Project Final Report
Fake News Detection Using Machine Learning Techniques

Team Members:

Rakshitha Reddy Gannu	: 11705430
Maanvi Reddy Muchantula	: 11705770
Akshaya Immadi	: 11717355
Nazeer Ahmed Mohammed	: 11645178
Gopi Muppuri	: 11689593

ABSTRACT

The emphasis of this initiative is developing an automatic system based on machine learning tools to identify false news. Spotting fake news stories has become a major difficulty in an age in which online channels can spread misinformation very fast. We start this problem as a binary classification challenge using text preprocessing, feature extraction with TF-IDF, and training several machine learning models to differentiate actual from false news. Along with Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, and XGBoost, our experiments feature models. Among these, SVM and Logistic Regression achieved the highest accuracy and ROC AUC scores, demonstrating that traditional machine learning methods combined with natural language processing can provide an effective solution for combating misinformation online.

INTRODUCTION

In the digital age, where data moves quickly and reaches huge groups of viewers in seconds, the spread of fake news is becoming one of the most urgent problems. Public opinion, elections, panic, or critical issues such as health and science are all influenced by misinformation that may also affect people. Particularly when confronted with the vast volume of internet material produced daily, traditional news verification techniques are slow and not feasible at scale. This underscores the pressing need for sophisticated, automatic systems able to help discriminate between false data.

The main objective of our project involved creating a machine learning system that would recognize true or fake news articles. Our team designed a typical text classification system that included data preprocessing after which we performed TF-IDF vectorization before executing various classification algorithms. We trained six different algorithms including Logistic Regression, Naive Bayes and Support Vector Machine (SVM) and Decision Tree, Random Forest and XGBoost. Environmental evaluation of multiple algorithms took place through the assessment of accuracy metrics plus precision and recall using F1-score and ROC AUC to measure performance excellence. Logistic Regression with Support Vector Machine proved to generate the best performance among all classified models.

our work belongs more generally to natural language processing and applied machine learning. Though many of these previous studies were devoid of comparative analysis or relied too much on deep learning, it further develops former research that has shown success with classical ML methods for text classification. Although deep learning shows promise, we concentrated on conventional approaches that are often quite powerful when supported by strong preprocessing and feature engineering, quicker to train, and simpler to interpret.

Finally, we show that using n-gram TF-IDF features and good text cleaning, conventional machine learning algorithms especially Logistic Regression and SVM can excel in spotting fake news. This paves the way for scalable, lightweight solutions able to help platforms, publishers, and the public fight internet falsification.

BACKGROUND

Modern digital platforms and social media networks have boosted information availability while generating a widespread dissemination of misleading information which people today refer to as "fake news." Fake news distributed through internet channels spreads like wildfire to thousands up to millions of readers within brief periods. Rapid information dissemination results in wrong perceptions that also manipulate people's feelings as well as occasional real-world outcomes.

Many bogus news stories about fake virus cures and virus minimization emerged on the internet throughout the COVID-19 pandemic. Fake headlines during election time claiming candidate withdrawals or scandals succeeded in misinforming massive groups of people before corrections from official sources became available. Highly dangerous fake news proves its capacity to deceive people through its professional-looking journalistic presentation style.

Fake news publishers create content that matches established credible sources through attention-grabbing emotional or sensational headlines to promote sharing. The delicate nature and diverse manifestations of fake news make its identification challenging for humans. An assessment must combine linguistic analysis of content structure with contextual indicators to gauge its authenticity. The domain of fake news detection operates between natural language processing (NLP) and machine learning alongside information retrieval for its implementation.

Research teams along with engineers use data mining along with machine learning methods to address this challenge during the past few years. The initial methods employed rule-based human intervention and keyword detection to identify fake news but these methods became inadequate because language and deception techniques adopted by writers continue to evolve. Complex systems extract text data patterns and analytical methods like word frequency analysis and phrase detection to determine fake or genuine article content through classification models.

These information systems aim to provide accurate classification in addition to fast processing with high scalability and adaptive analysis of misinformation patterns. A background review provides the foundation for our project which tests conventional machine learning models combined with TF-IDF feature extraction techniques on this important social responsibility task.

EXPERIMENT METHODOLOGY

A typical modern machine learning procedure led us to develop our fake news detection system through steps of data preparation followed by text preprocessing then feature extraction and model training and finally performance assessment.

