iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.59877

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Fake Job Detection System

Prof. Sanjivni Kale[1], Advait Manohar[2], Pranav Vitankar[3], Sahil Hadke[4]
[1]*Assistant Professor at P R Pote College of Engineering and Management Amravati*
[2, 3, 4]*Student at P R Pote College of Engineering and Management, Amravati*

Abstract: ***The technology has been updated to one level up, and the idea of hiring the employees the business companies, through online procedure is carried out. This makes the companies to get the employees of required post more immediate and in a faster way. It will be cost-effective as well. By exploring the internet, one can get the job easily of their qualifications and the field they wish to work in it. The posted jobs may be fake or legitimate, which are unaware by the people. To get rid of these kind of problems we come up with a new software which is designed to predict the job posts, as a result producing whether it is fake or legit one. We are designing a system as Fake job Post prediction using the concept of machine learning, in that we are using Random Forest classifier that produces accurate results in an efficient manner. The designed algorithm achieves the result of 98% as compared to the previously used algorithms. The students or users who search for a job may find difficulties in identifying the job posts that are fake and apply for the jobs, entering all the personal information without knowing about it. In some case they may get into the scams like paying money in the form of application fees in the need of job or the assurance of getting job after paying the money. The framework helps us to detect the posted jobs are fake not.***

Keywords: Fraud Classification, Fraud Detection Techniques, Machine learning, Decision tree, Random forest Logistic regression, Fraud detection and prediction.

I. INTRODUCTION

In recent years, the proliferation of online job portals has revolutionized the job search process, providing individuals with unprecedented access to a plethora of employment opportunities. However, amidst this convenience lies a growing concern: the prevalence of fake job postings. These deceptive listings not only waste job seekers' time and effort but also pose potential risks such as identity theft and financial scams. As such, the need for robust mechanisms to identify and combat fake job posts has become increasingly urgent. This project aims to address this pressing issue by leveraging the power of machine learning algorithms. By harnessing the vast amounts of data available on job portals, social media platforms, and other online sources, machine learning models can be trained to distinguish between legitimate job postings and fraudulent ones. Through the analysis of various textual, visual, and contextual features, these models can learn to detect patterns indicative of deceptive practices, thereby enabling the automated identification of fake job posts with high accuracy. The significance of this research extends beyond the realm of job seekers and recruiters. Businesses operating online job platforms stand to benefit significantly from the implementation of effective fake job post detection systems. By enhancing the integrity and credibility of their platforms, they can foster greater trust among users and safeguard their reputation in the competitive job market landscape.

II. EASE OF USE

1) *Job Description Analysis:* Look for red flags in the job description such as vague job responsibilities, unrealistic salary offers, spelling and grammatical errors, and overly generic language. Fake job postings often lack specific details about the job role and company.
2) *Research the Company:* Investigate the company advertising the job. Check their website, social media presence, and online reviews to ensure they are legitimate. Fake job postings may lead to nonexistent or fraudulent companies.
3) *Contact Information Verification:* Legitimate job postings should provide clear and accurate contact information for the company. Verify the email address, phone number, and physical address provided in the job posting.
4) *Check for Consistency:* Ensure that the job requirements, qualifications, and responsibilities are consistent throughout the posting. Discrepancies or inconsistencies may indicate a fake job.
5) *Use Online Tools and Databases:* There are several online tools and databases available for verifying job postings. These tools can help you identify common scam patterns and cross-reference job details with known legitimate listings.
6) *Trust Your Instincts:* If something feels off or too good to be true, it probably is. Trust your instincts and proceed with caution.
7) *Report Suspected Fraud:* If you encounter a suspicious job posting, report it to the platform or job board where you found it. This helps prevent others from falling victim to the scam.

