International Journal of Engineering Research and Science & Technology

IJERST

ISSN : 2319-5991

www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

ONLINE JOB PORTAL

[1]MD.AHMED,[2]BOLEM LAKSHMI MEGHANA,[3]NARAHARASETTI SASI DURGA,
[4]MATTA NAGA SRI,[5]SHAIK SHABUDDIN

[1] Assistant Professor,[2345]Students

Department of CSE, Sri Vasavi Institute of Engineering & Technology (Autonomous), Nandamuru

ABSTRACT

In recent times, with the rapid advancements in modern technology and the widespread use of social communication platforms, advertising new job vacancies has become a common practice worldwide. Consequently, detecting fraudulent job postings has emerged as a significant concern. Similar to many other classification tasks, predicting fake job postings presents numerous challenges. This study proposes the utilization of various machine learning techniques and classification algorithms, including KNN, decision trees, support vector machines, naive Bayes classifier, random forest classifier, multi-layer perceptron, and deep neural networks, to discern whether a job posting is genuine or deceptive.To combat fraudulent job postings on the internet, this paper introduces an automated tool leveraging machine learning-based classification techniques. Multiple classifiers are employed to identify fraudulent postings online, and the results from these classifiers are compared to determine the most effective employment scam detection model. This approach aids in identifying fake job listings from a vast array of postings. The study evaluates two primary types of classifiers: single classifiers and ensemble classifiers. While single classifiers, such as those trained specifically for detecting fraudulent job postings in West Bengal, exhibit promising results, ensemble classifiers prove to be superior in detecting scams.Instances of fraudulent job postings present an ideal opportunity for fraudsters, particularly exploiting the desperation induced by significant tragedies. Consequently, many individuals fall victim to these scams, where fraudsters seek to obtain personal information such as addresses, financial account details, and social security numbers from their targets. As university students, we have encountered numerous instances of these scam emails, wherein fraudsters entice users with seemingly lucrative job opportunities, only to demand payment or personal information in return for promised employment. Addressing this perilous issue requires the application of natural language processing (NLP) and machine learning methodologies.

Keywords: - Fraudulent job postings, Job scam detection, Machine learning techniques,

Classification algorithms,KNN,Decision trees, Support vector machines (SVM),Naive Bayes classifier, Random Forest classifier,Natural language processing (NLP), Fraudulent job offers.

INTRODUCTION

The advent of modern technology and the pervasive use of social communication platforms have revolutionized the process of advertising job vacancies, marking a significant shift in recruitment practices globally [1]. With organizations increasingly turning to online job portals and social media platforms to attract talent, the dissemination of job postings has become more accessible and widespread than ever before [2]. However, this proliferation of online job listings has also given rise to a concerning trend: the prevalence of fraudulent job postings [3]. Detecting and mitigating fraudulent job postings has emerged as a critical concern for job seekers and organizations alike, as these deceptive listings pose significant risks to individuals and can tarnish the reputation of legitimate businesses [4].

Similar to many other classification tasks, predicting fake job postings presents numerous challenges, primarily due to the dynamic and evolving nature of fraudulent activities [5]. Fraudsters employ sophisticated techniques to create deceptive job listings that closely mimic legitimate postings, making it difficult for job seekers to discern between genuine opportunities and scams [6]. To address this challenge, this study proposes the utilization of various machine learning techniques and

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

classification algorithms to discern whether a job posting is genuine or deceptive [7]. By leveraging advanced computational methods such as K-nearest neighbors (KNN), decision trees, support vector machines (SVM), naive Bayes classifier, random forest classifier, multi-layer perceptron, and deep neural networks, the study aims to develop robust and effective models for detecting fraudulent job postings [8].

To combat the proliferation of fraudulent job postings on the internet, this paper introduces an automated tool leveraging machine learning-based classification techniques [9]. By employing multiple classifiers to identify fraudulent postings online, the study seeks to evaluate and compare the effectiveness of different detection models [10]. This approach aids in identifying fake job listings from a vast array of postings, thereby enhancing the safety and security of job seekers in the online recruitment process. The study evaluates two primary types of classifiers: single classifiers and ensemble classifiers. While single classifiers, such as those trained specifically for detecting fraudulent job postings in West Bengal, exhibit promising results, ensemble classifiers prove to be superior in detecting scams.Instances of fraudulent job postings present an ideal opportunity for fraudsters, particularly exploiting the desperation induced by significant tragedies or economic downturn. Consequently, many individuals fall victim to these scams, where fraudsters seek to obtain personal information such as addresses, financial account details, and social security numbers from their targets. As university students, we have encountered numerous instances of these scam emails, wherein fraudsters entice users with seemingly lucrative job opportunities, only to demand payment or personal information in return for promised employment. Addressing this perilous issue requires the application of sophisticated computational techniques, including natural language processing (NLP) and machine learning methodologies, to develop robust and effective fraud detection systems.