Dataset

The available dataset included 1,267 news articles which were clearly classified as REAL or FAKE. The articles received classification as REAL or FAKE. Both an original text and cleaned version of text appeared in the dataset. We utilized original article content to replace any missing clean-up sections when necessary for maintaining complete texts.

Preprocessing

- Text conversion throughout the analysis used a lowercase format.
- Punctuation together with digits as well as special characters and URLs received removal during this step.
- A filtering process removed stopwords from the text to maintain its informative terms without noise.

Feature Extraction

The data conversion into numerical vectors used TF-IDF vectorization with Term Frequency–Inverse Document Frequency as its core algorithm.

Key settings:

- The ngram_range parameter value of (1, 2) allows for extracting unigrams and bigrams.
- The exclusion of frequently appearing words in the dataset occurred through max_df = 0.7.
- The application used 'english' as stop_words parameter to eliminate frequent words from the English language.

TEXT VISUALIZATION

The analysis of cleaned textual data structure required two straightforward visualizations for exploring frequent words within the database. The chart reveals the twenty most used words in the text. A word cloud presents an overview of the main dominant words found within the data set. The visual data confirms names alongside political keywords and emotional terminology appear throughout actual news content and false news content.

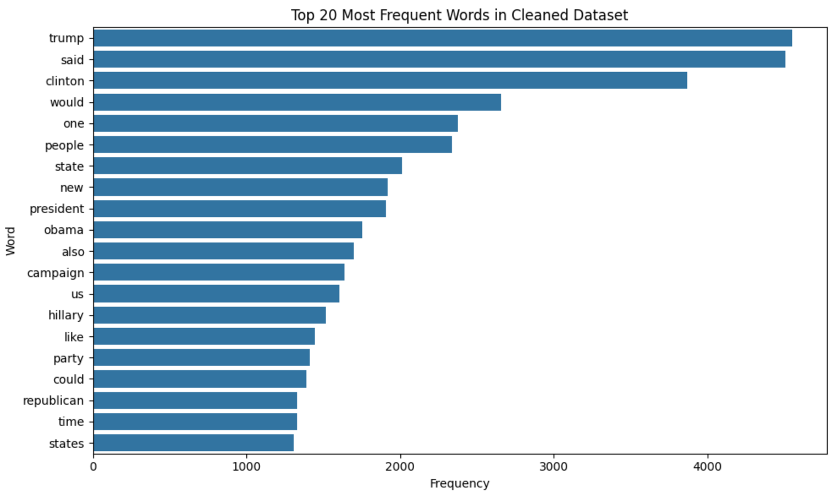

Figure 1: Top 20 Most Frequent Words in the Cleaned Dataset

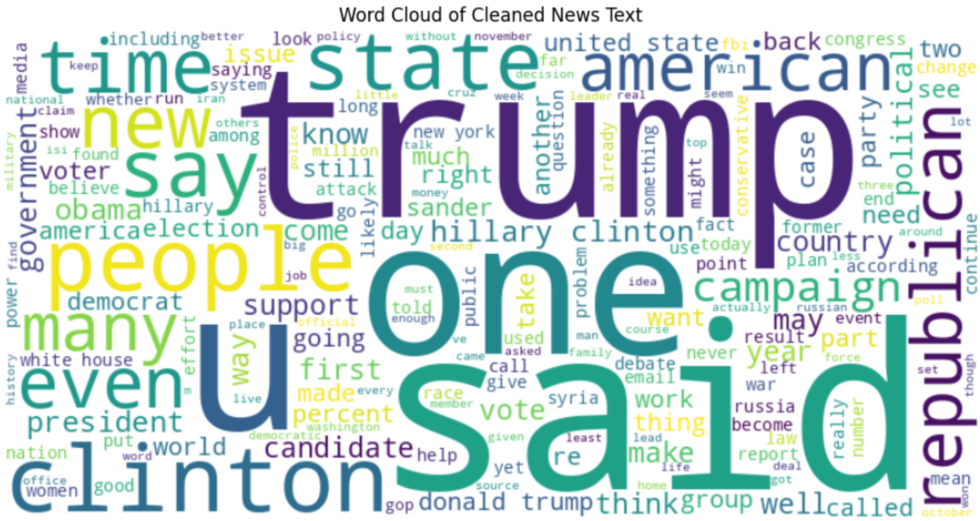

Figure 2: Word Cloud of Cleaned News Text

TRAIN - TEST SPLIT

Sampling distributed 80% of data points for training purposes while keeping 20% for testing to sustain the initial label distribution of the data.

