

Online Recruitment Fraud Detection Using Machine Learning and Deep Learning Techniques

S. Mano Venkat¹, P. Madhuri², P. Chandrika³, M. Prasanth⁴, K. Harika⁵

¹Assistant Professor, Computer Science And Engineering, Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh, India

^{2,3,4,5th} B.Tech Final Semester, Bachelor of Technology, Computer Science And Engineering, Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh, India

Abstract - The rapid growth of online recruitment platforms has transformed the hiring process for both employers and job seekers. However, it has also increased the number of fraudulent job postings and fake recruiters, leading to financial loss and misuse of personal information. Therefore, detecting recruitment fraud has become essential to ensure the reliability and security of online job portals. Existing systems rely on manual verification and traditional machine learning algorithms such as Support Vector Machines (SVM) to identify fraudulent job postings by analyzing features like job descriptions, company profiles, salary details, and job requirements. Although these methods assist in fraud detection, their performance is limited when handling complex data patterns, achieving an accuracy of around 70%.

To address these limitations, the proposed system utilizes advanced machine learning techniques including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Random Forest algorithms. These models effectively learn complex patterns from large datasets and improve classification performance. By analyzing recruitment-related features, the proposed system achieves an accuracy of approximately 90%, thereby enhancing fraud detection efficiency and ensuring a safer and more reliable online recruitment environment.

Key Words: Online Recruitment Fraud (ORF), Deep Learning, Fraud Detection, Neural Networks, Python, Flask, Web Application, Machine Learning, Data Classification, Recruitment Systems

1. INTRODUCTION

With the increasing digitization of recruitment processes, online job portals have become a primary medium for job seekers and employers. These platforms provide easy access to job opportunities; however, they have also led to a rise in fraudulent activities. Fake job postings and scam recruiters are increasingly targeting applicants, resulting in financial loss and misuse of personal information. This makes online recruitment fraud a significant concern and highlights the need for a secure recruitment environment.

In the modern digital era, advanced technologies such as deep learning can be utilized to address these challenges effectively. By analyzing textual and behavioral patterns in recruitment data, intelligent systems can identify suspicious activities and reduce the risk of fraud. However, ensuring accuracy and efficiency in detecting complex fraud patterns remains a critical challenge.

To address this issue, the project titled "Online Recruitment Fraud Detection System using Artificial Neural Networks" focuses on developing a smart and reliable solution for identifying fraudulent job postings. The system analyzes recruitment-related data such as job descriptions, company details, and user inputs to classify job listings as genuine or fraudulent. By leveraging advanced techniques including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Random Forest algorithms, the system improves detection accuracy and efficiency. This approach not only reduces the risk of fraud but also enhances user trust and ensures a safer and more reliable online recruitment environment.

2. Existing System

Currently, most online recruitment fraud detection systems rely on traditional methods such as manual verification and basic rule-based approaches. These systems involve checking employer credentials, analyzing job postings, and using user reports to identify suspicious activities. However, these approaches are not effective in detecting advanced and evolving fraud patterns.

In addition to manual methods, machine learning techniques such as Support Vector Machines (SVM) are used to classify job postings based on features like job description, company profile, salary details, and job requirements. Although SVM helps in identifying fraudulent activities, it has limitations when dealing with large and unstructured datasets.

Many existing solutions depend on predefined rules and manual feature extraction, which limits their flexibility and scalability. Manual verification is time-consuming and makes it difficult to handle large volumes of recruitment data. These systems often fail to capture hidden patterns, leading to lower detection accuracy.

Additionally, existing systems cannot adapt quickly to new fraud techniques, as fraudsters continuously change their strategies. As a result, many fraudulent job postings remain undetected, exposing job seekers to risks. These limitations highlight the need for a more efficient fraud detection system.

3. Challenges of Existing System

The major limitations of existing systems are as follows:

Existing systems rely on manual verification, which is time-consuming and prone to human errors. Traditional techniques such as Support Vector Machines (SVM) have limitations in handling large and unstructured data like job descriptions and company information. These systems depend on predefined rules, reducing their ability to detect complex fraud patterns. They also lack real-time detection and struggle to adapt to evolving fraud techniques, which affects overall accuracy. Additionally, these systems often fail to capture hidden patterns in data and may produce incorrect predictions, reducing reliability. They are also less efficient when processing large volumes of recruitment data and require continuous manual intervention for better performance.

