International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

JobScamShield-Fake Job Post Detection

Prof. Tejashree Pangare[1], Ms. Sanskruti Surve[2], Ms. Kishori Salokhe[3], Mr. Sarthak Hasbe[4]

[1,2,3,4]Dept. of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, India

Abstract

JobScamShield is an AI-based system designed to detect and prevent fake job postings, ensuring safer online recruitment. It verifies company authenticity using GST numbers, registration details, and CEO or employee validation. The system applies Natural Language Processing (NLP) and Machine Learning (ML) techniques to analyze job descriptions, detect inconsistencies, and flag unrealistic salary offers. Integration with Glassdoor helps cross-check reviews and real job listings, while HR and interview question analysis further expose fraudulent patterns. By combining company verification, linguistic analysis, and behavioral checks, JobScamShield provides an intelligent, automated, and reliable tool that protects job seekers from scams and enhances trust in digital hiring platforms.

Keywords: Fake Job Detection, Natural Language Processing (NLP), Machine Learning (ML), Company Verification, Glassdoor Integration, Job Scam Prevention

I. INTRODUCTION

In The online recruitment platforms has made hiring process easier by the growing availability of online hiring websites. This recent wave of online hiring has, however, led to an increase in online job recruitment scams whereby scammers are increasingly using online job sites to carry out their activities. In addition, online hiring sites currently in use make use of traditional methods of moderation, complaints, as well as basic filtering methods, which are inefficient in coping with online hiring fraud cases.

Through the integration of machine learning for automated classification, natural language processing for text analysis, and real-time company verification, the system provides an automated, scalable, and reliable way to enhance security and trust in online recruitment process.deployment.

II. BACKGROUND

A. Growing Dependence on Online Recruitment Platforms-

With the rapid adoption of technology and development of information technology, online recruitment sites have become a important part of the employment process. Websites and online networking have become a major way of discovering employment opportunities and applying for jobs.as online recruitment sites become more widespread and prevalent, online recruitment has become a major way of seeking employment.

B. Complexity of Detecting Fraudulent Job Postings-

Job fraud detection is a complex process because the fraudulent techniques employed by scammers have become more advanced. Thus, the fraudulent jobs may appear to be more similar to the actual jobs by

using proper job titles. It may be more difficult to trace fraudulent jobs using traditional techniques, as it requires more inspection.

C. Limitations of Existing Website Builders-

The current job portals mostly use basic keyword filtering, human moderators, or end-user complaining to detect fraudulent job advertisement . Though they give some level of protection, they are not scalable or applicable to constantly varying scamming approaches. Also, current systems do not use any external data sources to verify the legitimacy of the organizations or detect behavioral characteristics in human resource conversations.

III. Motivation

1. Reducing Risks for Job Seekers

Job seeker browsing through online job portals does not have adequate knowledge about fraudulent job vacancies. Scam job vacancies appear to be real and authentic, and users might find difficulty while determining real and fake job vacancies.

2. Early Detection of Job Scams

Generally, the traditional method of scam detection involves a reactive approach where a person's complaints are triggered after suffering losses from a scam. The automated scam detection can detect suspicious job postings at an early stage to avoid loss or misuse of a person's information.

3. Minimizing Financial and Emotional Loss

Scams about jobs often cause economic loss and emotional agony, especially to young students or fresh graduates. However, the timely alerting of scams also aids in avoiding the economic loss or emotional distress caused to victims.

4. Improving Accuracy through Multi-layered Verification

Currently, it has been noticed that simply through text-based scams, it becomes difficult to detect scams. It requires integration with company verification and behavior analysis.

5. Leverage AI and Real Time Verification Technologies

Advancements in areas like Artificial Intelligence, Natural Processing Languages, and data verification APIs help in building intelligent systems that are adept in understanding online job postings. The ability to make use of these advanced technologies makes way for building highly efficient and scalable systems in preventing online recruitment scams.

IV. Literature Review

Recent As online job boards proliferate, the cybersecurity sector is increasingly concerned about job-related scams. To overcome this problem, the researchers are suggesting the use of Machine Learning (ML) and Natural Language Processing (NLP) techniques to identify fake job postings.

M.P. Manikandaprabhu [1] developed a method for identifying fraudulent job postings, employing machine learning techniques such as Naive Bayes and Logistic Regression. The system utilized TF-IDF features and analyzed job descriptions to ascertain the authenticity of the postings.

The model was mostly right, but it only looked at text dataand didn't check to see if the business was real.

Madhavi et al.[2]A Random Forests-based method was developed to find fake job ads online. The study's results showed that ensemble models performed better than single classifiers, especially when the data is imbalanced.

But the answer didn't let companies or recruiters see their information right away.

Aliedaani et al. [3] utilized NLP preprocessing techniques and ML classifiers to detect fraudulent job advertisements.This study showed the importance of both feature extraction and ensemble learning techniques.

But it only looked at what was in job descriptions and not at other ways that people could get a good reputation.

Quihui and Espinosa [4] looked at a few machine learning algorithms for finding fake job postings and found that the Random Forest and Gradient Boosting models worked better. The research suggested that text-based models might not work when fraudsters use job descriptions that are both real and well-written.