©IJRASET: All Rights are Reserved

III. RELATED WORK

Table I
Previous Related Research Works

Ref	Title	Author	Result
[1]	Performance comparison of two class boosted decision tree and two class decision forest algorithms in predicting fake job postings	Shibly, F., Sharma, U., & Naleer, H	The two-class boosted decision tree algorithm was more effective than the two-class decision forest algorithm in detecting fake job postings.
[2]	Fake job recruitment detection using a machine learning approach. International Journal of Engineering Trends and Technology	Dutta, S., & Bandyopadhyay, S. K.	The paper suggests using machine learning, with the Random Forest classifier, to detect employment scams. It achieved an accuracy of 98.27%, outperforming existing methods
[3]	Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence	Cook J, Lewandowsky S, Ecker	Inoculation neutralizes misinformation's impact on consensus perception and moderates polarization on climate change beliefs and policy support. Inoculations, focusing on general misinformation techniques, effectively complement consensus messaging in combating misinformation.
[4]	Fake job detection and analysis using machine learning and deep learning algorithms	Anita, C., Nagarajan, P., Sairam, G. A., Ganesh, P., & Deepakkumar	Machine learning and deep learning algorithms to distinguish between fake and real job postings. It finds that Bi-Directional LSTM yields the most accurate results compared to other classification algorithms.

IV. LITERATUREREVIEW

Detecting fake job postings has become increasingly important due to the rise of online job platforms and the prevalence of job scams. Researchers have explored various approaches to address this issue, employing techniques from natural language processing, machine learning, and data mining. This literature review provides an overview of key studies in the field of fake job detection.

A. Text Analysis Techniques

Yi et al. (2018) proposed a method for identifying fake job postings using text analysis techniques such as keyword extraction, sentiment analysis, and topic modeling. They found that fake job postings often contain exaggerated language, misspellings, and inconsistencies.

B. Machine Learning Models

Deng et al. (2020) developed a machine learning model based on features extracted from job descriptions, such as word frequencies and syntactic patterns, to classify job postings as genuine or fake. Their model achieved high accuracy in detecting fraudulent postings.

C. Social Network Analysis

Zhang et al. (2019) investigated the use of social network analysis to identify fake job postings on online platforms. They analyzed the connections between job posters, applicants, and other entities to uncover suspicious patterns indicative of fraudulent activity.

D. Crowd sourcing and User Feedback:

Shah et al. (2017) explored the use of crowd sourcing and user feedback mechanisms for fake job detection. They developed a system where users could report suspicious postings and provide feedback on the legitimacy of job offers, which improved the overall accuracy of the detection process.

E. Deep Learning Approaches

Wang et al. (2021) proposed a deep learning framework for fake job detection, leveraging techniques such as convolution neural networks (CNNs) and recurrent neural networks (RNNs) to extract features from job postings and classify them as genuine or fake.

F. Cross-platform Analysis

Liu et al. (2018) conducted a cross-platform analysis of job postings to detect fake listings. They compared job postings across multiple online platforms and identified inconsistencies and discrepancies that were indicative of fraudulent activity.

V. METHODOLOGY

The methodology for conducting a feasibility study for implementing the fake job post detection system involves a structured approach to evaluating its technical, economic, and operational feasibility. Below are the key steps involved in assessing the feasibility of integrating the fake job post detection system into existing job platforms. The target of this study is to detect whether a job post is fraudulent or not. Identifying and eliminating these fake job advertisements will help the job-seekers to concentrate on legitimate job posts only. In this context, a dataset from Kaggle [13] is employed that provides information regarding a job that may or may not be suspicious. This dataset contains 17,880 number of job posts. This dataset is used in the proposed methods for testing the overall performance of the approach. For better understanding of the target as a baseline, a multistep procedure is followed for obtaining a balanced dataset. Before fitting this data to any classifier, some pre-processing techniques are applied to this dataset. Preprocessing techniques include missing values removal, stop-words elimination, irrelevant attribute elimination and extra space not be sufficient enough. Adjustment of these parameters enhances the reliability of this model which may be regarded as the optimised one for identifying as well as isolating the fake job posts from the job seekers. This framework utilised MLP classifier as a collection of 5 hidden layers of size 128, 64, 32, 16 and 8 respectively. The K-NN classifier gives a promising result for the value k=5 considering all the evaluating metric. On the other hand, ensemble classifiers, such as, Random Forest, AdaBoost and Gradient Boost classifiers are built based on 500 numbers of estimators on which the boosting is terminated. After constructing these classification models, training data are fitted into it. Later the testing dataset are used for prediction purpose. After the prediction is done, performance of the classifiers are evaluated based on the predicted value and the actual value.