Overall, these challenges reduce the efficiency and reliability of fraud detection systems and increase the risk of fraudulent job postings.

4. Proposed System

The proposed system, Online Recruitment Fraud Detection System using Artificial Neural Networks (ANN), is designed to overcome the limitations of existing fraud detection methods. The system focuses on providing an automated, accurate, and scalable solution for identifying fraudulent job postings and improving the security of online recruitment platforms.

The system collects and processes job-related data such as job descriptions, company details, salary information, and other relevant features. Advanced machine learning and deep learning techniques are applied to analyze this data and detect fraudulent patterns. Artificial Neural Networks (ANN) are used to learn complex relationships within the dataset and classify job postings effectively. Convolutional Neural Networks (CNN) are utilized to extract meaningful features from unstructured textual data such as job descriptions, while Random Forest algorithms enhance classification accuracy by combining multiple decision trees and reducing overfitting.

The main components of the system are designed to ensure efficient processing and reliable performance. The system performs data preprocessing, feature extraction, and classification using trained models to identify whether a job posting is genuine or fraudulent. It is capable of handling large volumes of recruitment data and provides faster and more accurate predictions compared to traditional methods. The system also supports efficient data handling and improves overall model performance through optimized processing techniques. This ensures better scalability and consistency in fraud detection across different types of recruitment data.

Users can input job-related information into the system, and the model analyzes the data to predict whether the job posting is real or fake. Based on this prediction, the system helps users identify fraudulent listings and avoid potential scams. By enabling real-time fraud detection and improving accuracy, the proposed system enhances user safety and builds trust in online recruitment platforms.

5. Advantages of Proposed System

1. High Detection Accuracy: The system uses Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Random Forest algorithms to improve classification performance and accurately identify fraudulent job postings.

2. Automation of Fraud Detection: The system automates fraud detection, reducing manual effort, saving time, and minimizing human errors.

3. Real-Time Detection: It detects fraudulent job postings instantly as they are uploaded, ensuring quick response and reducing the risk of exposure to fake opportunities.

4. Efficient Handling of Large Data: The system processes large volumes of recruitment data efficiently, making it suitable for large-scale platforms and continuous data flow.

5. Deep Feature Extraction: CNN extracts meaningful features from unstructured text data, improving the detection of hidden and complex fraud patterns.

6. Reduced False Predictions: The system minimizes false positives and false negatives, improving reliability compared to traditional methods.

7. Scalability and Flexibility: It is scalable and can handle increasing data efficiently while supporting future enhancements, ensuring consistent system performance.

8. Enhanced User Safety: The system protects job seekers by identifying and filtering fraudulent job postings, reducing the chances of scams. It helps users make safer decisions by providing accurate predictions and minimizing exposure to fake job opportunities in online platforms.

9. Adaptability to New Patterns: It learns from new data and adapts to evolving fraud techniques, ensuring long-term effectiveness. This continuous learning capability allows the system to identify new and unseen fraud patterns, improving detection performance over time and maintaining robustness in dynamic environments.

10. Improved Platform Trust: The system enhances trust and credibility in online recruitment platforms by providing reliable fraud detection. It helps users confidently interact with job portals and ensures

a safer recruitment ecosystem by reducing fraudulent activities and improving overall platform reliability.

6. Algorithm and Techniques

6.1 Machine Learning-Based Fraud Detection Algorithm

The proposed system uses a machine learning and deep learning-based algorithm to analyze and classify job postings as genuine or fraudulent. The step-by-step workflow of the algorithm is explained below:

The working process of the system is carried out in a structured sequence of steps.

In the first step, the system accepts job posting data as input, including job descriptions, company details, salary information, and other relevant features provided by the user or dataset.

In the second step, data preprocessing is performed. The input data is cleaned by removing noise, handling missing values, and converting text into a suitable format for further analysis.

In the third step, text processing and feature extraction are carried out. Techniques such as tokenization and encoding are applied to convert unstructured textual data into structured numerical form.

In the fourth step, the processed data is fed into deep learning and machine learning models including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Random Forest algorithms.

