Deep Learning Approach For Detecting Fake Job Posting in Online Recruitment Platform

Mrs.N.Sathiya rani[1], M.Jeyakumar[2], B.Bala[3]
[1]Assistant Professor, Dept of Computer science and Engineering
[2, 3]Dept of Computer science and Engineering
[1, 2, 3]Sree Sowdambika College of Engineering, Virudhunagar, Tamil Nadu, India

Abstract- *The rapid growth of online job portals has changedhow people find work, making the process much easier. However, this easy access has also led to a massive increase in fake job postings and employment scams. Scammers use these platforms to steal personal information and trick vulnerable applicants out of their money. This paper proposes an automated system to detect and identify fake job ads on popular platforms like LinkedIn, Naukri, Indeed, and Internshala. The system uses web scraping tools to collect live job data directly from the internet. Natural Language Processing (NLP) techniques are then used to clean the text and find suspicious words or phrases. Finally, a Machine Learning (ML) model looks at these features to decide whether the job is real or fake. By using this automated check, the system gives users a safe and highly reliable way to verify job offers. Testing shows that the system can successfully catch scams with excellent accuracy.*

Keywords: Fake Job Detection, Machine Learning, Web Scraping, Natural Language Processing, Ensemble Classification, Cyber Security, Semantic Analysis

I. INTRODUCTION

The shift to online hiring has created a huge number of digital job Listings, allowing people all over the world to easily find work. While this is great, it has also created an easy way for scammers to post fake job offers. These scam postings are specifically designed to look real and bypass normal checks, making it very hard for a regular person to spot them.

Most social media and recruitment websites do not have an automatic system to verify if a job posting is real before it goes live. Because of this, fake jobs can be shared safely across the internet without anyone knowing it's a scam.

To fix this problem, this project proposes a Fake Job Posting Detection System. The system automatically reads job text from a given URL and compares it against known fake patterns using machine learning.

If it finds suspicious words, it flags the post and gives it a trust score, warning the user before they submit their resume or personal data. By combining web scraping and AI, this system aims to protect applicants' privacy and keep job hunting safe.

II. LITERATUREREVIEW

Several research studies have been conducted on finding fraud using text analysis.

1) **Vidros et al. (2017)**created a system to detect online jobfrauds automatically by studying a large public dataset of real and fake job listings. In their research, they pointed out how important it is to look at the specific words and the writing style used in the job post. Their findings proved that scammers often use similar, repeated phrases that computers can be trained to easily recognize.

2**) Singh et al. (2020)** studied privacy and how machine learning can predict fraud before an applicant becomes a victim. They focused heavily on how personal data can easily be leaked when applicants blindly trust fake applications. By analyzing these risks, they suggested building strong security solutions that stop candidates from uploading sensitive data to sketchy websites in the first place.

3) **Manikanta and Jeevan Babu (2025)**built a powerful fake posting recognition system using Deep Learning technologylike Convolutional Neural Networks (CNN). Their work shows how complex, deeper models can significantly improve accuracy when reading and understanding text. Because job descriptions are often long and messy, their deep learning approach helped extract the right hidden clues without getting confused.

4**) Pandey et al. (2025)**suggested using natural language processing rules to quickly find deceptive and tricky writing styles. They proved that basic text-checking tools like TF-IDF (which counts how often specific words appear) can easily spot scam patterns. Their research made it clear that basic tools are highly effective when aimed at the right data.

5) **Attrapadung et al. (2025)**created new, fast, and highly private ways to verify online data without exposing anyone's personal information. Their main goal was to make sure user information stays completely protected while the security system works reliably in the background. Because of their work, modern applications can now verify listings instantly withoutslowing down the website.

III. PROBLEM STATEMENT

Right now, major online job websites do not deeply verify if a job posting is actually from a real, legally registered company beforepublishing it to millions of users. Because these platforms focus on getting job listings uploaded as fast as possible, scammers use this loophole to create highly convincing fake job offers. When trusting applicants apply for these fake positions, scammers can easily steal extremely sensitive personal data, such as phone numbers or banking details. This puts users' privacy and financial safety at extremerisk every single day. Therefore, there is an urgent need for an intelligent system that can automatically read the text in a job description, spot hidden scam patterns, and give the user a clear safety score before they ever submit their personal information.