S. Brightwood [5] examined the efficacy of various fake job detection models and identified elevated false-positive rates as a significant issue. The study indicated that the findings would be more reliable if content analysis were employed in conjunction with additional verification methods.

Patel et al.[6]The use of natural language processing (NLP) and behavioral patterns to detect online fraud was examined by their study showed that looking at how recruiters communicate can help find suspicious activities, even though the system wasn't made for job portals in the first place.

Singh and Verma [7] developed a fraud detection method using machine learning and data validation. While the existing system helped reduce general online fraud, it didn't specifically address recruitment scams.

V. LIMITATIONS

While the efficiancy of JobScamShield in combating fraudulent job posting detection has been established, some limitations are evident. The limitations are data-oriented, incorporating the dependency of data on generalized modeling, along with the constraints that require verification. Recognizing these limitations aids in creating the scope of the system.

1. **Dependence on Quality of Training Data-**

Additionally, the performance of machine learning models in JobScamShield would largely depend on its training data. If a training dataset does not capture various newly developed scam trends adequately, JobScamShield would not perform very well.

2. Challenges in Detecting Highly Sophisticated Scams-

Some fraudsters also make use of advanced tactics like imitating genuine firms or using actual details of genuine vacancies and conducting genuine interviews. The use of text- based evaluation alone is not effective in these situations, and some reliance upon external verification is needed.

3. Limited Behavioral Analysis Scope-

The system analyzes HR communications and interview patterns based on behavioral criteria. However, human behavior is quite diverse. A number of real recruiters inadvertently display patterns that are similar to those that indicate scamming. This accounts for some false positives.

4. Scalability Constraints with Real-Time Verification-

Real-time company verification carried out with help of GST validation and external APIs could result in a lag if a high number of job postings are carried out at a time. High traffic situations may need optimizations.

5. Reliance on External Information Sources

The success of JobScamShield also depends upon third-party resources like company register

information and Glassdoor reviews through API calls for verification purposes. Any change in their availability or API access policy will affect its overall reliability.

VI. Existing Systems

Various approaches already exist to detect and identify fraudulent job posts on these platforms. All these approaches are based on some artificial intelligence or simple techniques of machine learning. The table given below describes the prospects of fraud existing in these approaches or systems and our proposed JobScamShield system.

1. Manual Moderation and User Reporting

Currently, job portals heavily rely on user complaints and administration reviews in order to detect cases of falsifying jobs, a practice that is reactive in nature as cases of scams will only be detected after a user has been a victim of a scam.

2. Rule-Based Scam Detection Systems

These systems are designed to detect suspicious job postings by using predetermined keywords and rules. These systems are simple to use, though they can be evaded by modifying the patterns of words.

3. ML-Based Text Classification Systems

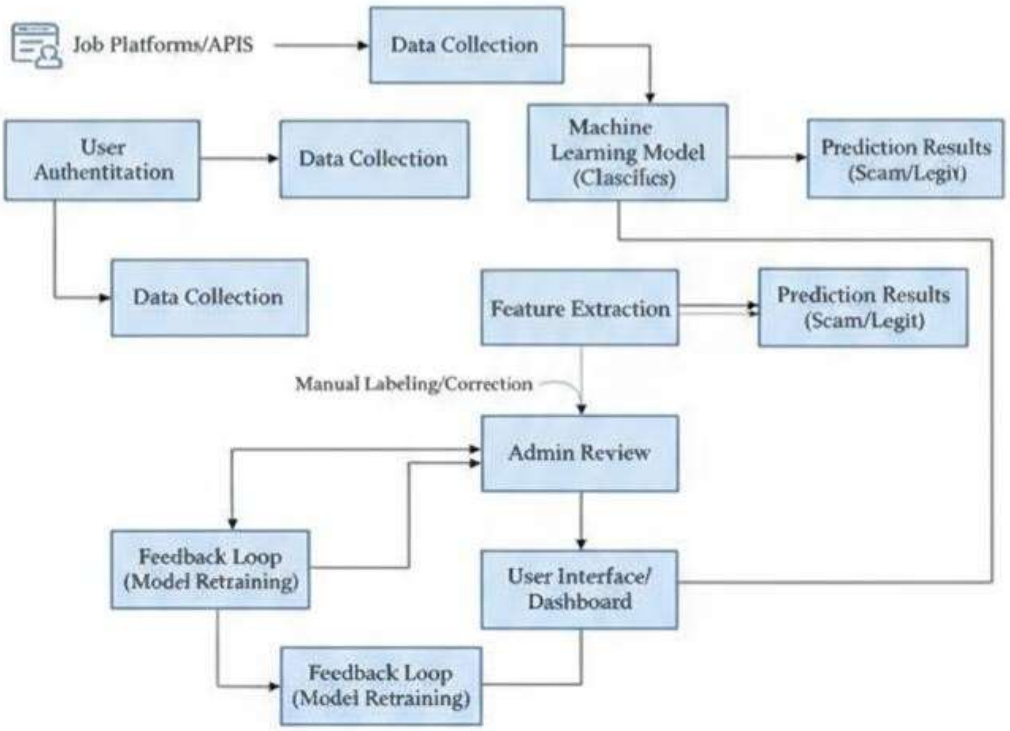
Machine learning models analyze descriptions of jobs to ascertain if a post is genuine or fraudulent. This increases accuracy levels; however, there is no verification with companies or behavioral analysis.