In the fifth step, ANN learns complex relationships within the data, while CNN extracts deep features from textual inputs, and Random Forest improves classification accuracy by combining multiple decision trees.

In the sixth step, the system analyzes the processed data and classifies the job posting as genuine or fraudulent based on learned patterns.

In the seventh step, the system generates the prediction result and displays whether the job posting is real or fake to the user.

Finally, in the eighth step, all input data and prediction results are stored in the database for future analysis, monitoring, and system improvement.

6.2 Fraud Detection Scoring Logic

The Fraud Detection Score represents the likelihood of a job posting being genuine or fraudulent. The scoring logic is defined as follows:

- If the **job posting is classified as genuine** by the model
→ **Score = 80-100 (High Reliability Score)**
- If the **job posting shows suspicious patterns**
→ **Score = 40-70 (Moderate Risk Score)**
- If the **job posting is classified as fraudulent**
→ **Score = 0-30 (High Fraud Risk Score)**

This scoring mechanism helps users understand the reliability of job postings and identify whether they are safe, suspicious, or fraudulent.

6.3 Fraud Risk Estimation

The system estimates the fraud risk of job postings using predefined patterns and learned features from recruitment data. When suspicious indicators such as unrealistic salary, incomplete company details, or unusual contact information are detected, the system assigns a high fraud risk score, indicating a strong possibility of fraud. In contrast, if valid company information and consistent job details are present, a low fraud risk score is assigned.

For job postings that do not clearly match either genuine or fraudulent patterns, the system assigns a moderate risk score based on overall analysis. This mechanism helps users evaluate the reliability of job postings and supports them in making safer decisions while applying.

6.4 Technologies Used

The system is developed using the following technologies:

web interface for user interaction, data input, and result display.

Backend:

Python (Flask) – Handles application logic, user requests, and fraud detection processing.

Database:

SQLite / MySQL – Used for storing user data, job details, and prediction results.

Machine Learning & Deep Learning:

TensorFlow, Keras, Scikit-learn – Used to implement ANN, CNN, and Random Forest models.

Security:

User Authentication System – Ensures secure login and protects user data.

Visualization:

Web Interface – Displays prediction results in a clear format.

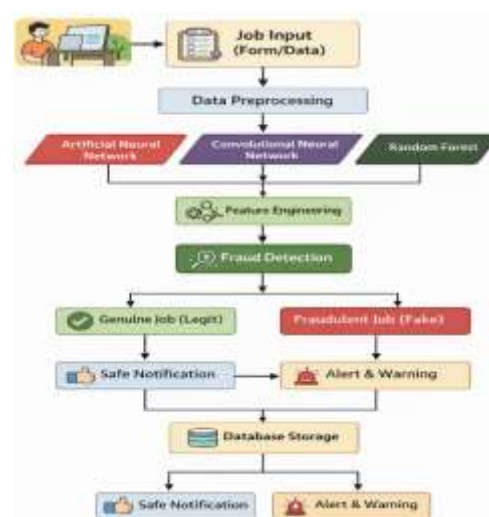


Fig -1: Work-flow

7. Methodology

7.1 Input

Frontend:

HTML, CSS – Used to design a simple and user – friendly

The system begins with user authentication, where the user registers and logs in securely. After login, the user enters job-related details through a web-based form, including job title, salary, location, and company information. The system ensures that all required fields are properly filled for accurate analysis.

The entered data is used to determine whether the job posting is genuine or fraudulent. This simple input method makes the system easy and efficient.

7.2 Processing

Once the job details are submitted, the system processes the input data using preprocessing and feature extraction techniques. The input is cleaned, transformed, and converted into a structured format suitable for analysis. The system identifies important features such as job description, salary, and company details. The processed data is then fed into machine learning and deep learning models including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Random Forest algorithms. These models analyze patterns in the data and classify the job posting as genuine or fraudulent. Based on this analysis, the system generates the final prediction result.

7.3 Output

After processing, the system displays the analysis results on the user interface. The output includes the classification result indicating whether the job posting is genuine or fraudulent, along with the prediction status. In addition, the system provides a clear indication of the fraud risk level to help users understand the reliability of the job posting. The interface may also display model results and relevant details used for prediction. These outputs help users make informed decisions and avoid potential job scams.

