Kurdish Studies
December 2022
Volume: 10, No: 2, 906-916
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v10i2.3852

Strategies for Enhancing Fabricated News Detection: A Machine Learning Approach

Mohammad Rehbar Khan[1*] and Piyush Kumar Gupta[2]

[1*]Jamia Hamdard University
[2]Jamia Hamdard University

Abstract: Fake news, or fabricated news, poses a significant threat to trust in government and democracy by disseminating false information to influence political views. The issue under consideration receives considerable attention from researchers because of its increasing popularity. This review paper examines the existing advancements concerning countering fake news and introduces ensemble methodologies for the news articles binary classification. The research demonstrates the efficiency of machine learning algorithms, such as Passive Aggressive, Naïve Bayes, and Support Vector Machine classifiers, in the detection of fabricated news content. Nevertheless, it also underscores the constraints of basic classification techniques and underscores the necessity for specialized methodologies specifically designed for identifying fake news. This paper highlights the issue of limited data availability in this field and urges the need for effective techniques to differentiate between fake and authentic news sources.

Keywords: Fabricated News, Ensemble Techniques, Machine Learning, Binary Classification, Detection Methods, Trust, Democracy, Algorithm Evaluation.

1. Introduction

The identification and detection of fake news have become a topic of significant interest among researchers worldwide. Conversations on the internet about different events have become more common due to the widespread use of social networking platforms [4-9]. Individuals routinely engage in activities such as searching, retrieving, and discussing news events in their daily lives. Conversely, unpredictable events like environmental disasters or climate phenomena have the capacity to lead to unforeseen situations, such as the spread of misinformation. When people trust news from seemingly trustworthy sources like friends or family, spreading misinformation can cause confusion among the public.

There is an increasing need for automated systems capable of assessing the accuracy of news articles [25-27]. This is due to the difficulty in distinguishing between authentic and counterfeit news. One of supervised learning, Predictive active deep learning is a technique that allows controlling of the data by the learner used in the learning process. This is usually done by seeking advice from a knowledgeable human oracle [28-36]. This active learning process utilizes both labeled and unlabeled examples to create a small set of newly labeled data and a classifier. The main aim of this procedure is to accomplish classification accuracy of high level while simultaneously reducing the manual effort required for human annotation [35-41].

Fake news refers to deliberately spreading false information or hoaxes and has a historical origin in traditional print and broadcast media [40-50]. In the current era of digital technology, anyone can publish content on the internet, resulting in the widespread distribution of potentially misleading information [50]. Although machine learning and artificial intelligence are considered as potential solutions for detecting fake news, the task remains a complex and interdisciplinary challenge. Past studies have mainly determined on using machine learning methods to categorize online news and social media content, particularly in identifying fake news following events like the 2016 U.S. presidential elections. Straightforward content-based classification methods, along with techniques like n-gram and part-of-speech labeling, have been found to be inadequate for detecting fake news. 51-54 in the order mentioned. A more detailed examination of the content, incorporating context-free grammar and semantic analysis, has shown promise in enhancing classification accuracy [51-59]. Nevertheless, there is a need to enhance the efficiency of integrating machine learning techniques to tackle the complexities of detecting fake news.

Our hypothesis proposes that integrating machine learning with classification techniques can better the fake news detection, as evidenced by experiments carried out on publicly accessible datasets. By analyzing the frequency of words and using methods such as word cloud visualization, we can uncover trends within the news corpus. Furthermore, utilizing various sources of news data for testing and training allows us to assess the ability of our models to generalize [60-64].

Table 1 Fake News Classes

Sr.No.	Class	Description	Example
1	False Connection	Inconsistency with the content caused by headlines, images, or captions that could be misinterpreted.	A news article with a sensational headline but unrelated images.
2	False Context	Providing real information while incorporating inaccurate situational details, thereby distorting its original significance.	Sharing an old photo with a misleading caption to fit a narrative.
3	Manipulated Content	fabricating or distorting real data or images in order to mislead viewers.	Editing a image to include false details and mislead viewers.
4	Satire	Content intended for humor or entertainment but could be mistaken as genuine news.	A satirical news article with exaggerated claims for comedic effect.
5	Misleading Content	Presentation of information in a way that frames an issue or narrative to mislead the audience.	Selective quoting or misrepresentation of facts to sway opinion.
6	Imposter Content	Impersonation of genuine sources to spread false information.	Creating a fake social media account posing as a reputable news outlet.
7	Fabricated Content	Completely false content created with the intention to deceive and cause harm.	Fabricating a news story with fictitious events and sources.

Advancements in computer technology have facilitated the storage, processing, and access of vast amounts of data, laying the groundwork for machine learning applications [65-72]. Supervised, unsupervised, and reinforcement learning are the main approaches to learning from data, each serving distinct purposes in predictive modelling and pattern recognition [70-77]. Despite the potential of machine learning, challenges such as incomplete or biased data and algorithmic bias underscore the importance of careful implementation and continuous improvement.

2. Literature Review

Data mining, the utilization of machine learning methods to manage large databases, has gained significant popularity [1-16]. The primary aim of data mining is to extract valuable insights and patterns from extensive datasets, typically for the purposes of predictive modelling or decision-making. Nevertheless, it is imperative to comprehend that machine learning encompasses more than just the management of databases; it constitutes a fundamental element of artificial intelligence [17-25].

True intelligence encompasses the capacity to effectively adjust and respond to dynamic environments and varying circumstances, a characteristic that is naturally inherent in machine learning systems. These systems can enhance their performance without the need for explicit programming for each scenario by leveraging past experiences or example data [26-42]. The ability to adapt is especially vital in the current dynamic digital environment, where new challenges and contexts frequently arise.