Algorithms Used

We trained and evaluated the following six classification models:

1. Logistic Regression
2. Naive Bayes (MultinomialNB)
3. Decision Tree Classifier
4. Random Forest Classifier
5. Support Vector Machine (SVM)
6. XGBoost Classifier

All models were trained using default parameters, with one exception:
For XGBoost, we set use_label_encoder=False and eval_metric='logloss' for compatibility and performance tracking.

Evaluation Metrics

To assess and compare model performance, we used the following metrics:

- Accuracy – Overall correctness of predictions
- Precision – Ability to avoid false positives
- Recall – Ability to correctly capture true positives
- F1 Score – Harmonic mean of precision and recall
- ROC AUC – Area under the ROC curve; how well the model distinguishes between classes

These metrics allowed us to evaluate both the accuracy and robustness of each model. All evaluations were performed on the test set, and the results were stored for side-by-side comparison.

RESULTS

The six different models trained on equal datasets under the same TF-IDF feature set received standardized performance evaluations. All models get compared in a single visual representation through this table:

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC
Logistic Regression	0.913	0.957	0.867	0.910	0.978
SVM	0.913	0.965	0.859	0.909	**0.981**
Random Forest	0.906	0.919	0.891	0.905	0.962
XGBoost	0.902	0.912	0.891	0.901	0.957
Naive Bayes	0.898	0.925	0.867	0.895	0.964
Decision Tree	0.756	0.766	0.742	0.754	0.756

Table 1: Model Performance Metrics

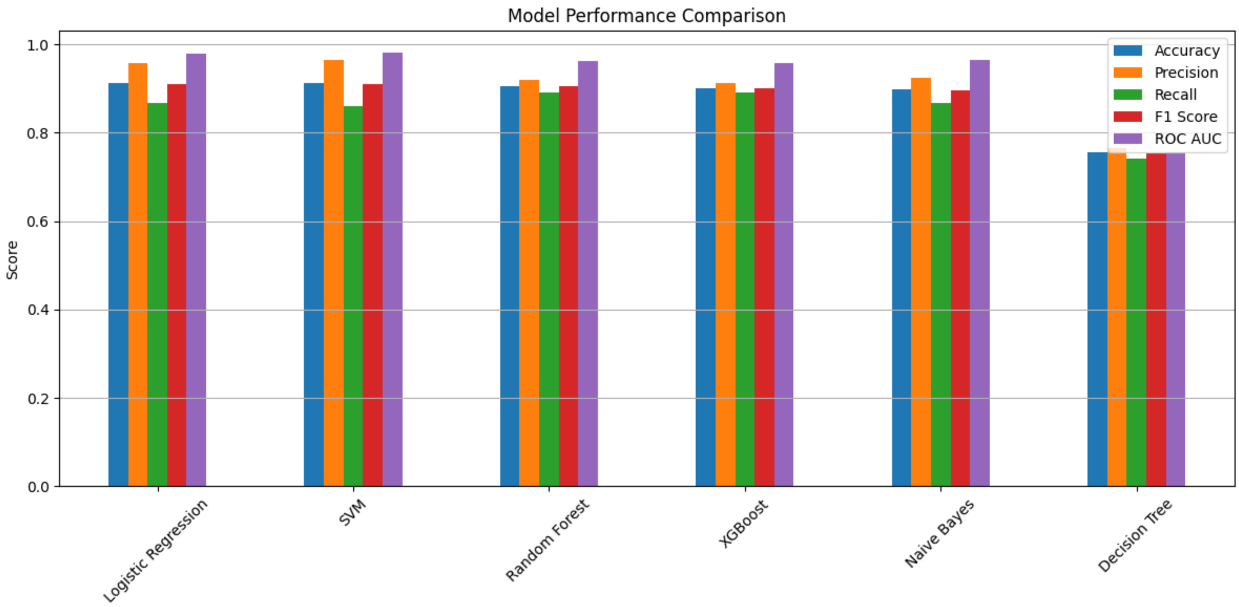

Figure 3: Model Performance Comparison (Bar Chart)

The figure demonstrates a graphical overview of five essential metrics including Accuracy alongside Precision and Recall and F1 Score and ROC AUC which track the performance of six different models. All performance metrics reflect a superior outcome for Logistic Regression and SVM while Decision Tree demonstrates significantly worse results according to the data.