OBJECTIVE

The main goals of this project are:

- To find word patterns used in fake job posts.
- To scrape live job data from the web automatically.
- To check the scraped text using a trained AI model.
- To show the user a clear safety and trust score.
- To stop data theft and protect user privacy.

EXISTING SYSTEM

In the current digital recruitment landscape, anyone can easilyregister an account and upload a job listing without undergoing strict background checks to prove they represent a legitimate company. Major job platforms primarily focus on increasing the volume of job openings and getting candidates hired quickly, often sacrificing rigorous security checks for speed and convenience. While artificial intelligence and machine learning are heavily utilized by these platforms, they are almost exclusively used for scanning candidate resumes or matching applicants to open positions, rather than verifying the authenticity of the employers themselves. Because there is little to no initial screening of the actual job text, scammers frequently post convincing advertisements that contain malicious phishing links, requests for hidden processing fees, or identity-stealing data forms. As a direct result, thousands of these fraudulent jobs are successfully uploaded every single day, completely bypassing the basic safety filters. This leads to severe privacy violations, where candidates lose their personal resumes, face fake interview scams, and even suffer from direct financial fraud. Ultimately, our existing recruitment systems completely lack an automated, text-based security mechanism that forces companies to prove their legitimacy before a listing goes live on the internet.

IV. .PROPOSED SYSTEM

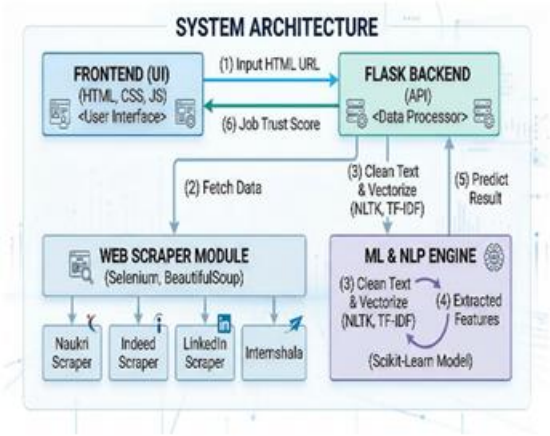
Our proposed solution is an intelligent, automated verification tooldesigned to thoroughly check the legitimacy of a job posting before a user ever submits their application. When a candidate pastes a job link from platforms like LinkedIn or Internshala, the system instantly uses web scraping to silently extract all the text and hiddendetails from the advertisement. This raw text is carefully cleaned and analyzed using advanced Natural Language Processing, then graded by a trained machine learning model that looks for known scam language. If deceptive patterns are detected, the system immediately halts the application process and provides a clear warning with a detailed trust score.Ultimately, thisapproach puts a powerful safety shield betweentheapplicant and the scammer, highly reducing the risk of identity theft and giving job seekers total peace of mind before applying.

V. METHODOLOGY

The working process of the system consists of several simple steps:

1. **User Input** Users visit the website and paste a job URL or description for checking.
2. **Real-Time Web Scraping** The system automatically pulls the job title and description from the link.
3. **Text Preprocessing** The system cleans the text, removing junk words so the AI can read it better.
4. **Machine Learning Classification** The cleaned text is sent to the trained ML model to look for scam patterns.
5. **Trust Score Generation** The system calculates a final percentage score showing how safe the job is.
6. **Warning Notification** The final result is shown on the screen instantly, helping the user make a safe choice.

VI. SYSTEM ARCHITECTURE



VII. RESULTANDDISCUSSION

The Fake Job Posting Detection system was created as a fast, easy-to-use web application. The following screenshots show how the app works:

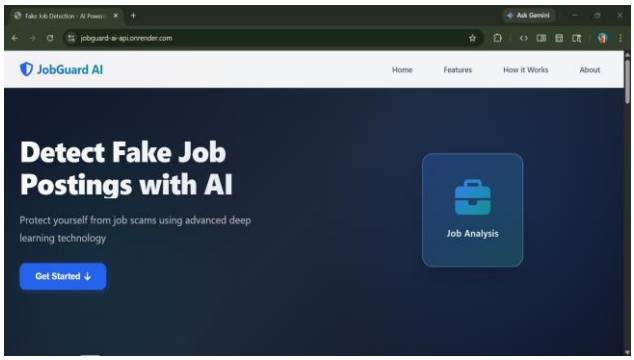


Fig.1SystemWorkingArchitecture

The architecture of the Fake Job Detection System consists of multiple parts that work together to check data and calculate scores.