Machine learning algorithms function by using parameters to construct models and subsequently optimizing them through the utilization of training data or prior experiences [43-49]. These models have the capability to generate predictions regarding future outcomes or offer descriptive insights into patterns observed in existing data. The inference process is guided by fundamental statistical principles, which enable algorithms to identify significant patterns from sample data.

Recently, there has been a growing trend toward the utilization of machine learning, deep learning, and natural language processing (NLP) methodologies for the purpose of identifying and identifying instances of misinformation [54-59]. This trend has been observed in recent times. Scholars have been compelled to investigate novel methodologies for detecting and reducing the propagation of inaccurate information on digital platforms as a result of the widespread dissemination of such information. This is an example of the application of machine learning algorithms, which can be seen in systems such as FakeNewsTracker. These algorithms are used to analyze data pertaining to social context and news. According to [1-11], this makes it possible to identify fabricated news using automated systems.

Several different methods of feature extraction and machine learning classifiers have been investigated by researchers in order to develop effective models for identifying instances of fake news [1-23]. The ability to accurately differentiate between genuine and fabricated content is possessed by these models. This capability is achieved through the examination of linguistic cues, social context, and propagation patterns. In spite of this, there are still challenges that need to be conquered, such as problems associated with the bias of the dataset, the selection of features, and the robustness of the algorithm. These difficulties bring to light the importance of performing continuous research and making improvements [12-50].

Furthermore, advancements in deep learning and natural language processing (NLP) have led to the development of more sophisticated models for identifying instances of fake news [31-50]. The intricate neural network architectures and language representations that these models employ are what allow them to extract nuanced semantic information from textual data.

The ability of these models to determine whether or not news articles and posts on social media are true can be improved by taking into account a variety of factors, including context, sentiment, and syntax.

Regardless these advancements, the detection of inaccurate information remains a formidable obstacle owing to the dynamic nature of misinformation tactics and the extensive volume of digital content accessible. Current research endeavors are focused on tackling these challenges through the investigation of innovative methodologies, the enhancement of current models, and the creation of comprehensive frameworks for the detection and mitigation of fake news [27-48]. Taking everything into consideration, machine learning continues to be an essential component in the fight against the problem of false information and fake news in the context of the digital industry. Scholars and professionals have the ability to develop strategies that are both resilient and expandable, with the goal of protecting the authenticity of digital information systems [1-50]. These strategies can be created through the utilization of data and algorithms. The achievement of this objective, on the other hand, calls for the collaboration of experts from a variety of fields, the ongoing development of innovative ideas, and a commitment to the implementation of responsible and ethical practices among machine learning professionals.

Table 2 Comparative Evaluation of Identifying Fake News techniques

Citation No.	Title of Article	Techniques Used	Uses
[6]	"A proposed deep learning model based on ensemble learning is recommended for detecting fake news about COVID-19 on Twitter."	BERTweet, Roberta, and Ensemble-based Deep Learning	Identifying misinformation relating to COVID-19 on Twitter using cutting-edge methods
[7]	"Fabrication of a False Information Identification Model Using N-Gram Analysis and Machine Learning Methods"	N-Gram Analysis, TF-IDF, Linear Support Vector Machine	Identifying false information with a precision of 92%
[10]	"Using Graph Convolutional Networks to detect and address misinformation and incidents in social networks."	Graph Convolutional Networks (GCNs)	Identifying misinformation, fraudulent accounts, and rumors in social networks with Graph Convolutional Networks (GCNs)
[12]	"Dual-language fake news detection model utilizing TF-IDF and N-Gram analysis"	TF-IDF, N-Gram Analysis, Linear Support Vector Classification	Identifying false information in Bengali and English languages with a precision of 93.29%.
[17]	"Using Machine Learning Classifiers to Identify False Information on Social Media Platforms"	Cognitive, Visual, Affective, and Behavioural Cues, Machine Learning Classifiers	Detecting misinformation on social media sites with an accuracy of 80%
[20]	"Investigating and Recognizing False Information on Social Platforms"	Data Collection, Visualization Techniques	Identifying and understanding misleading data through efficient visualization methods

3. Model Development

In this case, various classification models are considered, and to choose the best option and adjust its parameters, a series of experiments are conducted across different models. Initially, we explore classification models known for their effectiveness and success in similar sentence classification tasks. Some models, such Logistics Regression, are removed due to inadequate findings, while others such as SVM, naïve Bayes, as well as the Passive Aggressive show promising outcomes, leading to additional testing [1-23]. To assess accuracy, comparisons are made with results from other datasets using performance metrics.

3.1 Naïve Bayes

The model is recognized for its robustness in classification tasks, especially when applied to small datasets, owing to its efficient storage requirements. Nevertheless, the result may be compromised when presented with correlated words. The functioning of this model is predicated upon the assumption of feature independence. In simple terms, it computes the probability of an attribute being part of a particular class without taking into account the impact of other classes. Equation 1 depicts the Naïve Bayes formula, which provides an additional description [17-36].

$$P(\mathrm{q} \mid \mathrm{w}) = \frac{P(\mathrm{w} \mid \mathrm{q}) P(\mathrm{q})}{P(\mathrm{w})} \quad (1)$$

Where, Posterior Probability $P(\mathrm{q} \mid \mathrm{w})$ is calculated using the likelihood $P(\mathrm{w} \mid \mathrm{q})$, Class Prior Probability (P(q)), and Predictor Prior Probability (P(w)).

$$P\big((\mathrm{W} \mid Q_i\big) = \prod_{k=1}^{n} P(w_k \mid q_i) = P(q_1 \mid w_i) \times P(w_2 \mid q_i) \times \ldots \times P(q_n \mid w_i) \quad (2)$$

The classification method entails calculating the maximum posterior probability, P(q_i.| w), which indicates the highest probability of a specific class based on the observed data, using Bayes' theorem. This assumption implies that characteristics are independent, which simplifies calculations by just considering the class distribution and decreasing computing costs greatly. Naive Bayes is a commonly employed technique for discriminating between true and bogus news stories by assessing their correctness, specifically utilizing multinomial Naïve Bayes. It is crucial to acknowledge that there are other algorithms created to accomplish identical goals, and Naive Bayes is but one of them. The reason for its popularity lies in its efficacy in training classifiers for this specific purpose. Thus, Naive Bayes may be used to verify the credibility of news stories, along with other approaches.