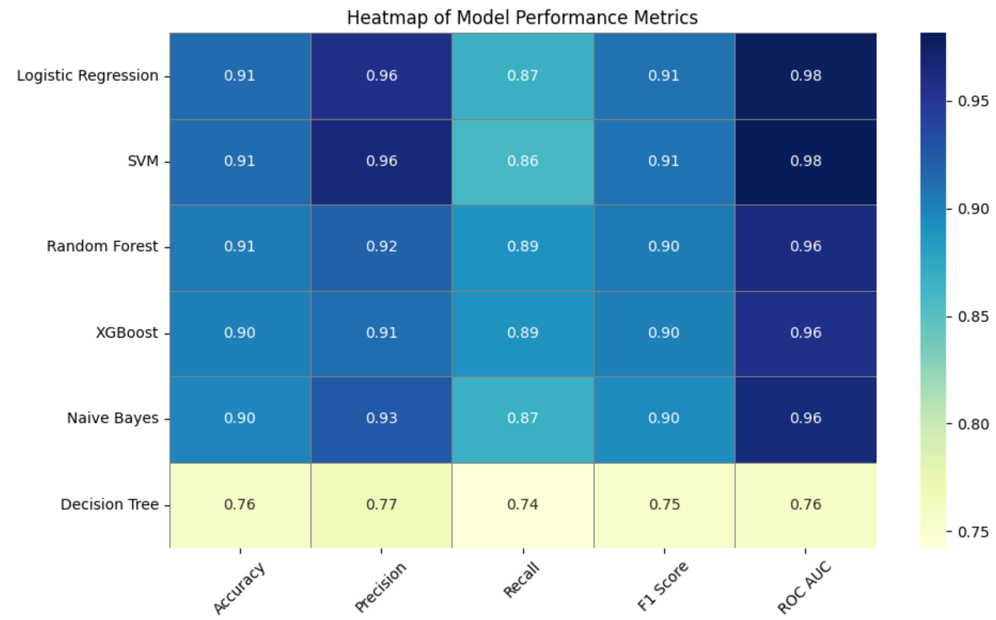

Figure 4: Heatmap of Model Performance Metrics

A color heat mapping approach enhances the visual perception from the bar chart by using darker shades to represent superior model metrics performance. Darker shades indicate better performance. The ROC AUC values and F1 Scores of SVM and Logistic Regression approach perfection at 0.98 because of their outstanding results.

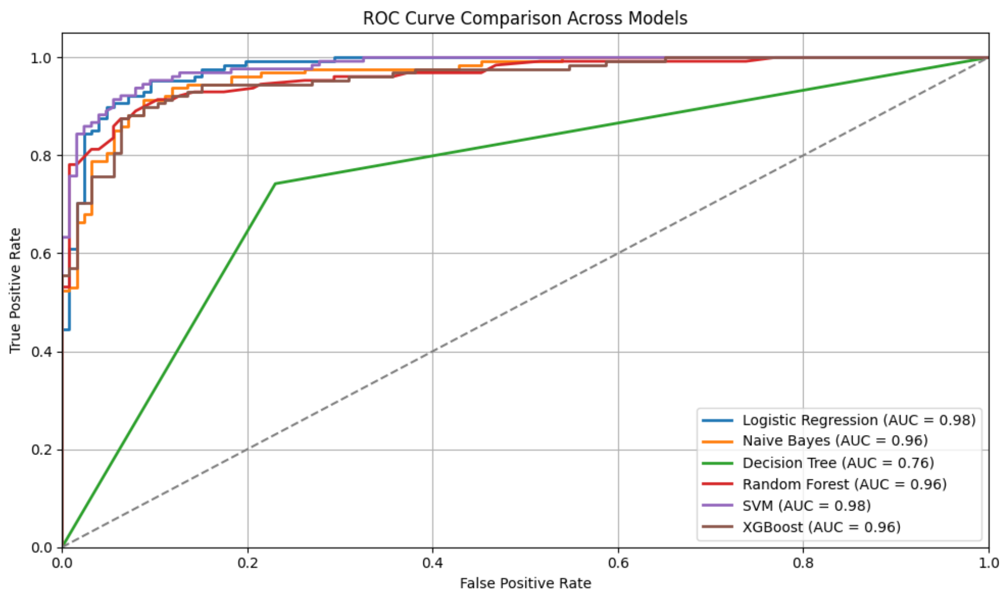

Figure 5: ROC Curve Comparison Across Models

The ROC curves in this visual representation allow viewers to evaluate how sensitive and specific each model performs at different classification threshold points. The regression curve patterns of Logistic Regression and SVM show excellent performance since they strongly occupy the top-left side of the visualization. AUC values demonstrate that these models deliver peak performance with 0.98 and 0.981 respectively.