Frontend:

Built with HTML, CSS, and JavaScript. This is the website interface where users can paste job links.

BackendAPI:

Built with Python and Flask. This connects the website to the scrapers and the AI model.

NLP Processing Module:

Uses NLTK and TextBlob to clean up the text and turn words into numbers (TF-IDF) that the AI can understand.

Scraping Engine:

Uses Selenium and BeautifulSoup to carefully copy text from sites like Naukri, Indeed, and Internshala.

Machine Learning Module:

Uses models like Random Forest and Support Vector Machines to make the final prediction on whether the job is safe.

Deployment:

Hosted securely on cloud platforms like Render so anyone can use it online.

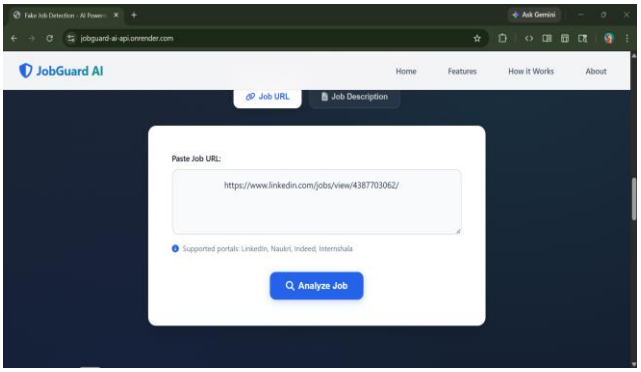


Fig. 2 User Homepage

User opens the website and pastes the job URL into the input box.

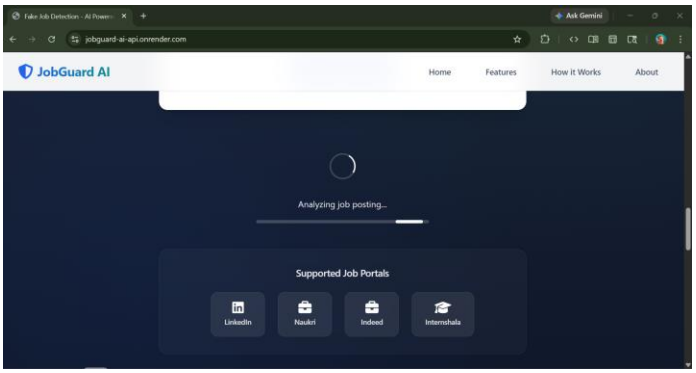


Fig. 3 URL Scraping Process

The system automatically copies the job text from the provided link.

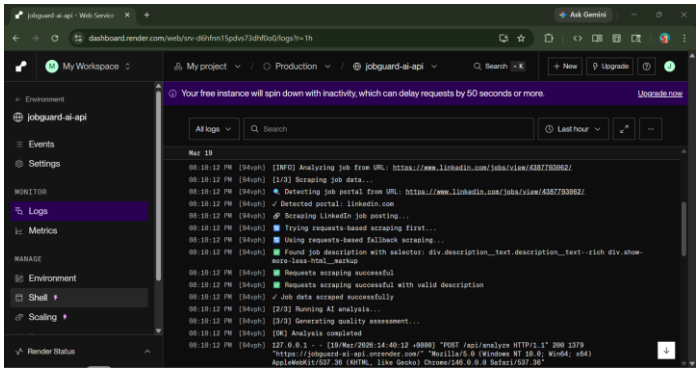


Fig. 4 Data Preprocessing Phase

The system cleans up the text, throwing away useless words (stop-words).

