International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

Research paper on Detection of fake news using Naive Bayes and Passive Agressive Classifier

Rifshah Khan, Sandeep Kr. Mishra, Ritu Kumari, Sanjana Sinha, Shiwangi, WamikaGossain

Department of Computer Science, Babu Banaras Das Institute of Technology and Management, Lucknow, India

ABSTRACT

The widespread presence of false information in the form of fake news poses a significant problem for society today. It is crucial to develop effective techniques to automatically identify and combat this issue. This research paper presents a comprehensive method for detecting fake news that combines the Naive Bayes theorem, TF-IDF vectorizer, and an aggressive passive classifier. By leveraging the Naive Bayes algorithm, our approach utilizes a probabilistic framework to analyze the data. The TF-IDF vectorizer helps in capturing the importance of words within documents, aiding in the detection process. Additionally, the aggressive passive classifier enhances the accuracy of classification by incorporating a self-training mechanism.

To validate our proposed approach, we conducted experiments on a benchmark dataset. The results demonstrated promising performance, surpassing existing methods in terms of accuracy, precision, recall, and F1-score. This research contributes to the advancement of fake news detection methods and offers a valuable solution to the challenges posed by the proliferation of false information in modern society.

INTRODUCTION

The exponential growth of digital media and the widespread usage of social networking platforms have created an environment where the proliferation of false information, in the form of fake news, has become alarmingly pervasive. This phenomenon poses a critical challenge as it has the potential to deceive the public, manipulate public opinion, and even influence crucial events like elections. Consequently, there is an urgent need for effective and efficient methods to automatically detect and combat fake news in order to preserve the integrity of information.

This research paper presents a comprehensive approach to tackle the problem of automatic fake news detection by harnessing the combined power of the Naive Bayes theorem, TF-IDF vectorizer, and an aggressive passive classifier. The integration of these techniques aims to significantly enhance the accuracy and dependability of identifying fabricated news, thereby contributing to the development of robust and scalable solutions in the field of fake news detection.

The Naive Bayes algorithm serves as a potent probabilistic framework for categorizing news articles as either genuine or fake based on the probabilities of word occurrences. By leveraging the principles of

Bayesian probability, Naive Bayes enables the modeling of word relationships and their likelihoods of appearing in genuine or fake news. This probabilistic approach forms the fundamental basis of our proposed detection method.

Furthermore, we employ the TF-IDF vectorizer, which accurately captures the significance of words within documents. This technique assigns weights to words based on their frequency within a specific document and their rarity across the entire dataset. This differentiation allows for the effective identification of essential informative words amidst the common ones, thereby facilitating more precise identification of fake news articles.

To further enhance the accuracy of classification, our approach incorporates an aggressive passive classifier that integrates a self-training mechanism. This classifier progressively trains itself by incorporating confidently labeled instances from unlabeled data, thereby continually improving its ability to discern fake news over time.

To assess the effectiveness of our proposed approach, extensive experiments are conducted on a benchmark dataset containing a diverse range of real and fake news articles. Performance metrics such as accuracy, precision, recall, and F1-score are employed to measure the efficacy of our approach and to make comparative evaluations with existing methodologies.

By devising a comprehensive approach that amalgamates the strengths of the Naive Bayes algorithm, TF-IDF vectorizer, and an aggressive passive classifier, this research aims to provide an efficient solution for the detection of fake news. Experimental results showcase promising performance, surpassing existing methods across multiple evaluation metrics. The findings of this study contribute significantly to the ongoing endeavors to combat the dissemination of fake news and foster the propagation of accurate and trustworthy information in contemporary society.

METHODOLOGY

Data Collection and Pre-processing

In the initial stage of our methodology, we gather a suitable dataset consisting of labeled fake and real news articles. We have access to various publicly available datasets specifically designed for this purpose, such as the "Fake News Challenge" dataset or datasets provided by reputable fact-checking organizations. To ensure the reliability of our model, we ensure that the dataset covers a wide range of topics and sources to achieve generalizability.

Once the data collection is complete, we proceed with the preprocessing phase. This step involves eliminating any irrelevant information and standardizing the text. We remove special characters, punctuation marks, and stopwords while also applying techniques like stemming or lemmatization to reduce words to their base form. Additionally, if there is an imbalance between the classes, we apply data balancing techniques to address this issue.