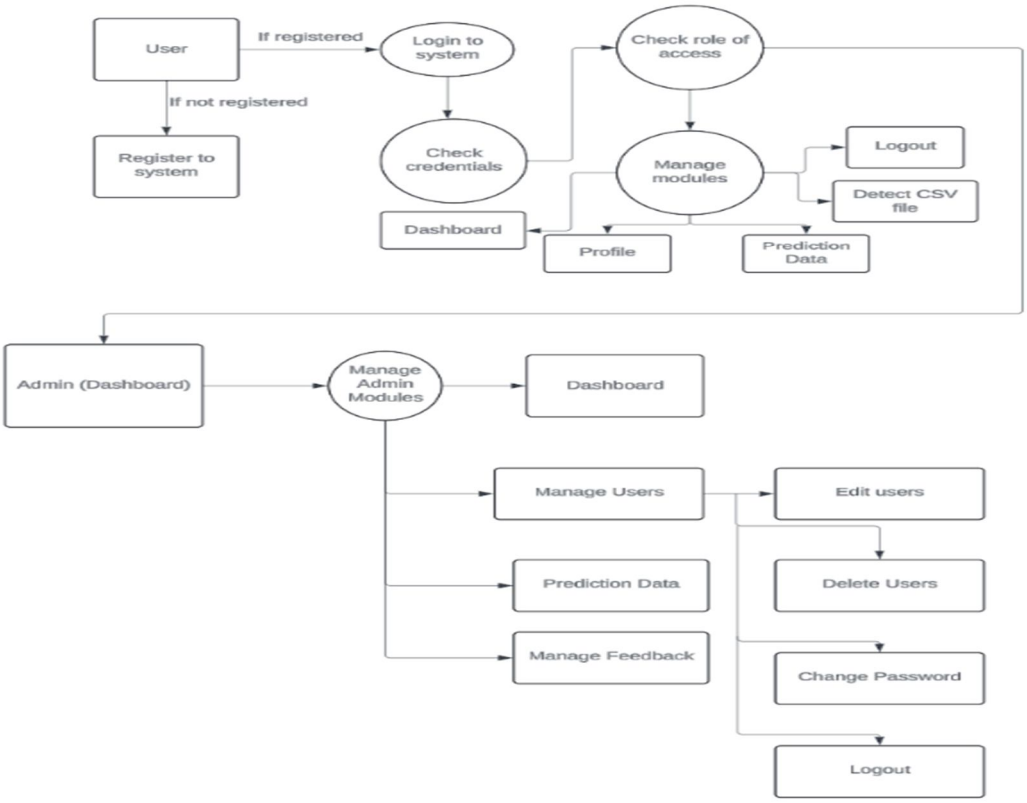


Fig.1:Data Flow Diagram

VI. RESULTS AND DISCUSSION

A. Machine Learning Models

1) Studies employing machine learning models for fake job detection have shown promising results. These models leverage features extracted from job postings, such as textual content, syntactic patterns, and metadata, to classify postings as genuine or fake.
2) Discussions around machine learning-based approaches often revolve around the choice of features, the selection of appropriate algorithms, and the balance between precision and recall in classification.

B. Text Analysis Techniques

1) Text analysis techniques, including keyword extraction, sentiment analysis, and topic modeling, have been effective in identifying linguistic cues indicative of fake job postings.
2) Discussions focus on the identification of key linguistic features that distinguish genuine job postings from fraudulent ones, as well as the robustness of text analysis methods across different types of job listings.

