Facebook* Announcements on How It Combats Disinformation. *The International Communication Gazette*, 82 (1): 60 - 81

Islam, R., Kabir, M. A., Rahman, A., & Tajrian, M. (2023) A Review of Methodologies for Fake News Analysis. *IEEE Access*, 11: 73879 – 73893

Jain, A. & Kasbe, A. (2018) Fake News Detection. In: *2018 IEEE International Students' Conference on Electrical, Electronics, and Computer Sciences*. IEEE: 1 - 5

Jackson, B. F. (2014) Censorship and Freedom of Expression in the Age of Facebook. *New Mexico Law Review*, 44(1): 121 – 168

Jacobs, L. G. (2022) Freedom of Speech and Regulation of Fake News. *Society of Comparative Law*. Oxford: 1 - 34

Karppi, T. & Nieborg, D. B. (2021) Facebook Confessions: Corporate Abdication and Silicon Valley Dystopianism. *new media & society*, 23 (9): 2634 -2649

Kozak, J. Probierz, B., & Stefański, P. (2021) Rapid Detection of Fake News Based on Machine Learning Methods. *Procedia Computer Science*, 192: 2893 - 2902

Lee, K., Lin, J., Mou, G., Tremblay-Taylor, G., & You, D. (2019) Detecting Fake News Articles. In: *2019 IEEE International Conference on Big Data*. IEEE: 3021 – 3025

Lim, Z. W., Ling, R., Tandoc Jr., E. C. (2018) Defining "Fake News": A Typology of Scholarly Definitions. *Digital Journalism*, 6 (2): 137 - 153

Ling, R. & Yao, M. Z. (2020) “What is Computer-Mediated Communication?” -- An Introduction to the Special Issue. *Journal of Computer-Mediated Communication*: 4 – 8

Lopez, M. A., Mattos, D. M. F., de Medeiros, D. S. V., de Oliveira, N. R., & Pisa, P. S. (2021) Identifying Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges. *Information*, 12 (38): 1 – 32

Mathew, T. E. (2019) Simple and Ensemble Decision Tree Classifier Based Detection of Breast Cancer. *International Journal of Scientific & Technology Research*, 8 (11): 1628 - 1637

Meta (2023) *Meta reports second quarter 2023 results*, *Meta*. Available at: https://investor.fb.com/investor-news/press-release-details/2023/Meta-Reports-Second-Quarter-2023-Results/default.aspx (Accessed: 01 August 2023).

Patil, S. M., Saran, S., & Sharma, U. (2021) Fake News Detection Using Machine Learning Algorithms. *International Journal of Engineering Research & Technology*, 9 (3): 509 - 518

Pennycook, G., Rand, D. G., & Ross, R. M. (2021) Beyond “Fake News”: Analytical Thinking and the Detection of False and Hyperpartisan News Headlines. *Judgment and Decision Making*, 16 (2): 484 – 504

Repede, Ș. E. (2023) Researching Disinformation Using Artificial Intelligence Techniques: Challenges. *Bulletin of “Carole I” National Defense University*, (2): 70 - 85

Zafarani, R. & Zhou, X. (2020) A Survey of Fake News: Fundamental Theories, Detection, Methods, and Opportunities. *ACM Computer Surveys*, 53 (5): 1 - 40