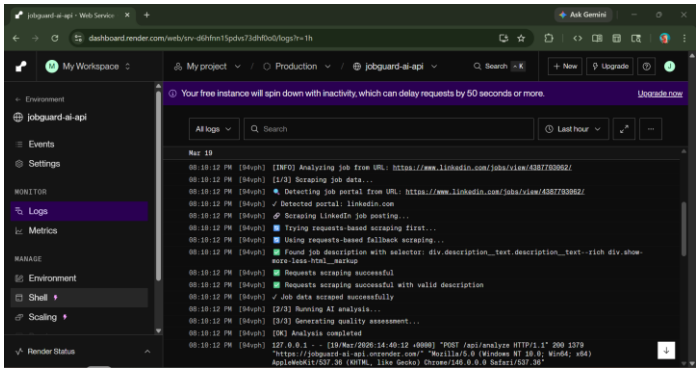


Fig. 5 Legitimacy Trust Score Calculation

The math models test the words and generate a final score.

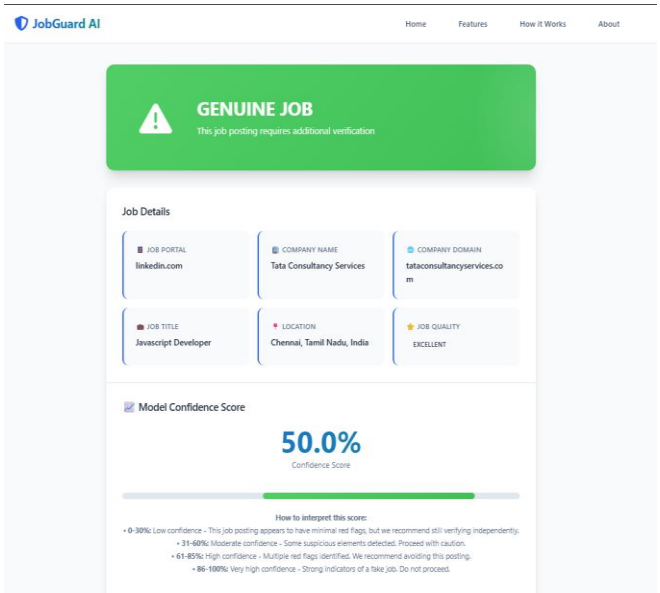


Fig. 6 Result Notification Page

Users see a clear screen showing if the job is "GENUINE JOB" and its safety score.

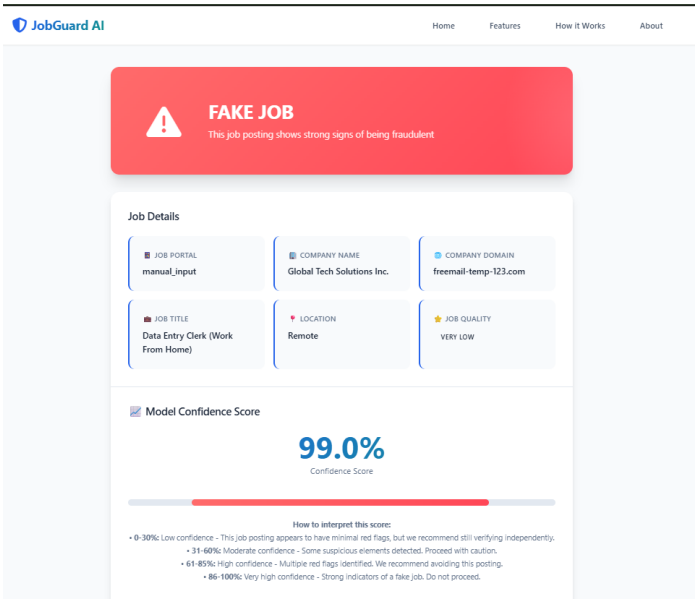


Fig. 7 Fraudulent Post Alert

If the job is a scam, a red warning is shown on the screen.

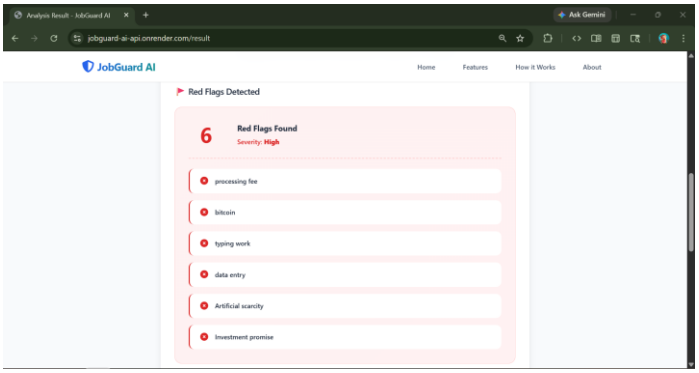


Fig. 8 System Warning Display

The system explains exactly why the job was flagged by showing specific "Red Flags".