C. Social Network Analysis

1) Social network analysis approaches examine the connections between job posters, applicants, and other entities to uncover patterns of fraudulent behavior.
2) Discussions highlight the importance of considering the social context in which job postings are shared and the role of network structures in detecting fake job listings.

D. Crowd-sourcing and User Feedback

1) Crowd-sourcing and user feedback mechanisms have proven valuable in supplementing automated detection methods by leveraging the collective intelligence of users to identify suspicious postings.
2) Discussions often focus on the reliability of user-reported data, the design of effective feedback mechanisms, and the integration of human judgment with machine learning algorithms.

E. Deep Learning Approaches

1) Deep learning frameworks, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated strong performance in fake job detection tasks.
2) Discussions center around the interpretability of deep learning models, the scalability of training procedures, and the generalization capabilities across different job domains.

F. Semantic Analysis

1) Semantic analysis techniques, such as word embeddings and semantic similarity measures, offer insights into the underlying meanings of job postings, enabling the detection of deceptive language.
2) Discussions explore the nuances of semantic analysis in capturing subtle linguistic cues and the challenges associated with semantic ambiguity in job descriptions.

G. Blockchain-based Solutions

1) Blockchain-based solutions aim to create transparent and tamper-proof records of job postings, enhancing the traceability and authenticity of job listings.
2) Discussions often center around the scalability of blockchain technology, the integration with existing job platforms, and the potential impact on data privacy and security.

VII. CONCLUSION

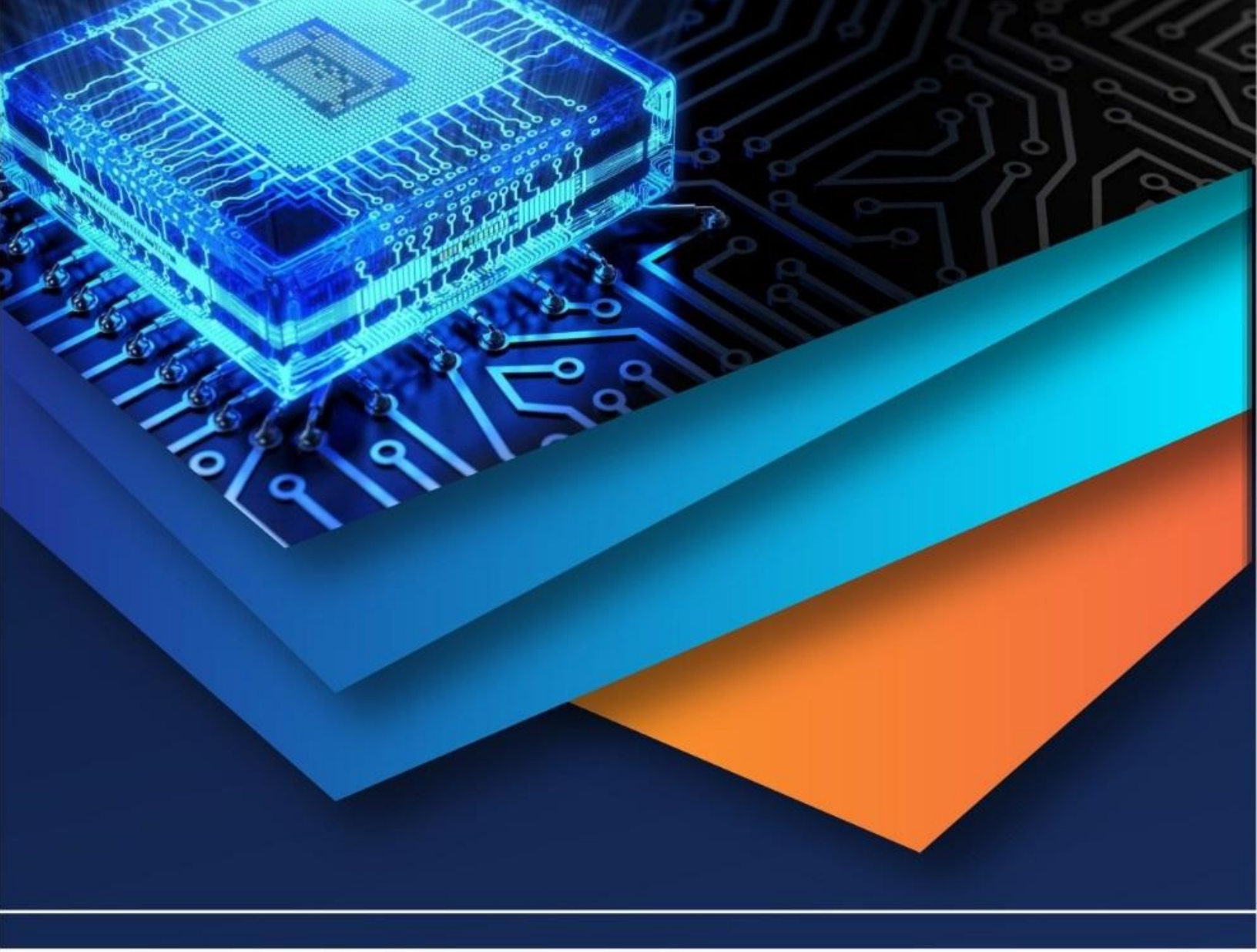
In conclusion fake job post detection using machine learning presents a critical solution to combat the proliferation of fraudulent activities in online job markets. By leveraging machine learning algorithms and advanced text analysis techniques, these systems can identify suspicious job postings and protect job seekers from falling victim to scams and deceitful practices.