3.2 The Support Vector Machine

SVM algorithm is utilized for supervised learning in both regression and classification tasks through the analysis of data. This algorithm analyses the data and categorizes it into separate groups according to its specific attributes. Support Vector Machines (SVM) which is shown in Figure 1, provide numerous benefits, such as rapid learning rate, exceptional precision, efficient categorization, and robustness against irrelevant features. It is notable for being one of the most thoroughly studied classifiers in current research. Furthermore, Support Vector Machines (SVM) exhibit exceptional efficacy in tackling the issue of identifying false information, rendering it a prominent option for such endeavours [24-49].

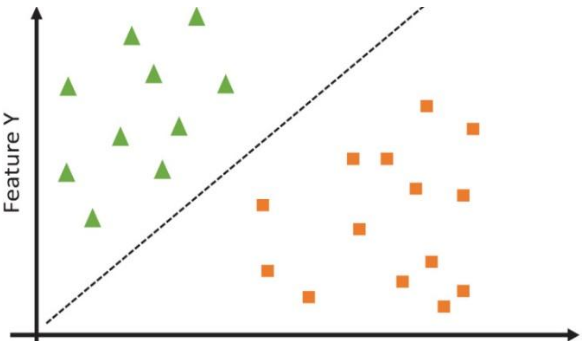

Figure 1 Support Vector Machine

3.3 Passive Aggressive algorithms

The algorithms are a class of machine learning techniques predominantly employed for classification purposes. Despite their simplicity in design, they have exhibited noteworthy performance across various domains. When compared to alternative methods like Online Perceptron and MIRA, Passive Aggressive algorithms have consistently proven their effectiveness [47-50]. Their versatility and reliability make them a valuable tool in tackling classification challenges.

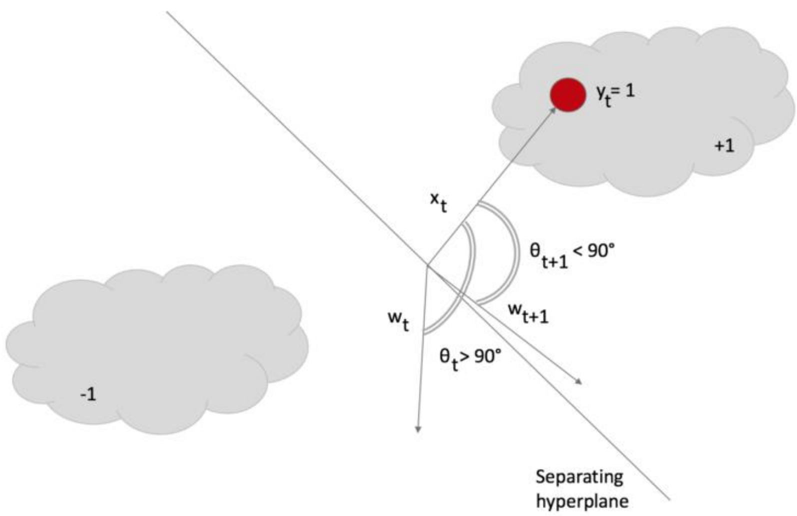

Figure 2 Passive Aggressive technique

3.4 Logistic Regression

The statistical technique known as LR is employed to estimate the association between variables. Especially efficient in binary classification tasks, it performs exceptionally well in situations where the classification entails two separate classes. This approach functions by constructing a mathematical representation, shown in equation 3, of the likelihood of a specific category or occurrence taking place, taking into account the values of independent variables. In order to commence the process of classification, Logistic Regression generally requires a substantial sample size in order to guarantee both robustness and accuracy. Due to its versatility and wide range of applications, this tool holds significant value across multiple domains, with a particular emphasis on binary classification tasks [99].

$$\text{Ln}\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \tag{3}$$

3.5 Natural Language Processing

The accuracy and overall performance of a classifier in Natural Language Processing (NLP) models can be significantly hindered by the presence of irrelevant or redundant features within a dataset. In order to address this problem, feature reduction techniques are utilized with the goal of optimizing the size of the text features. The process of reduction entails the exclusion of frequently occurring, low-information words such as "the," "and," "there," and "when," with the intention of directing attention towards words that possess greater significance. To enhance computational efficiency and model performance, the dataset's dimensionality is effectively reduced by setting a threshold for word frequency, commonly represented as "n," and implementing techniques such as lowercasing as well as the stop word removal. Due to the pressing nature of the issue, which is constantly changing, it is crucial to utilize deep learning algorithms such as TF-IDF and CountVectorizer. A number of algorithms have been developed to enhance the efficiency of text processing tasks and simultaneously enhance the overall performance of the classifier [99].