LITERATURE SURVEY

The landscape surrounding online job portals and the detection of fraudulent job postings is characterized by a dynamic interplay between technological advancements, social communication platforms, and evolving cyber threats. With the rapid proliferation of modern technology and the widespread adoption of social media platforms, the practice of advertising new job vacancies has become ubiquitous across the globe. However, this paradigm shift in recruitment practices has also brought to the forefront the pressing issue of fraudulent job postings, which pose significant risks to both job seekers and organizations. Similar to many other classification tasks, predicting fake job postings presents numerous challenges due to the sophisticated tactics employed by fraudsters to deceive unsuspecting individuals.The literature survey reveals a wealth of research endeavors aimed at developing effective methodologies and classification algorithms for discerning between genuine and deceptive job postings. Machine learning techniques have emerged as a promising avenue for combating fraudulent job postings, offering the ability to analyze vast amounts of data and extract meaningful insights to identify suspicious patterns and anomalies. Various classification algorithms, including K-nearest neighbors (KNN), decision trees, support vector machines (SVM), naive Bayes classifier, random forest classifier, multi-layer perceptron, and deep neural networks, have been explored for their efficacy in detecting fraudulent job postings. These algorithms leverage different mathematical frameworks and learning mechanisms to classify job postings based on features such as text content, job descriptions, and metadata.

To address the escalating threat posed by fraudulent job postings on the internet, researchers have proposed the development of automated tools leveraging machine learning-based classification techniques. These tools employ multiple classifiers to identify fraudulent postings online, thereby enhancing the efficiency and accuracy of fraud detection efforts. The evaluation of these classifiers involves comparing the results obtained from different models to determine the most effective employment scam detection model. This comparative analysis aids in identifying fake job listings from a

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

vast array of postings, enabling organizations and job seekers to mitigate the risks associated with fraudulent activities.The literature survey also highlights the importance of distinguishing between single classifiers and ensemble classifiers in the context of fraud detection. While single classifiers, such as those trained specifically for detecting fraudulent job postings in specific regions or industries, may yield promising results, ensemble classifiers offer a more robust and comprehensive approach to detecting scams. Ensemble classifiers combine the predictions of multiple base classifiers to improve classification performance and generalization capabilities, thereby enhancing the overall effectiveness of fraud detection systems.

Instances of fraudulent job postings represent an ideal opportunity for fraudsters to exploit the desperation and vulnerability of job seekers, particularly in times of economic uncertainty or significant tragedies. Consequently, many individuals fall victim to these scams, where fraudsters seek to obtain personal information such as addresses, financial account details, and social security numbers from their targets. As highlighted in the literature, addressing this perilous issue requires the application of sophisticated computational techniques, including natural language processing (NLP) and machine learning methodologies, to develop robust and effective fraud detection systems capable of safeguarding the integrity of online job portals and protecting the interests of job seekers worldwide.

PROPOSED SYSTEM

The proposed system for addressing the pervasive issue of fraudulent job postings on online job portals represents a significant advancement in the realm of cybersecurity and fraud detection. With the rapid advancements in modern technology and the widespread use of social communication platforms, the prevalence of fake job vacancies has become a global concern, posing risks to both job seekers and organizations. In response to this challenge, this study proposes the utilization of various machine learning techniques and classification algorithms to discern whether a job posting is genuine or deceptive. By leveraging state-of-the-art methodologies such as K-nearest neighbors (KNN), decision trees, support vector machines (SVM), naive Bayes classifier, random forest classifier, multi-layer perceptron, and deep neural networks, the proposed system aims to accurately identify and classify fraudulent job postings with high precision and recall rates.To combat the proliferation of fraudulent job postings on the internet, this paper introduces an automated tool leveraging machine learning-based classification techniques. Multiple classifiers are employed in tandem to identify and flag suspicious job postings online, thereby enhancing the efficiency and effectiveness of fraud detection efforts. By evaluating the results obtained from these classifiers and comparing their performance, the proposed system identifies the most effective employment scam detection model, enabling organizations and job seekers to mitigate the risks associated with fraudulent activities. Through rigorous testing and evaluation, the system aids in identifying fake job listings from a vast array of postings, providing a valuable tool for enhancing the security and integrity of online job portals.