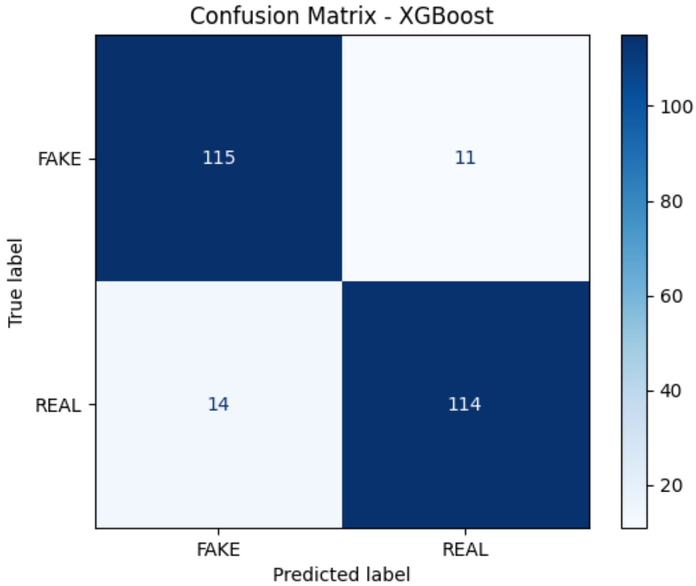

Figure 6: Confusion Matrix – XGBoost

XGBoost achieved a solid performance according to its confusion matrix which showed 115 FAKE and 114 REAL correct classifications out of 254 total examples.

The classification model shows balanced prediction precision because it misidentifies only 14 REAL cases while mistaking 11 FAKE cases as REAL. The model shows exceptional precision and recall in its previous tables because of this result.

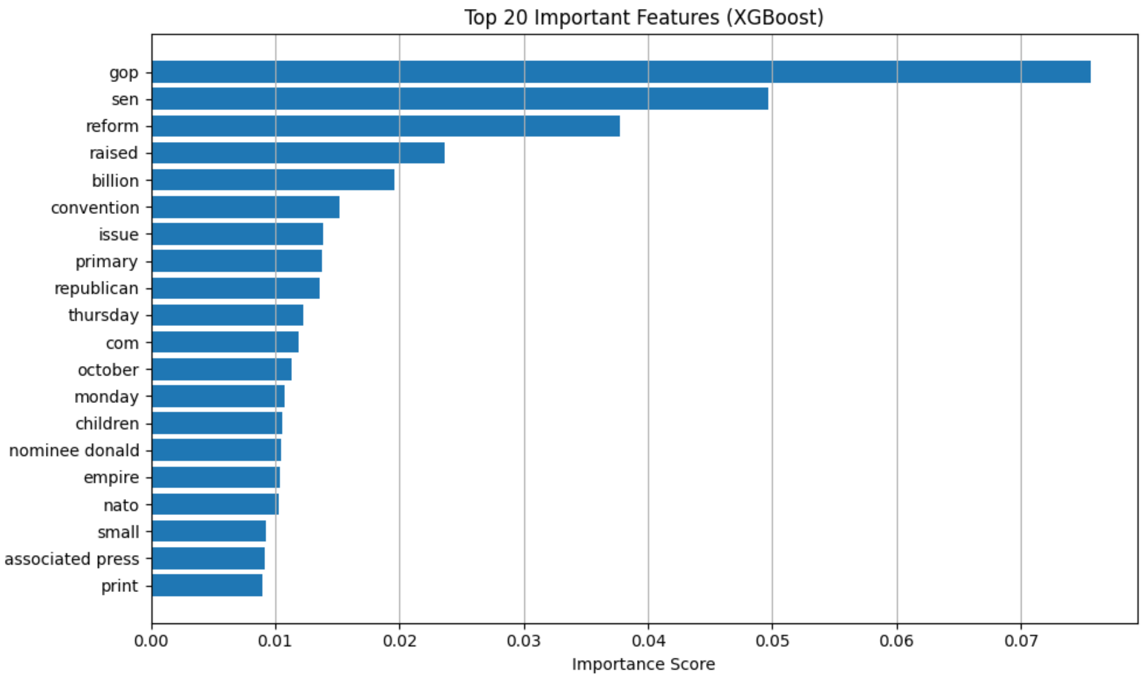

Figure 7: Top 20 Important Features (XGBoost)