Feature	Manual & Rule- Based Systems	ML-Base d Syst ems	JobSca m- Shield (propos ed)
Manual Intervention Required	Yes	Limit ed	No
AI-Based Detection	No	Yes	Yes
NLP Text Analysis	No	Yes	Yes
Company Verification	No	No	Yes

VII. Problem Statement

The rise in fraudulent job postings online presents serious risks to job seekers. Most current detection systems are supported by text classification or manual reporting; neither is adequate to track sophisticated scams that use language and company identities that appear legitimate. What is required is an automated system to analytically review job descriptions intelligently, validate company authenticity, and detect anomalies in behavior in real time. The objective of JobScamShield is to design and implement a comprehensive AI-based solution that accurately classifies job postings as genuine or fraudulent while minimizing user effort and improving online recruitment security.

VIII. PROPOSED ARCHITECTURE



IX. METHODOLOGY

The proposed system, named JobScamShield, is designed to mitigate the emergence of fraudulent job posts using a combination of Machine Learning and Natural Language Processing concepts to process the related data to be categorized as legitimate or fraudulent job posts. Its key considerations include the need for high precision, speedy response, security, scalability, and ease of use.For its development, the proposed system would be created using the Python scripting language with the ability to be deployed in either a desktop or cloud environment. The system design provides a definition for the entire workflow or architecture of JobScamShield.This involves various modules for data collection/preprocessing, model training, and classification. Natural language processing techniques are used to obtain essential feature representations for jobs, where a machine learning model is used for data classification. The design also enables easy extension or update in the near future. All modules are combined to form a single system with integration using Python with Flask/Django during its implementation.

The simplicity of this interface will ensure that users can just input job description data to get results on its authenticity instantly. The use of a database will allow for storing results for subsequent continuous model monitorings. To ensure accurate and correct operation, a series of procedures known as "testing and validation" is carried out on the system. In addition, various methods and approaches of "testing" ensure the proper functionality and usability of the system. The performance of the system is determined through various standard metrics, including "accuracy," "precision," "recall," and "F1-Score." The security model for this system includes data anonymization, communication between parties, and access.The system then becomes ready for running after testing and can run locally or virtually. The system can then be updated for accuracy and for accommodating emerging job scam models

X. RESULT ANALYSIS

The effectiveness of the proposed system was validated using existing data sets, where real and phishing information was used. The data set used for evaluation includes real and phishing information collected from various online job portals. It was preprocessed using various Natural Language Processing techniques. TF-IDF vectorization was used for vector representation.

Various machine learning models, such as Logistic Regression, Random Forest, and XG Boost, have been trained, tested, and implemented on the processed data set. The performance of the system has been tested by utilizing metrics such as accuracy, precision, recall, and F score, where it has been observed that ensemble-based models provide better accuracy than standalone models.

Besides classification through texts, company validation has also been incorporated into this system to make it even more reliable. Organizational details were also verified through GST number validation and reputation checking, increasing the accuracy of detection against false positives.

The user testing shows that, indeed, the output of JobScamShield can be understood by job seekers. The overall picture revealed by this assessment is that the proposed multi- layered approach outperforms the traditional text-based detection approach.

XI. CONCLUSION AND FUTURE SCOPE

The JobScamShield project aims to create an automated system which will detect and prevent user from fraudulent online job postings through its complete AI-driven solution. The system will be using Natural Language Processing (NLP) and Machine Learning (ML) technologies to analyze job descriptions and find linguistic errors . The project also includes an advanced module which uses external data points for validating the authenticity of the posting organization through GST numbers, addresses, CEO verification, employee identification.

The system uses sentiment analysis to check user reviews from Glassdoor against actual job postings while studying HR practices and interview questions to identify detection of suspicious or fraudulent hiring practices. The system uses XGBoost and Random Forest algorithms for classification which results in better accuracy than the current systems ,allowing job seekers to use a trustworthy platform for finding genuine job openings.

REFERENCES

1. M. P. Manikandaprabhu, “Machine Learning Algorithm for Fake Job Detection Systems,” 2024.

2. D. R. Madhavi et al., "Detection of Online Employment Scam Through Fake Jobs Using Random Forest Classifier," 2022.
3. MDN A. Aliedaani et al., "Detection of Fake Job Postings Using Machine Learning and NLP Approaches," 2022.
4. P. Quihui and G. P. Espinosa, "Fake Job Detection with Machine Learning: A Comparative Study," 2023.
5. S. Brightwood, "Evaluating the Performance of Fake Job Detection Models," 2024.
6. R. Patel et al., "Behavioral Analysis for Online Fraud Detection Using NLP," 2021.
7. A. Singh and P. Verma, "Machine Learning-Based Framework for Online Fraud Detection," 2020.