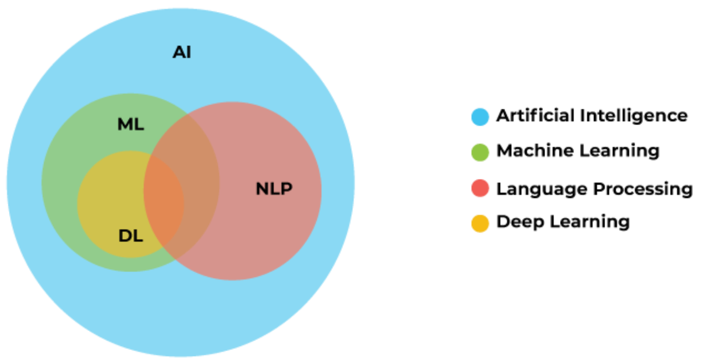

Figure 3 NLP Models

3.6 Decision Tree

The primary method used by the Decision Tree (DT) to achieve its main goal of effectively separating a region of potential observations is through a series of recursive dividing operations. Decision Trees are intended to mimic human cognitive processes, unlike algorithms like Support Vector Machines and neural networks. Consequently, they provide a classification method that is clear and transparent. The researchers utilized the Classification and Regression Trees (CART) model to achieve their objective of addressing the binary classification issue. CART uses the Gini index as a cost measure to assess the efficiency of feature splits. The Gini index is a statistical tool used to measure inequality within a data set. Computing the total squared probability for every class first. An evaluation is conducted on the information gain associated with each attribute to determine the most suitable attribute for classifying the dataset. The assessment commences with the computation of entropy. Every iteration chooses the characteristic to split the set according to the lowest entropy value. Entropy quantifies the amount of chaos in a system and is utilized to assess its organization. The reduction in entropy is quantified by information Gain that results from splitting a dataset according to a specific attribute. The change in entropy is a result of the division of the data. The attribute with the highest Information Gain value is selected for the split in the current iteration.

$$\text{Gini Index } = 1 - \sum_{i=1}^{n} p_i^2 \tag{4}$$

$$\text{Entropy } H(S) = \textstyle\sum_{c\in C} - p(c)\log_2 p(c) \tag{5}$$

3.7 Logistic Regression

Logistic Regression is a widely used classification algorithm that is well-suited for the task of assigning labels to observations into distinct classes. Logistic Regression is a suitable choice for the classification task due to its binary nature, demonstrating its effectiveness. The Logistic Regression variable is applied to a sigmoid function, which assigns a value for the probability that is then assigned to a particular of the classes in a discrete set. In regression analysis, logistic regression is a statistical technique used to estimate the parameters of a logistic model. These parameters are represented by the coefficients of a linear combination. The logistic model, as illustrated in Figure 4, is a statistical tool utilized to simulate the likelihood of an event occurring. A linear combination of several independent variables is used to represent the log-odds of the event. In the field of regression analysis, logistic regression entails estimating these parameters in order to forecast the likelihood of a particular event, like voting or not voting, based on a dataset of variables that are independent. Due to the probability of the outcome, the dependent variable has a bound between 0 and 1. [99].

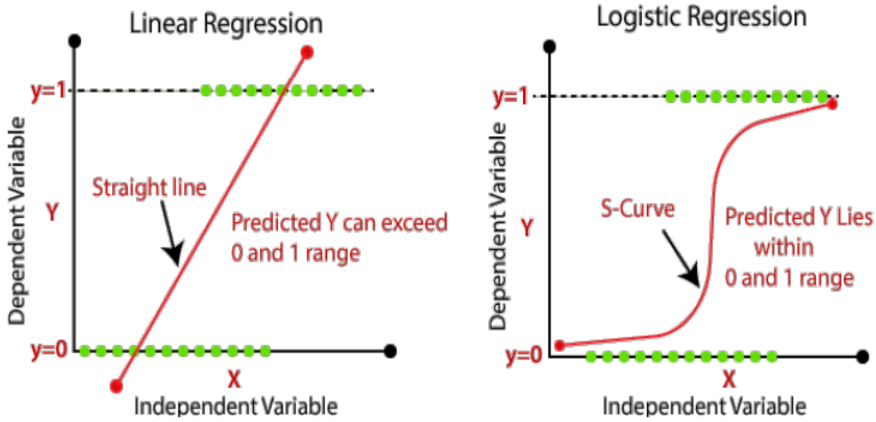

Figure 4 Graphical representation of Logistic Regression & Linear Regression

3.8 Bagging Classifier

The Classifier, which is also known as the Bootstrap aggregating classifier, is a widely recognized ensemble meta-estimator that enhances variance reduction. Figure 5 represents the bagging Classifier technique. The methodology involves implementing the base method on various subsets of the data, utilizing methods such as voting or averaging to combine individual predictions into a final output. Bagging is commonly employed in situations where a black-box technique, like decision tree, is utilized. Its purpose is to mitigate the significant variance that is frequently observed in these models [67-77].

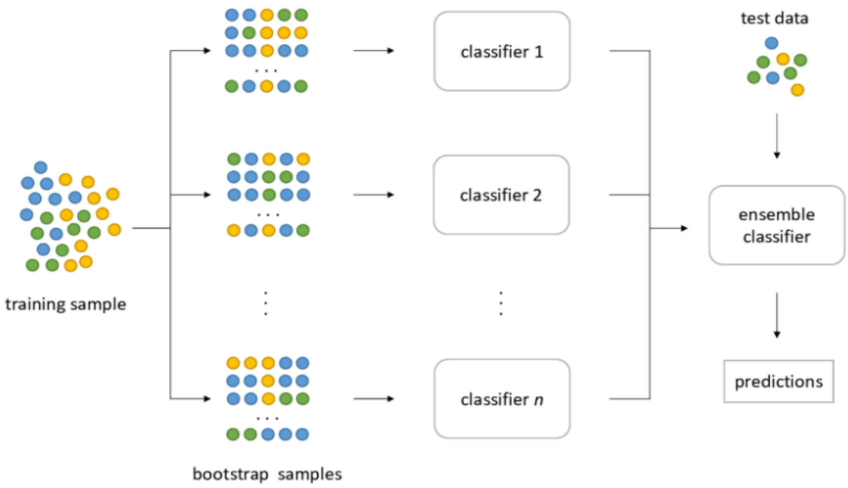

Figure 5 Bagging Classifier

4. Results

In our study, we conducted a comprehensive evaluation of different classification models for detecting fake news. The objective was to determine the most effective amalgam of features and classifiers to deal with the growing problem of misleading in online platforms.