Through this project, we have explored various machine learning models, feature engineering methods, and evaluation metrics to develop effective detection systems.

While the journey has been insightful, there remain numerous opportunities for further research and innovation to advance the field. Throughout this endeavor, we have delved into various machine learning models ,such as logistic regression, support vector machines, decision trees, and ensemble methods, to construct robust detection frameworks. Furthermore, feature engineering methodologies, including TF-IDF vectorization and word embeddings, have enhanced the systems' ability to capture subtle indicators of fraudulent behavior within job postings.

REFERENCES

[1] Shibly, F., Sharma, U., & Naleer, H. (2021). Performance comparison of two class boosted decision tree and two class decision forest algorithms in predicting fake job postings. Annals of the Romanian Society for Cell Biology, 25(4), 2462–2472.
[2] Dutta, S., & Bandyopadhyay, S. K. (2020). Fake job recruitment detection using a machine learning approach. International Journal of Engineering Trends and Technology, 68(4), 48–53.
[3] Cook J, Lewandowsky S, Ecker UK (2017) Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence. PLoS One 12(5):e0175799.
[4] Anita, C., Nagarajan, P., Sairam, G. A., Ganesh, P., & Deepakkumar, G. (2021). Fake job detection and analysis using machine learning and deep learning algorithms. Revista GEINTEC-Gestao, Inovacao e Tecnologias, 11(2), 642–650
[5] Aljedaani, W., Javed, Y., & Alenezi, M. (2020). LDA categorization of security bug reports in chromium projects. In Proceedings of the 2020 European Symposium on Software Engineering (pp. 154–161).
[6] Aljedaani, W., Nagappan, M., Adams, B., &Godfrey, M. (2019). A comparison of bugs acrosstheiOSandAndroidplatformsoftwoopen-sourcecross-platformbrowserapps. In 2019 IEEE/ACM 6th International Conference on Mobile Software Engineering and Systems (MOBILESoft) (pp. 76–86). IEEE.
[7] Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759.
[8] Rustam, F., Ashraf, I., Mehmood, A., Ullah, S., & Choi, G. S. (2019). Tweets classification on the base of sentiments for US airline companies. Entropy,21(11),1078.
[9] Sugumar, R. (2018). Improved performance of stemming using an efficient stemmer algorithm for information retrieval. Journal of Global Research in Computer Science, 9(5), 01–05.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)