The plot indicates which words XGBoost selected as critical factors during its classification process. The analysis showed that terms such as "gop" "sen" and "reform" contributed most to the determination process about fake or real news content.

RELATED WORK

Fake news detection has become a priority issue because social media platforms now function as main news platforms. Different researchers have utilized machine learning algorithms and natural language processing methods for this task.

Rubin et al. (2016) conducted one of the initial studies regarding news deception detection through analysis of linguistic indicators alongside rhetorical elements. The study established key approaches to determine stylistic features between genuine and fraudulent news content. The combination of n-gram and TF-IDF features with Logistic Regression for fake news article classification led to effective results according to Ahmed et al. (2017).

The LIAR dataset introduced by Wang (2017) contains numerous political statements for truth validation purposes which both CNNs and LSTMs analyze for sequential patterns. Deep learning models possessed good detection capability yet needed enormous datasets together with massive computational power to achieve results. The research by Rashkin et al. (2017) applied RNNs with stylistic features to study fake, satirical and true news resulting in the discovery that fake news demonstrates poor lexical diversity and lack of objectivity.

The latest research of Shu et al. (2020) provides an extensive overview of fake news detection approaches which stresses the significance of content-based as well as context-based features that include social engagement metrics with user credibility assessments. We concentrate on content-based methods through TF-IDF and classical machine learning algorithms because we want to determine the limitations of these lightweight approaches.

Standardized preprocessing and feature engineering combine with constant dataset analysis to serve as the core aspects of our project as opposed to previous works. The concept remains that text-based classification with interpretable models delivers highly accurate results while omitting deep learning and social network data.

CONCLUSION

In this work, we designed a machine learning-based system for the detection of spurious news articles based on natural language processing techniques. Having framed the problem as a binary classification task, we compared the performance of the baseline models such as Logistic Regression, Naive Bayes, SVM, Decision Tree, Random Forest, and XGBoost. Through preprocessing and TF-IDF-based feature extraction, raw news text was transformed into structured data suitable for classification.

We have seen that the best performing models are Logistic Regression and SVM, where both of them have achieved over 91% accuracy along with an ROC AUC of over 0.97. Our results demonstrate that fake news can be identified with high efficiency without deep learning, without social context, but with only classical methods combined with high quality engineered features.

There are some limitations to the project. While the dataset is balanced, it is small in the continually evolving context of modern-day standards and may not generalize to more diverse or multilingual sources of misinformation, either. Also importantly, our approach analyzes the content wholly in isolation, without also considering external signals, such as publisher credibility or users' engagement patterns.

Future work may include increased size and diversity of the dataset, adding context-based features (like the metadata for articles, or which articles are socially shared), and/or exploring deep learning methods like Transformers (and other helping models, such as BERT or others) to gain a better sense of semantics. Even better, real-time news feeds could allow for live testing of the proposed systems and move them closer to implementation in practice.

In general, this work shows that traditional machine learning is still a useful and scalable tool to address online misinformation, when accompanied by careful text processing.

REFERENCES

- Ahmed, H., Traore, I., & Saad, S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques. *Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, 127–138. https://doi.org/10.1007/978-3-319-69155-8_9
- Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2921–2927. https://doi.org/10.18653/v1/D17-1317
- Rubin, V. L., Conroy, N. J., Chen, Y., & Cornwell, S. (2016). Fake news or truth? Using satirical cues to detect potentially misleading news. *Proceedings of NAACL-HLT 2016*, 7–17. https://aclanthology.org/N16-2006/
- Shu, K., Wang, S., & Liu, H. (2020). A survey on fake news detection. *ACM SIGKDD Explorations Newsletter*, 19(1), 22–36. https://doi.org/10.1145/3137597.3137600
- Wang, W. Y. (2017). "Liar, Liar Pants on Fire": A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 422–426. https://doi.org/10.18653/v1/P17-2067
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785