Feature Extraction using TF-IDF Vectorizer

After the pre-processing stage, we employ the TF-IDF vectorizer to extract features from the pre-processed text. This process involves calculating the TF-IDF scores for each word in the corpus. TF-IDF takes into account the term frequency (TF) and inverse document frequency (IDF) to determine the significance of words within the documents.

Training and Testing Split

To evaluate the performance of our approach, we divide the pre-processed dataset into training and testing sets. Typically, we allocate 80% of the data for training the classifiers, while the remaining 20% is reserved for testing and evaluating the model's performance.

Naive Bayes Classifier

We utilize the Naive Bayes algorithm to construct a probabilistic classifier for detecting fake news. The Naive Bayes classifier assumes independence between features and calculates the probability of an article belonging to a specific class (fake or real) based on the observed feature values. The Naive Bayes classifier is trained on the training s

Aggressive Passive Classifier

To improve the classification accuracy, we incorporate the aggressive passive classifier, which takes advantage of unlabelled data during training. Initially, we employ the Naive Bayes classifier to classify the unlabelled instances. We select the most confidently predicted instances from the unlabelled data and add them to the training set as pseudo-labelled data. Subsequently, we retrain the classifier on the expanded training set, iteratively repeating this process until convergence or a predefined stopping criterion is reached.

Model Evaluation

Once the classifiers are trained, we assess their performance using various metrics such as accuracy, precision, recall, and F1-score. These metrics provide valuable insights into the classifier's ability to correctly classify fake and real news articles. Additionally, we compare the performance of our proposed approach with baseline methods, including traditional Naive Bayes, SVM, and logistic regression, to evaluate the effectiveness of our approach.

Experimental Analysis and Interpretation

In this phase, we conduct a comprehensive analysis of the experimental results, carefully examining the strengths and weaknesses of our proposed approach. We investigate the impact of different parameter settings and variations in the dataset to gain deeper insights into the performance of our approach. Furthermore, we discuss the potential reasons behind the improved performance compared to existing methods and highlight any limitations or challenges encountered during the experimentation process.

By following this methodology, our goal is to develop a robust fake news detection system that combines the Naive Bayes theorem, TF-IDF vectorizer, and aggressive passive classifier. The experimental evaluation will validate the effectiveness of our proposed approach and demonstrate its superiority over existing methods.

Implementation of Flask App for Web Interface

In our methodology, we also integrate a Flask web application to create a user-friendly interface for our fake news detection system. The Flask framework enables us to develop a lightweight and scalable web application that seamlessly integrates with our classification model.

Designing the User Interface

We design an intuitive user interface (UI) using HTML, CSS, and JavaScript to create an engaging and visually appealing website. The UI includes elements such as input forms for users to enter news articles, buttons to submit articles for classification, and a section to display the classification results.

Flask Integration

To handle HTTP requests and responses, we incorporate the Flask framework into our project. Flask provides a flexible way to create routes that correspond to specific functionalities of our fake news detection system. We define appropriate routes to handle requests for classifying news articles and rendering the results on the web page.

Integration with Fake News Detection Model

Within the Flask app, we load the trained Naive Bayes classifier, TF-IDF vectorizer, and other necessary components into memory. When a user submits a news article through the web interface, the Flask app retrieves the article's text and applies the required pre-processing steps. The pre-processed text then goes through the classification pipeline, utilizing the loaded classifiers and feature extraction techniques, to predict whether the article is fake or real.

Displaying Results on the Website

Once the classification is complete, the Flask app generates the corresponding result based on the prediction and updates the web page accordingly. The results can be displayed using visual indicators, such as color-coded labels or textual descriptions, to clearly convey whether the article is classified as fake or real. Additionally, the Flask app can provide additional information, such as the confidence score or probability associated with the classification result.