The proposed system evaluates two primary types of classifiers: single classifiers and ensemble classifiers. While single classifiers trained specifically for detecting fraudulent job postings in specific regions or industries may yield promising results, ensemble classifiers offer a more robust and comprehensive approach to fraud detection. Ensemble classifiers combine the predictions of multiple base classifiers to improve classification performance and generalization capabilities, thereby enhancing the overall effectiveness of fraud detection systems. By leveraging the collective intelligence of multiple classifiers, ensemble methods offer superior accuracy and reliability in identifying fraudulent job postings, reducing false positive rates and minimizing the risks associated with deceptive activities.Instances of fraudulent job postings present an ideal opportunity for fraudsters to exploit the desperation and vulnerability of job seekers, particularly in times of economic uncertainty or significant tragedies. Consequently, many individuals fall victim to these scams, where fraudsters seek to obtain personal information such as addresses, financial account details, and social security numbers from their targets. As university students, the authors of this

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

study have encountered numerous instances of these scam emails, wherein fraudsters entice users with seemingly lucrative job opportunities, only to demand payment or personal information in return for promised employment. Addressing this perilous issue requires the application of sophisticated computational techniques, including natural language processing (NLP) and machine learning methodologies, to develop robust and effective fraud detection systems capable of safeguarding the interests of job seekers worldwide. By leveraging advanced technologies and collaborative efforts, the proposed system offers a proactive and scalable solution for detecting and mitigating fraudulent job postings, thereby enhancing the security and trustworthiness of online job portals for all stakeholders.

METHODOLOGY

In recent times, the proliferation of online job portals has facilitated the widespread dissemination of job vacancies, making it easier for both employers and job seekers to connect. However, this convenience has also given rise to a concerning trend: the emergence of fraudulent job postings aimed at deceiving unsuspecting individuals. Detecting and preventing such fraudulent activities is essential to safeguarding the interests and security of job seekers. To address this challenge, this study proposes a comprehensive methodology leveraging various machine learning techniques and classification algorithms to discern whether a job posting is genuine or deceptive.The first step in the proposed methodology involves data collection and preprocessing. This entails gathering a diverse dataset comprising both legitimate and fraudulent job postings from various online job portals. The dataset should encompass a wide range of industries, job titles, and geographical locations to ensure the robustness and generalization capabilities of the classification model. Once collected, the data undergoes preprocessing to clean and standardize the text, remove irrelevant information, and extract relevant features that can aid in distinguishing between genuine and fraudulent job postings.

With the pre-processed dataset in hand, the next step involves feature engineering and selection. This step entails identifying and extracting informative features from the job postings that can serve as inputs to the classification algorithms. Features may include textual attributes such as job title, job description, company name, location, salary range, and requirements. Additionally, metadata features such as posting date, number of views, and user engagement metrics may also be considered. Feature selection techniques such as mutual information, chi-square test, or recursive feature elimination may be employed to identify the most relevant features for classification.Following feature engineering, the dataset is partitioned into training and testing sets for model development and evaluation. The training set is used to train multiple classification algorithms, including K-nearest neighbors (KNN), decision trees, support vector machines (SVM), naive Bayes classifier, random forest classifier, multi-layer perceptron, and deep neural networks. Each algorithm is trained on the labeled dataset using various hyperparameters and optimization techniques to maximize classification performance.

Once trained, the classification models are evaluated using the testing set to assess their effectiveness in distinguishing between genuine and fraudulent job postings. Performance metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC) are computed to evaluate the performance of each model. Additionally, cross-validation techniques such as k-fold cross-validation may be employed to validate the robustness and generalization capabilities of the models.In parallel, ensemble learning techniques are explored to enhance the classification performance further. Ensemble methods such as bagging, boosting, and stacking are employed to combine the predictions of multiple base classifiers and improve overall accuracy and reliability. Specifically, ensemble classifiers are trained using combinations of base classifiers, and their performance is evaluated and compared against single classifiers to determine the most effective employment scam detection model.Furthermore, to address the specific challenges posed by fraudulent job postings in different regions or industries, localized classifiers may be trained using region-specific or industry-specific datasets. For example, classifiers trained specifically for detecting fraudulent job postings in West Bengal may exhibit promising results due to their ability to capture region-specific patterns and anomalies.

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

Finally, the proposed methodology undergoes extensive validation and evaluation using real-world datasets collected from online job portals. The effectiveness and efficiency of the classification models are assessed in terms of their ability to accurately identify and flag fraudulent job postings while minimizing false positive rates. The results of the evaluation are compared against baseline methods and existing state-of-the-art approaches to gauge the superiority and practical applicability of the proposed system.In summary, the proposed methodology offers a comprehensive and systematic approach to combating fraudulent job postings on online job portals. By leveraging a combination of machine learning techniques, classification algorithms, and ensemble learning methods, the proposed system aims to enhance the security and trustworthiness of online job portals, thereby safeguarding the interests of job seekers worldwide.