The comparison table includes three classifiers: Naïve Bayes, Passive Aggressive, and Support Vector Machine (SVM). Out of the various classifiers considered, the Passive Aggressive classifier demonstrated superior performance in terms of accuracy, attaining a noteworthy accuracy rate of up to 93%. While Naïve Bayes and SVM also demonstrated competitive accuracy levels, Passive Aggressive exhibited the highest accuracy rate, indicating its efficacy in accurately classifying fake news articles.

Table 3 Comparison Table of Classifiers Utilized in Previous Studies

Classifiers	Accuracy (%)	Precision (%)	Recall (%)
Naïve_Bayes	85	89	87
Passive_Aggressive	93	92	89
Support_Vector_Machine	84	82	87

Table 3 presents a comparison of the performance traits of various classifiers utilized in our study to identify fake news. Each classifier is evaluated based on three main metrics: accuracy, precision, and recall. Accuracy refers to the classifier's overall correctness, precision is the ratio of correctly identified false news articles among all articles classified as fake, and recall is the percentage of actual fake news articles correctly identified by the classifier.

The Naïve Bayes algorithm attained an accuracy of 87%, a precision of 89%, and a recall rate of 85%. The Passive Aggressive model demonstrated a notable accuracy rate of 92%, accompanied by precision of 93% and recall of 89%. Achieving an accuracy of 84%, the Support Vector Machine model demonstrated precision and recall rates of 82% and 87%, respectively.

The measurements provide crucial insights into the effectiveness of each classifier in distinguishing genuine from deceptive news articles. The Passive Aggressive classifier achieved the highest accuracy rate compared to all other classifiers, as well as competitive precision and recall scores. The results demonstrate the effectiveness of Passive Aggressive in accurately categorizing fake news articles.

5. Conclusion

The review paper extensively explored different classification models to detect fake news. Various experiments were carried out to assess different feature sets and model selections in order to identify the most optimal combination for precise classification. This thorough investigation aimed to provide reliable methods to combat the spread of misinformation on online platforms.

Passive Aggressive (PA) stood out as the top performer among the classifiers examined, demonstrating impressive accuracy and precision. PA showed a small decrease in recall rate, suggesting an area that could be enhanced.

In comparison to the existing literature, the proposed classifiers demonstrated comparable levels of accuracy. Notably, the Passive Aggressive classifier exhibited an exceptional accuracy rate of up to 93%. The study excelled beyond previous methods in specific areas, highlighting the effectiveness of the suggested blend of characteristics and classifiers.

The study analyzed the impact of different classifier combinations on accuracy and discovered intriguing results regarding ensemble methods. Through meticulous experimentation, it was found that the Hard Voting Ensemble model was the most successful. A Decision Tree Classifier and Logistic Regression were incorporated into the model, resulting in an accuracy rate surpassing 88%. Applying logistic regression for binary classification and harnessing Bagging Classifier's variance-reducing capabilities notably enhanced classification accuracy.

The review study significantly contributes to the field of false news detection by offering valuable insights on the efficacy of various classifiers, as well as the importance of feature selection and model combination. Further research in this growing field should explore additional classifiers and datasets to improve the precision and resilience of fake news detection methods.

4.1 Future scope

The main goal of upcoming projects is to develop an automated fact-checking system to help individuals without experience verify news content. Various factors will be considered to conduct a comprehensive investigation into fake news. The factors encompass established facts, sources, subjects, URLs, geographical information, publication schedules, and the credibility of the sources. The aim is to increase the identification of false news by tackling existing challenges. This will be achieved by using advanced algorithms, large datasets, and interdisciplinary methods. This endeavor enhances the ongoing efforts to combat misinformation and uphold the integrity of information.

References

1. N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, 2015.
2. Arega, Kedir Lemma, and S. KotherMohideen. "Grouping and Detection of Fake News via web-based media Using Machine Learning in Amharic Language." CENTRAL ASIAN JOURNAL OF THEORETICAL & APPLIED SCIENCES 3.5 (2022): 89-110.
3. Arushi Gupta, Rishabh Kaushal, "Improving Spam Detection in Online Social Networks", 978-1-4799-7171-8/15/$31.00 ©2015 IEEE.
4. Mladenova, Tsvetelina, and Irena Valova. "Research on the Ability to Detect Fake News in Users of Social Networks." 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). IEEE, 2022.
5. Arvinder Pal Singh Bali, Mexson Fernandes, Sourabh Choubey, Mahime Goel, "Comparative Performance of Machine Learning Algorithms for Fake News Detection", Part of the Communications in Computer and Information Science book series (CCIS, volume 1046) (2022).
6. Kumar, Abhinav, Jyoti Prakash Singh, and Amit Kumar Singh. "COVID-19 Fake News Detection Using Ensemble-Based Deep Learning Model." IT Professional 24.2 (2022): 32-37.
7. H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," Lecture Notes in Computer Science Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pp.127–138,Oct.2017.
8. Khalil, Ashwaq, et al. "AFND: Arabic fake news dataset for the detection and classification of articles credibility." Data in Brief 42 (2022): 108141.