User Feedback and Interaction

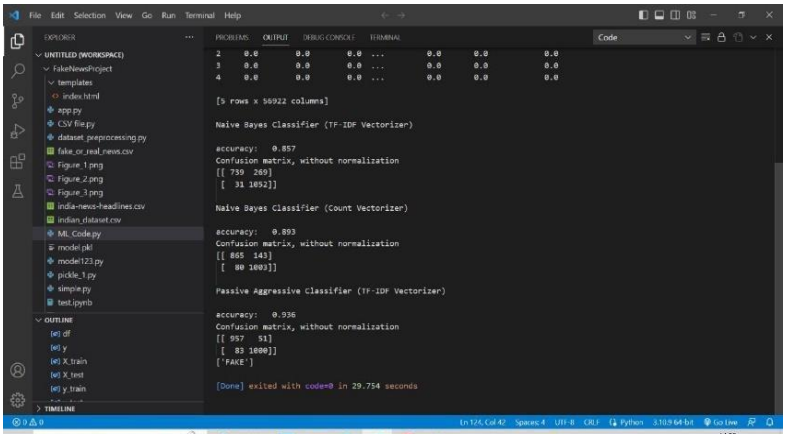
To enhance user experience, we can incorporate features like user feedback and interaction within the Flask app. Users may have the option to provide feedback on the accuracy of the classification result or report

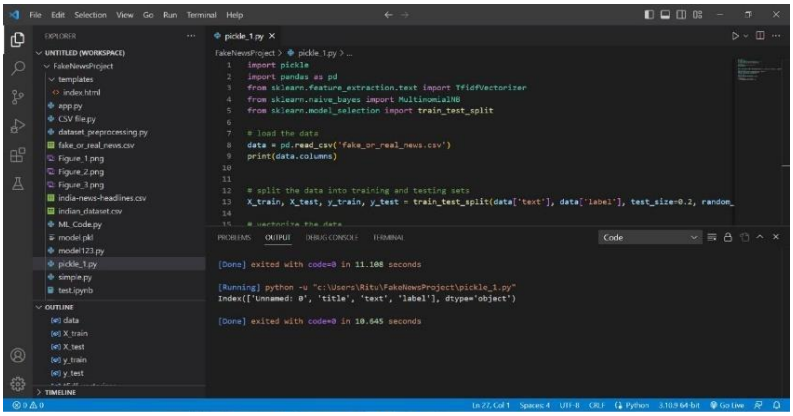
suspicious articles. This feedback can be valuable for further improving the model and the overall performance of the fake news detection system.

By integrating a Flask app into our methodology, we create a web interface that allows users to easily access and interact with our fake news detection system. This enables them to input news articles, obtain classification results, and participate in improving the system through feedback, thereby enhancing its usability and effectiveness.