8. Conclusion and Future Enhancements

The Online Recruitment Fraud Detection System presents a structured and scalable approach to identifying fraudulent job postings using machine learning and deep learning techniques. By integrating Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Random Forest algorithms with a web-based architecture, the system effectively analyzes job-related data and classifies postings as genuine or fraudulent.

Experimental results demonstrate improved detection accuracy and reliable classification performance. The system reduces dependency on manual verification by providing automated and real-time fraud detection. It enhances user safety by helping job seekers identify fraudulent postings.

Future enhancements may include the integration of advanced deep learning models, real-time data processing, and improved feature extraction techniques. The system can also be extended with explainable AI methods and multi-language support, making it more robust and effective.

9. Acknowledgement

I express my sincere gratitude to my project guide, faculty members, and institution for their valuable support and guidance throughout the development of this project. Their continuous encouragement, guidance, and technical assistance greatly contributed to the successful completion of this work.

10. References

1. Chakraborty, S., Pandey, S., & Ranjan, R. (2022). Deep learning-based job scam detection.
2. Li, Y., Xu, J., & Wang, L. (2021). Fraudulent job post detection using NLP and deep learning.
3. Jindal, N., & Liu, B. (2020). Detecting deceptive job listings using neural networks.
4. Kumar, A., Gupta, H., & Sharma, R. (2019). Hybrid models for job fraud detection.
5. Patil, V., & Mehta, R. (2018). Deep learning for recruitment fraud detection.
6. Brown, R., & Davis, E. (2020). AI-driven fraud detection in job portals.
7. Johnson, A., & Lee, M. (2021). Deep learning models for online recruitment fraud.
8. Wilson, D., & Carter, R. (2018). Neural network approaches for fraud detection.
9. Green, M., & White, S. (2019). Enhancing job security using machine learning.

10. Singh, P., & Verma, A. (2022). Machine learning techniques for cyber fraud detection.
11. Sharma, K., et al. (2023). Classification models for fraud detection systems.
12. Gupta, R., & Jain, S. (2021). Fraud detection using ML.
13. Ahmed, M., et al. (2016). ML fraud detection survey.
14. Ngai, E., et al. (2011). Data mining applications in fraud detection.
15. Phua, C., et al. (2010). Comprehensive survey of fraud detection techniques.
16. Dal Pozzolo, A., et al. (2015). Credit card fraud detection
17. Bahnsen, A., et al. (2016). Cost-sensitive fraud detection models.
18. Carcillo, F., et al. (2019). Machine learning for financial fraud detection.
19. Randhawa, K., et al. (2018). Fraud detection using ensemble learning.
20. Whitrow, C., et al. (2009). Transaction aggregation for fraud detection.
21. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning.
22. Bishop, C. (2006). Pattern Recognition and Machine Learning.
23. Mitchell, T. (1997). Machine Learning.
24. Han, J., et al. (2011). Data Mining Techniques.
25. Géron, A. (2019). Hands-On Machine Learning.



P. Madhuri is a final year, second semester student pursuing Bachelor of Technology in Computer Science and Engineering at Sanketika Vidya Parishad Engineering College. She has actively participated in academic projects and technical activities. Her areas of interest include software development and emerging technologies.



P. Chandrika is a final year, second semester student pursuing Bachelor of Technology in Computer Science and Engineering at Sanketika Vidya Parishad Engineering College. She has been involved in project development and practical learning activities. Her interests include programming and modern computing



M. Prasanth is a final year, second semester student pursuing Bachelor of Technology in Computer Science and Engineering at Sanketika Vidya Parishad Engineering College. He has contributed to academic projects and technical initiatives. His areas of interest include software engineering and application



K. Harika is a final year, second semester student pursuing Bachelor of Technology in Computer Science and Engineering at Sanketika Vidya Parishad Engineering College. She has participated in project-based learning and technical activities. Her interests include programming and emerging technologies.

BIOGRAPHIES



Mr. S. Mano Venkat holds an M.Tech and serves as an Assistant Professor in the CSE Department at Sanketika Vidya Parishad Engineering College. He teaches various core computer science subjects and guides students in academic projects and research activities.