9. S. Feng, R. Banerjee, and Y. Choi, “Syntactic stylometry for deception detection,” in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, Association for Computational Linguistics, 2012, pp. 171–175.
10. Varlamis, Iraklis, et al. "A Survey on the Use of Graph Convolutional Networks for Combating Fake News." Future Internet 14.3 (2022): 70.
11. Tandoc, E. C., & Lim, Z. W., & Ling, R. (2018). Defining “fake news.” Digital Journalism, 6, 137-153.
12. Piya, Fahmida Liza, Rezaul Karim, and Mohammad Shamsul Arefin. "BDFN: A Bilingual Model to Detect Online Fake News Using Machine Learning Technique." Soft Computing for Security Applications. Springer, Singapore, 2022. 799-816.
13. Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., & Krause, A. (2018). Fake News Detection in Social Networks via Crowd Signals, 517–524. https://doi.org/10.1145/3184558.3188722
14. Bhartiya, Priyanka, et al. "Fake News Predictor Model-Based on Machine Learning and Natural Language Processing." Handbook of Research on Machine Learning. Apple Academic Press 479-501, (2022).
15. Dubey, Yogita, et al. "Framework for Fake News Classification Using Vectorization and Machine Learning." Combating Fake News with Computational Intelligence Techniques. Springer, Cham, 2022. 327-343.
16. Lorent, S. (2019). Master thesis: Fake news detection using machine learning.
17. Janze, C., Risius, M., “Automatic detection of fake news on social media platforms.” Pacific Asia Conference on Information Systems (2017)
18. Agarwal, Swati, and Adithya Samavedhi. "Profiling fake news: Learning the semantics and characterisation of misinformation." International Conference on Advanced Data Mining and Applications. Springer, Cham, 2022.
19. Kozik, Rafał, et al. "Technical solution to counter potential crime: Text analysis to detect fake news and disinformation." Journal of Computational Science 60 (2022): 101576.
20. Kai Shu, Deepak Mahudeswaran, Huan Liu, “FakeNewsTracker: a tool for fake news collection, detection, and visualization”, Computational & Mathematical Organization Theory, vol. 25 issue 1, pp. 60-71, (2022).
21. De Magistris, Giorgio, et al. "An explainable fake news detector based on named entity recognition and stance classification applied to covid-19." Information 13.3 (2022): 137.
22. Sedik, Ahmed, et al. "Deep fake news detection system based on concatenated and recurrent modalities." Expert Systems with Applications 208 (2022): 117953.
23. Marco L. Della Vedova, Eugenio Tacchini, Stefano Moret, Gabriele Ballarin, Massimo Di Pierro, Luca de Alfaro, “Automatic Online Fake News Detection Combining Content and Social Signals”, ISSN 2305- 7254, 2017, (2022).
24. Mykhailo Granik, Volodymyr Mesyura, “Fake News Detection Using Naive Bayes Classifier”, 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)
25. Megıas, David, et al. "Architecture of a fake news detection system combining digital watermarking, signal processing, and machine learning."
26. N. J. Conroy, V. L. Rubin, and Y. Chen, “Automatic deception detection: Methods for finding fake news,” Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, Nov. 2015.
27. Ahmad, Tahir, et al. "Efficient Fake News Detection Mechanism Using Enhanced Deep Learning Model." Applied Sciences 12.3 (2022): 1743.
28. Pranav Bhardwaj, Zongru Shao, “Fake News Detection with Semantic Features and Text Mining”, International Journal on Natural Language Computing (IJNLC) Vol.8, No.3, June 2019
29. Chen, Mu-Yen, Yi-Wei Lai, and Jiunn-Woei Lian. "Using Deep Learning Models to Detect Fake News About COVID-19." ACM Transactions on Internet Technology (2022).
30. Jain, A., & Kasbe, A. (2018, February). Fake news detection. In 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) (pp. 1-5). IEEE.
31. Al-Asadi, Mustafa A., and Sakir Tasdemir. "Using artificial intelligence against the phenomenon of fake news: a systematic literature review." Combating Fake News with Computational Intelligence Techniques (2022): 39-54.
32. Rubin, V., Conroy, N., Chen, Y., & Cornwell, S. (2016). Fake news or truth? Using satirical cues to detect potentially misleading news. In Proceedings of the Second Workshop on Computational Approaches to Deception Detection (pp. 7-17).
33. Patil, Dharmaraj R. "Fake News Detection Using Majority Voting Technique." arXiv preprint arXiv:2203.09936 (2022).
34. Michal Lukasik, Trevor Cohn, and Kalina Bontcheva. 2015a. Classifying tweet level judgements of rumours in social media. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP‟15). 2590–2595.
35. Bozuyla, Mehmet, and AKIN ÖZÇİFT. "Developing a fake news identification model with advanced deep language transformers for Turkish COVID-19 misinformation data." Turkish Journal of Electrical Engineering and Computer Sciences 30.3 (2022): 908-926.
36. Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. Detection and resolution of rumours in social media: A survey. arXiv preprint arXiv:1704.00656, 2017
37. Kumar, Sudhanshu, and Thoudam Doren Singh. "Fake News Detection on Hindi News Dataset." Global Transitions Proceedings (2022).
38. Zhou, X., Cao, J., Jin, Z., Xie, F., Su, Y., Chu, D., ... & Zhang, J. Real-Time News Certification System on Sina Weibo. In Proceedings of the 24th International Conference on World Wide Web (pp. 983-988). ACM.
39. Nassif, Ali Bou, et al. "Arabic fake news detection based on deep contextualized embedding models." Neural Computing and Applications.