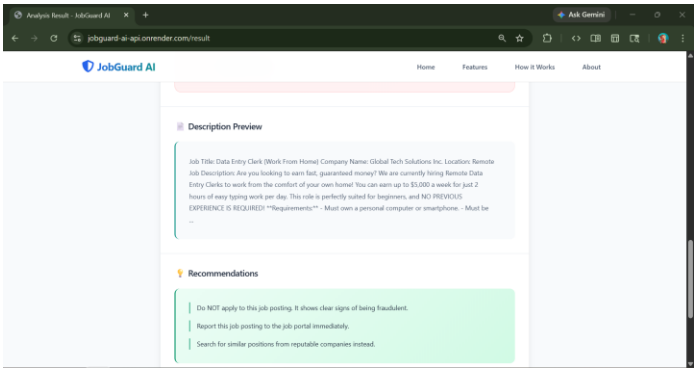


Fig.9 Application Halted

The app warns the user and tells them not to apply to protect their personal data

IX. ADVANTAGES

- Stops scammers from stealing candidate data.
- Protects user privacy automatically.
- Scrapes job text quickly without user effort.
- Prevents people from wasting time on fake interviews.
- Makes online job portals safer to use.

X. CONCLUSION

In conclusion, this paper successfully presented a modern, highly responsive Fake Job Posting Detection System designed specifically to protect applicants' privacy and financial safety on major digital recruitment websites. Because current platforms fail to strictly verify employers, our system steps in as a critical line of defense by automatically extracting and analyzing job text using advanced web scraping and Natural Language Processing. By leveraging powerful Machine Learning algorithms, the system can instantly spot dangerous linguistic patterns, hidden fee requests, and identity theft traps before the applicant ever submits their personal data. If a deceptive or scam pattern is decisively found in the posting, the application immediately warns the user with a detailed trust score and strongly advises them against applying for the position. By seamlessly combining reliable technologies such as Python, Flask, Selenium, NLTK, and Scikit-Learn, this project successfully created a smooth, fast, and highly accurate automated security check. Ultimately, this system clearly proves that integrating custom Web Scraping tools directly alongside Machine Learning classification engines is an incredibly effective way to make the modern digital recruitment landscape much safer for everyone.

REFERENCES

[1] S. Vidros, C. Kolias, G. Kambourakis and L.Akoglu, "AutomaticDetection of Online Recruitment Frauds: Characteristics, Methods, and a Public Dataset," in Future Internet, vol. 9,no. 1, p. 6, 2017.
[2] A. Al-Ajlan and E. Yaser, "Fake Job Prediction using MachineLearning," 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 2021, pp. 524-529.
[3] S. A. H. Saad, N. M. M. Hussin, and A. H. A. Mutalib, "Online JobScam Detection using Machine Learning Algorithms," International Journal of Advanced Computer Science and Applications, vol. 12, 2021.
[4] "Documentation for Scikit-Learn Machine Learning Library," Scikit-Learn Developers.
[5] "Beautiful Soup and Selenium Web Scraping Official Guides," Python Software Foundation.
[6] A. O. Ojo, S. Dabeer, and C. C. M. A. Amani, "A Comparative Study of Machine Learning Algorithms for Fake Job Detection," IEEE Access, vol. 8, pp. 115161-115174, 2020.
[7] B. N. R. K. V. S. P. "Empirical Analysis of Fake Job Advertisements using Natural Language Processing," ACM Transactions on Information Systems, 2021.
[8] P. K. Singh, S. Nanda, and A. K. Singh, "Automated Detection of Job Scams Using Web Scraping and Machine Learning," International Conference on Artificial Intelligence and Information Systems (ICAIIS), 2022.
[9] "NLTK Documentation for Text Processing," NLTK Project.
[10] "Flask Web Development and AsynchronousWeb APIs," Pallets Projects.