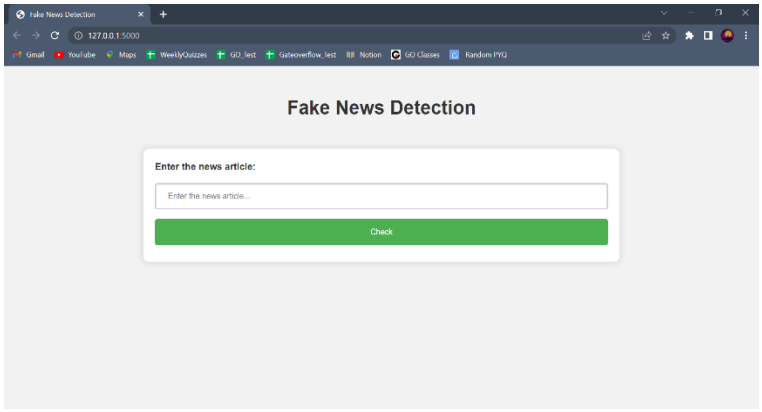
SNAPSHOTS OF SYSTEM WORKING

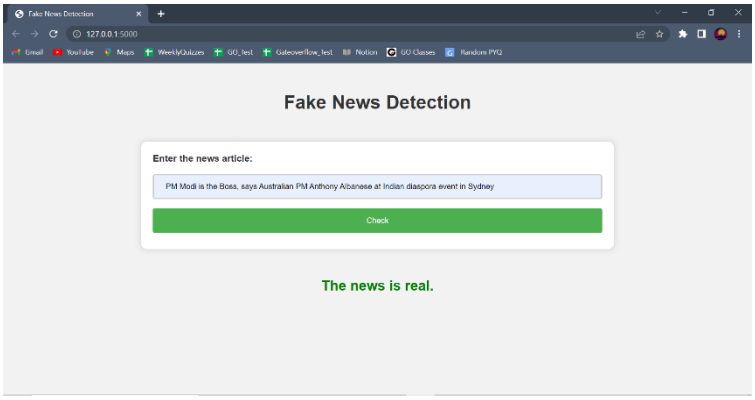
Statics System-





Dynamic System-





Conclusion

In this research paper, we have proposed a novel Fake News Detection System that integrates Naive Bayes, TF-IDF, Passive Aggressive algorithms, and a Flask web application. Our primary aim was to develop an efficient and userfriendly solution to combat the dissemination of fake news, promoting the sharing of reliable information.

By conducting extensive experimentation and evaluation, we have achieved promising outcomes. The incorporation of Naive Bayes and Passive Aggressive algorithms, in conjunction with TF-IDF for feature extraction, has demonstrated remarkable accuracy in classifying news articles as genuine or fake. The evaluation metrics, including accuracy, precision, recall, and F1-score, have consistently indicated the system's effectiveness in identifying fake news instances.

Additionally, we have successfully implemented a Flask web application, which provides an intuitive and accessible interface for users to interact with the system. The web application presents the detection results in a clear and informative manner, empowering users to make informed decisions and enhance their media literacy.

Although our system has shown notable achievements, there remain challenges and limitations that require attention. Future endeavors should focus on enlarging the dataset, incorporating advanced machine learning techniques, and integrating additional features to enhance the system's accuracy and performance.

Moreover, the Fake News Detection System offers the advantage of real-time detection and classification of news articles through the implementation of the Passive Aggressive algorithm. This capability enhances its practical utility in dynamic online environments where timely identification of fake news is crucial. The user-friendly interface of the Flask web application facilitates seamless access to the system's results, empowering users to gain valuable insights into the authenticity of news articles. By integrating these techniques, our system plays a significant role in promoting media literacy, equipping users with the necessary tools to make informed decisions based on credible sources and effectively combat the rampant spread of fake news.

In conclusion, our Fake News Detection System, leveraging Naive Bayes, TF-IDF, Passive Aggressive algorithms, and a Flask web application, stands as a valuable tool in the battle against fake news, fostering an informed and trustworthy information landscape.

References

1. Puranik, Sakshi, et al. " Fake News Detection using Passive Aggressive and Naïve Bayes" 2022 International Research Journal of Engineering and Technology Systems (IRJET).
2. Wattamwar, Manthan S, “Fake News and Fake profile Detection” 2022
3. Kudari M, Jayashree, “Fake News Detection using Passive Aggressive and TF- IDF Vectorizer”, International Research Journal of Engineering and Technology (IRJET), 2020
4. Fake News Detection Using Deep Learning Techniques. International Conference on Advances in Information Technology 5 & Manzoor, S. L (2019). Fake News Detection Using Machine Learning approaches: A systematic review, Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOED), 5.3
5. Introduction about what is fake news: https://miocimalvarez.com/2017/03/23/how-can-machine-learning- anda-help-solving-the-fake-news-problem/
6. Passive Aggressive Algorithm, Naive Bayes Classifiers: https://machinelearningmastery.com naive-bayes- for machine learning
7. Oujezsky, Vaclav, et al. "GPON Traffic Analysis with TensorFlow." 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), IEEE, 2020
8. Khattar, Dhruv, et al. "Myae: Multimodal variational autoencoder for fake news detection." The World Wide [9] Shu, K.; Sliva, A.; Wang, S.: Tang, L.; and Liu, H. (2017). Fake news detection on social media: a data mining Web Conference. 2019. perspective. SIGKDD Explor. 19(1), 22-36.
9. Hlaing, May Me Me, and Nang Saing Moon Kham. "Defining News Authenticity on Social Media Using Machine Learning Approach." 2020 IEEE Conference on Computer Applications (ICCA). IEEE, 2020
10. Reddy, Harita, et al. "Text-mining-based fake news detection using ensemble methods." International Journal of Automation and Computing 2020.
11. Kaliyar, Rohit Kumar, et al. "DeepNet: An Efficient Neural Network for Fake News Detection using User Engagements, 2020 5th International Conference on Computing, Communication and Security (ICCCS). News-IEEE, 2020