40. Rubin, V. L., Chen, Y., & Conroy, N. J. Deception detection for news: three types of fakes. In Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community (p. 83). American Society for Information Science.
41. Mohtaj, Salar, and Sebastian Möller. "The Impact of Pre-processing on the Performance of Automated Fake News Detection." International Conference of the Cross-Language Evaluation Forum for European Languages. Springer, Cham.
42. Ruchansky, N., Seo, S., & Liu, Y. CSI: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (pp. 797-806). ACM.
43. Sciucca, Laura Della, et al. FakeNED: A Deep Learning Based-System for Fake News Detection from Social Media. International Conference on Image Analysis and Processing. Springer, Cham.
44. Janze, C., & Risius, M. Automatic Detection of Fake News on Social Media Platforms.
45. Nassif, Ali Bou, et al. "Arabic fake news detection based on deep contextualized embedding models." Neural Computing and Applications (2022): 1-14.
46. Hiramath, C. K., & Deshpande, G. C. Fake News Detection Using Deep Learning Techniques. In 2019 1st International Conference on Advances in Information Technology (ICAIT) (pp. 411-415). IEEE.
47. Sciucca, Laura Della, et al. "FakeNED: A Deep Learning Based-System for Fake News Detection from Social Media." International Conference on Image Analysis and Processing. Springer, Cham, 2022.
48. Bourgonje, P., Schneider, J. M., & Rehm, G. From clickbait to fake news detection: an approach based on detecting the stance of headlines to articles. In Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism (pp. 84-89).
49. Mondal, Subrota Kumar, et al. Fake News Detection Exploiting TF-IDF Vectorization with Ensemble Learning Models. Advances in Distributed Computing and Machine Learning. Springer, Singapore.
50. Priya S. Gadekar, "Fake News Identification using Machine Learning", Volume 7 Issue V, May 2019, International Journal for Research in Applied Science & Engineering Technology (IJRASET)
51. Del Ser, Javier, et al. "Efficient Fake News Detection using Bagging Ensembles of Bidirectional Echo State Networks." 2022 International Joint Conference on Neural Networks (IJCNN). IEEE.
52. Davuth, N., & Kim, S. R. (2013). Classification of malicious domain names using support vector machine and bi-gram method. International Journal of Security and Its Applications7(51-58)
53. Madani, Mirmorsal, Homayun Motameni, and Hosein Mohamadi. "Fake news detection using deep learning integrating feature extraction, natural language processing, and statistical descriptors." Security and Privacy: e264.
54. Banerjee, S., Chua, A. Y., & Kim, J. J. (2015, January). Using supervised learning to classify authentic and fake online reviews. InProceedings of the 9th International Conference on Ubiquitous Information Management and Communication (pp. 1-7).
55. da Silva, Fernando Cardoso Durier, Ana Cristina Bicharra Garcia, and Sean Wolfgand Matsui Siqueira. "A Systematic Literature Mapping on Profile Trustworthiness in Fake News Spread." 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2022.
56. Torunoğlu, D., Çakirman, E., Ganiz, M. C., Akyokuş, S., & Gü rbü z, M. Z. (2011, June). Analysis of preprocessing methods on classification of Turkish texts. In2011 International Symposium on Innovations in Intelligent Systems and Applications (pp. 112-117). IEEE.
57. Jain, Mayank Kumar, et al. "Review on Analysis of Classifiers for Fake News Detection." International Conference on Emerging Technologies in Computer Engineering. Springer, Cham, 2022.
58. Raulji, J. K., & Saini, J. R. (2016). Stop-word removal algorithm and its implementation for Sanskrit language. International Journal of Computer Applications, 150(2), 15-17.
59. Epstein, Ziv, et al. "Do explanations increase the effectiveness of AI-crowd generated fake news warnings?." Proceedings of the International AAAI Conference on Web and Social Media. Vol. 16. 2022.
60. Samonte, M. J. C. (2018). Polarity analysis of editorial articles towards fake news detection. ACM International Conference Proceeding Series, 108–112. https://doi.org/10.1145/3230348.3230354.
61. Bhatt, Shaily, et al. "Fake News Detection: Experiments and Approaches beyond Linguistic Features." Data Management, Analytics and Innovation. Springer, Singapore, 2022. 113-128.
62. Gencheva, P. et al. (2017) 'A context-aware approach for detecting worth-checking claims in political debates', in International Conference Recent Advances in Natural Language Processing, RANLP. doi: 10.26615/978-954-452-049-6-037.
63. Pranto, Protik Bose, et al. "Are You Misinformed? A Study of Covid-Related Fake News in Bengali on Facebook." arXiv preprint arXiv:2203.11669 (2022).
64. Patwari, A., Goldwasser, D. and Bagchi, S. (2017) 'TATHYA: A Multi- Classifier System for Detecting Check-Worthy Statements in Political Debates', in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. doi: 10.1145/3132847.3133150.
65. Verma, Pawan Kumar, et al. "MCred: multi-modal message credibility for fake news detection using BERT and CNN." Journal of Ambient Intelligence and Humanized Computing (2022): 1-13.
66. Gruppi, M., Horne, B. D., & Adalı, S. (2020). NELA-GT-2019: A Large Multi-Labelled News Dataset for the Study of Misinformation in News Articles. arXiv preprint arXiv:2003.08444.
67. Singhal, Shivangi, et al. "Leveraging Intra and Inter Modality Relationship for Multimodal Fake News Detection." Companion Proceedings of the Web Conference 2022. 2022.