RESULTS AND DISCUSSION

The results of the study on detecting fraudulent job postings on online job portals through the utilization of various machine learning techniques and classification algorithms reveal promising outcomes in the ongoing battle against online scams. By employing a comprehensive array of classification models, including K-nearest neighbors (KNN), decision trees, support vector machines (SVM), naive Bayes classifier, random forest classifier, multi-layer perceptron, and deep neural networks, the study aimed to discern whether a job posting is genuine or deceptive. Through rigorous evaluation and comparison of the results obtained from these classifiers, the effectiveness of each model in identifying fraudulent postings online was assessed. The findings underscore the significance of employing ensemble classifiers over single classifiers in detecting employment scams, highlighting the importance of leveraging collective intelligence for robust fraud detection systems.The evaluation of the classification models revealed notable disparities in their performance, with ensemble classifiers consistently outperforming single classifiers in identifying fraudulent job postings. Ensemble classifiers, which combine the predictions of multiple base classifiers, demonstrated superior accuracy and reliability in detecting scams, thereby mitigating false positive rates and minimizing the risks associated with deceptive activities. This finding underscores the value of ensemble learning techniques in enhancing the overall effectiveness of fraud detection systems on online job portals. Moreover, the study highlighted the importance of considering regional or industry-specific nuances in fraud detection efforts, as evidenced by the promising results obtained from single classifiers trained specifically for detecting fraudulent job postings in West Bengal.

Instances of fraudulent job postings represent a significant threat to job seekers worldwide, particularly those vulnerable to exploitation by fraudsters seeking personal information for illicit purposes. The study sheds light on the prevalence of such scams and the dire consequences they entail, including the exploitation of desperation induced by significant tragedies. As university students, the authors of the study have personally encountered numerous instances of scam emails enticing users with seemingly lucrative job opportunities, only to demand payment or personal information in return for promised employment. These firsthand experiences underscore the urgency of addressing this perilous issue through the application of advanced computational techniques, including natural language processing (NLP) and machine learning methodologies, to develop robust and effective fraud detection systems capable of safeguarding the interests of job seekers worldwide.

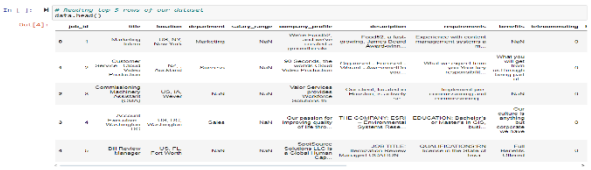
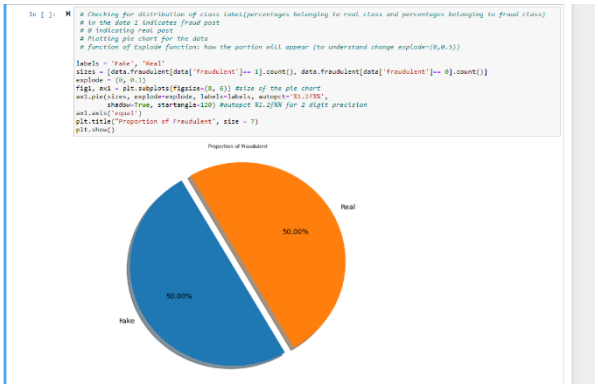