68. Jindal, N., & Liu, B. (2008, February). Opinion spam and analysis. In Proceedings of the 2008 international conference on web search and data mining (pp. 219-230).
69. Najadat, Hassan, Mais Tawalbeh, and Rasha Awawdeh. "Fake news detection for Arabic headlines-articles news data using deep learning." International Journal of Electrical & Computer Engineering (2088-8708) 12.4 (2022).
70. Pathak, A., & Srihari, R. K. (2019, July). BREAKING! Presenting Fake News Corpus for Automated Fact Checking. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop (pp. 357-362).
71. Garcia, Gabriel L., Luis Afonso, and João P. Papa. "FakeRecogna: A New Brazilian Corpus for Fake News Detection." International Conference on Computational Processing of the Portuguese Language. Springer, Cham, 2022.
72. Singh, Kavinder, et al. "A Comprehensive Study on Data-Driven Fake News Detection Methods." 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2022.
73. Yumeng Qin, Dominik Wurzer and Cunchen Tang. "Predicting future rumours", Chinese Journal of Electronics. 2018
74. Sharma, Aniket, Ishita Singh, and Vipin Rai. "Fake News Detection on Social Media." 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, 2022.
75. Liddy, E. D, "Natural language processing", In Encyclopedia of Library and Information Science, 2nd Ed, NY, Marcel Decker, Inc., 2001.
76. Nistor, Andreea, and Eduard Zadobrischi. "The Influence of Fake News on Social Media: Analysis and Verification of Web Content during the COVID-19 Pandemic by Advanced Machine Learning Methods and Natural Language Processing." Sustainability 14.17 (2022): 10466.
77. Amini, Mahyar, et al. "MAHAMGOSTAR.COM as a Case Study for Adoption of Laravel Framework As the Best Programming Tool for PHP Based Web Development for Small and Medium Enterprises." Journal of Innovation & Knowledge, ISSN (2021): 100-110.
78. Daumé III, Hal. "A Course in Machine Learning." Published by TODO, 2014.
79. Li, M., Clinton, G., Miao, Y., & Gao, F. (2020). Short text classification via knowledge powered attention with similarity matrix based CNN. arXiv:2002.03350.
80. Loey, Mohamed, Mohamed Hamed N. Taha, and Nour Eldeen M. Khalifa. "Blockchain Technology and Machine Learning for Fake News Detection." Implementing and Leveraging Blockchain Programming. Springer, Singapore, 2022. 161-173.
81. Amini, Mahyar, et al. "Agricultural development in IRAN based on cloud computing theory." International Journal of Engineering Research & Technology (IJERT) 2.6 (2013): 796-801.
82. Galli, Antonio, et al. "A comprehensive Benchmark for fake news detection." Journal of Intelligent Information Systems (2022): 1-25.
83. Li, Qi. "Literature survey: domain adaptation algorithms for natural language processing", Department of Computer Science The Graduate Center, The City University of New York, pp. 8-10, 2012.
84. Sun, C., Qiu, X., Xu, Y., & Huang, X. (2020). "How to fine-tune bert for text classification." In Proc. China Nat. Conf. Chin. Comput. Linguistics, vol. 11856. Cham, Switzerland: Springer, pp. 194–206.
85. Patel, Ritik H., et al. "Detecting Fake News Using Machine Learning." Intelligent Data Communication Technologies and Internet of Things. Springer, Singapore, 2022. 613-625.
86. Amini, Mahyar, and Nazli Sadat Safavi. "Cloud Computing Transform the Way of IT Delivers Services to the Organizations." International Journal of Innovation & Management Science Research 1.61 (2013): 1-5.
87. Garg, Sonal, and Dilip Kumar Sharma. "Linguistic features based framework for automatic fake news detection." Computers & Industrial Engineering 172 (2022): 108432.
88. Nanath, Krishnadas, et al. "Examination of fake news from a viral perspective: an interplay of emotions, resonance, and sentiments." Journal of Systems and Information Technology (2022).
89. Ghafoor, Hafiz Yasir, et al. "Fake News Identification on Social Media Using Machine Learning Techniques." Proceedings of International Conference on Information Technology and Applications. Springer, Singapore, 2022.
90. Rahman, Shagoto, et al. "Efficient Machine Learning Approaches to Detect Fake News of Covid-19." Machine Intelligence and Data Science Applications. Springer, Singapore, 2022. 513-525.
91. Shi, L.-L., et al. "Human-centric cyber social computing model for hot-event detection and propagation." IEEE Trans. Comput. Social Syst., vol. 6, no. 5, pp. 1042–1050, Oct. 2019.
92. Hangloo, Sakshini, and Bhavna Arora. "Combating multimodal fake news on social media: methods, datasets, and future perspective." Multimedia Systems (2022): 1-32.
93. Gupta, Anchal, et al. "Empirical Framework for Automatic Detection of Neural and Human Authored Fake News." 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2022.
94. Baria, Reeya, et al. "Theoretical Evaluation of Machine And Deep Learning For Detecting Fake News." 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS). IEEE, 2022.
95. Shahid, Wajiha, et al. "Detecting and Mitigating the Dissemination of Fake News: Challenges and Future Research Opportunities." IEEE Transactions on Computational Social Systems (2022).
96. Bozkir, Efe, et al. "Regressive Saccadic Eye Movements on Fake News." 2022 Symposium on Eye Tracking Research and Applications. 2022.
97. Rodriguez, Guillermo Romera, Sanjana Gautam, and Andrea Tapia. "Understanding Twitters behavior during the pandemic: Fake News and Fear." arXiv preprint arXiv:2202.05134 (2022).

98. Waghmare, Akash Dnyandeo, and Girish Kumar Patnaik. "Social Media Fake News Detection using mNB in Blockchain." 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). IEEE, 2022.
99. Sharifani, K., Amini, M., Akbari, Y., & Aghajanzadeh Godarzi, J. (2022). Operating Machine Learning across Natural Language Processing Techniques for Improvement of Fabricated News Model. International Journal of Science and Information System Research, 12(9), 1-4.