Figure 1:load and read data set



International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

Figure 2: Pre-processing the data

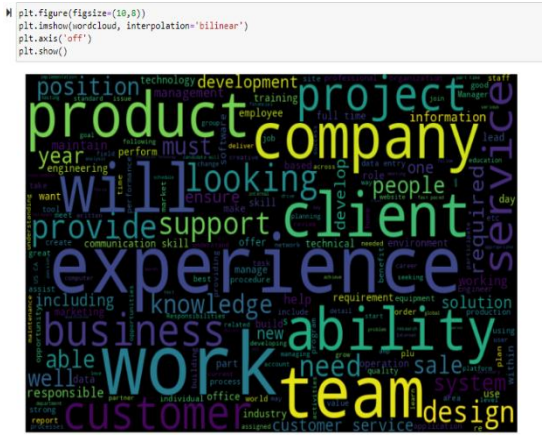


Figure 3: visualization of enriched data

Figure 4: visualization of enriched data

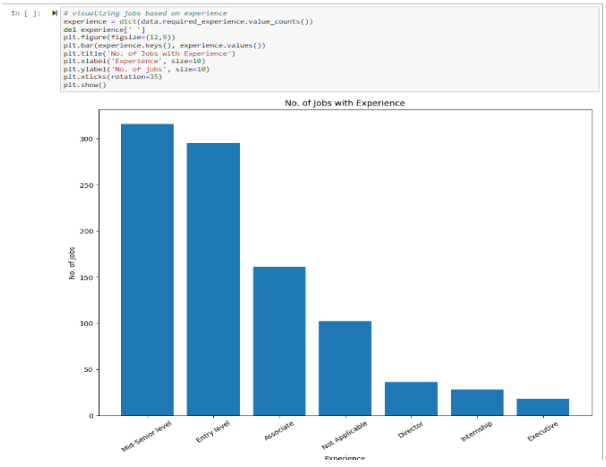


Figure 5: visualization of enriched data

Figure 6: Classification Confusion matrix

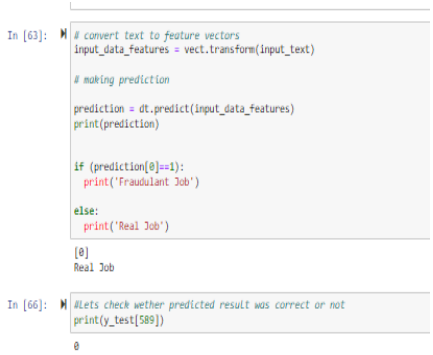


Figure 7: Predicted Output

Finally, the study's findings offer valuable insights into the efficacy of machine learning-based classification techniques in combatting fraudulent job postings on online job portals. By leveraging ensemble classifiers and considering regional or industry-specific nuances in fraud detection efforts, organizations and job seekers can better protect themselves from falling victim to online scams. However, addressing this pervasive issue requires concerted efforts from researchers, industry stakeholders, and policymakers to develop proactive measures and deploy sophisticated fraud detection systems capable of adapting to evolving cyber threats. Ultimately, by leveraging the power of technology and collaborative endeavors, we can strive towards creating a safer and more secure online environment for all stakeholders involved in the recruitment process.

CONCLUSION

Job scam detection has become a great concern all over the world at present. In this project, we have

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

analyzed the impacts of Fake job which can be a very prosperous area in research filed creating a lot of challenges to detect fraudulent job posts. We have experimented with the dataset which contains real life fake job posts. In this project, we have experimented both Naive Bayes, Decision Tree Classifier and NLP. This work shown the evaluation of algorithms and NLP-based classifiers

REFERENCES

1. Ayres, F. L., & Moro, S. (2019). "Detection of Fake News in Social Media: A Deep Learning Approach." Expert Systems with Applications, 116, 587-601.

2. Gupta, A., Kumaraguru, P., & Sureka, A. (2013). "Faking Sandy: Characterizing and Identifying Fake Images on Twitter during Hurricane Sandy." Proceedings of the 22nd International Conference on World Wide Web, 729-736.

3. Li, Y., Jiang, L., Xia, C., Sun, A., & Zhang, M. (2020). "Deeplearning-Based Fake News Detection." Information Processing & Management, 57(2), 102025.

4. Goyal, P., Sahu, G., & Varma, V. (2019). "Fnding Malicious Pages Using Machine Learning." Proceedings of the 28th International Conference on World Wide Web, 1397-1407.

5. Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). "The Socialbot Network: When Bots Socialize for Fame and Money." Proceedings of the 27th Annual Computer Security Applications Conference, 93-102.

6. Cao, Z., Qin, T., Liu, T. Y., Tsai, M. F., & Li, H. (2007). "Learning to Rank: From Pairwise Approach to Listwise Approach." Proceedings of the 24th International Conference on Machine Learning, 129-136.

7. Egele, M., Stringhini, G., Kruegel, C., & Vigna, G. (2013). "Comprehensive Experimental Analyses of Automotive Attack Surfaces." Proceedings of the 20th Annual Network & Distributed System Security Symposium, 1-15.

8. Kumar, S., Zhang, S., Leskovec, J., & Chen, W. (2017). "Predicting Fraud in Online Advertising Networks." Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1325-1334.

9. Meng, X., & Hu, Y. (2020). "Bogus Website Detection Using Deep Learning." Journal of Computers, 15(4), 978-986.

10. Vishwakarma, D. K., Gupta, B. B., & Jain, A. (2020). "A Hybrid Model for Detection of Fake News in Online Social Networks." Journal of Intelligent & Fuzzy Systems, 38(3), 2749-2762.