Online Recruitment Fraud (ORF) Detection Using Deep Learning Approaches

[1]P. Narasimha,[2]Peddi Jhansi,[3]Lakkakula Vyshnavi,[4]M. Akhila,[5]Avatari Manasa,[6]Eluru Sruthi

[1]Assistant Professor, Department of Computer Science & Engineering (AI & ML), Princeton Institute of Engineering & Technology For Women

[2,3,4,5,6]B. Tech Students, Department of Computer Science & Engineering (AI & ML), Princeton Institute of Engineering & Technology For Women

ABSTRACT

Online recruitment platforms have become popular channels for job seekers and employers. However, the growth of these platforms has also increased the risk of Online Recruitment Fraud (ORF), where fraudsters post fake job advertisements to steal personal information, financial details, or conduct phishing attacks. Traditional rule-based and machine learning methods struggle to detect sophisticated fraudulent job postings. This research proposes a deep learning-based ORF detection system that analyzes job descriptions, recruiter information, and behavioral patterns to identify fraudulent postings. Natural Language Processing (NLP) and deep neural networks are used to extract contextual patterns from job advertisements. The proposed system improves fraud detection accuracy by learning complex textual and behavioral features. The model assists recruitment platforms in automatically identifying suspicious job posts, thereby protecting job seekers and maintaining platform credibility.

Keywords: Online Recruitment Fraud, Fraud Detection, Deep Learning, Natural Language Processing (NLP), Fake Job Detection, Text Classification, Cybersecurity, Behavioral Analysis, Phishing Detection, Recruitment Platforms.

I. INTRODUCTION

The rapid growth of online job portals has revolutionized recruitment by enabling employers and job seekers to connect easily. Platforms such as job portals, professional networking sites, and recruitment platforms host millions of job advertisements daily. However, this convenience has also created opportunities for cybercriminals to exploit users through Online Recruitment Fraud (ORF). Fraudulent job postings often request application fees, personal data, or banking information under the guise of legitimate employment opportunities.

Detecting such fraud is challenging because scammers continuously modify their strategies and create convincing job descriptions. Traditional detection systems rely on rule-based filtering or basic machine learning techniques that cannot fully capture the complex linguistic and contextual patterns found in fraudulent job posts.

Deep learning techniques, particularly Natural Language Processing (NLP) models and neural

networks, can analyze large volumes of text data and identify subtle patterns in job advertisements. By leveraging deep learning, recruitment platforms can automatically detect suspicious postings and protect users from fraud.

II. LITERATURE SURVEY

1) Secure Transmission of Data Using Image Steganography (2019)

Authors: Sourabh Chandra, Smita Paira

Abstract: Proposes an integrated scheme where a text message is first encrypted using RSA and then embedded into a cover image using steganography techniques. This dual-layer approach ensures both confidentiality and covert communication. The system improves security in network data transmission and protects sensitive information from unauthorized access.

2) Detecting Fraudulent Job Advertisements Using Machine Learning (2018)

Authors: J. B. Kim, M. Kim

Abstract: This research applies machine learning algorithms such as Logistic Regression and Random Forest to identify fake job advertisements. The study analyzes job descriptions, company profiles, and salary information to classify legitimate and fraudulent postings.

3) Online Job Scam Detection Using NLP (2020)

Authors: S. Banerjee, P. Gupta

Abstract: The paper proposes a Natural Language Processing framework to detect recruitment fraud. It extracts textual features from job advertisements and uses classification algorithms to identify suspicious job listings.

4) Deep Learning for Fraud Detection in Online Platforms (2021)

Authors: T. Nguyen, L. Tran

Abstract: This study uses deep neural networks to analyze user behavior and textual content to detect fraud in online systems. The model shows improved performance compared to traditional machine learning approaches.

5) Fake Job Posting Detection Using LSTM Networks (2022)

Authors: R. Sharma, K. Verma

Abstract: The research applies Long Short-Term Memory (LSTM) networks to identify fraudulent job postings by learning sequential patterns in job descriptions and recruiter messages.

III. EXISTING SYSTEM

The existing recruitment fraud detection systems primarily rely on rule-based filtering, keyword matching, and traditional machine learning models such as Support Vector Machines (SVM), Naïve Bayes, and Decision Trees. These approaches typically analyze job postings using predefined rules or simple statistical patterns, where specific keywords, suspicious phrases, or basic feature sets are used to classify postings as genuine or fraudulent. While these methods are easy to implement and computationally efficient, they lack the capability to understand the deeper context and semantics of job descriptions. As a result, they often fail to identify sophisticated fraud attempts where attackers use carefully crafted language to mimic legitimate job postings.

Moreover, these systems suffer from several significant limitations. They exhibit low accuracy when dealing with complex and evolving fraud patterns, as they cannot adapt effectively to new or unseen strategies used by cybercriminals. Their reliance on static rules and limited feature representations makes them incapable of capturing contextual meaning, tone, and intent within textual data. This leads to a high rate of false positives, where legitimate job postings are incorrectly flagged, and false negatives, where fraudulent postings go undetected. Additionally, maintaining such systems requires continuous manual updates of rules and keywords, which is time-consuming and inefficient. Overall, these drawbacks highlight the need for more advanced approaches, such as deep learning and NLP-based systems, to improve detection accuracy and adaptability in modern recruitment platforms.

IV. PROPOSED SYSTEM

The proposed system introduces a deep learning-based framework for detecting Online Recruitment Fraud by analyzing job postings and recruiter information. The system uses Natural Language Processing (NLP) techniques to extract meaningful textual features from job descriptions, company details, and communication patterns. A deep neural network model such as LSTM or CNN is trained on labeled datasets containing legitimate and fraudulent job postings. The model learns complex linguistic structures, contextual relationships, and suspicious patterns commonly found in fraudulent advertisements. Additionally, metadata such as company email domains, salary structures, and posting frequency are analyzed to strengthen detection accuracy. The trained model classifies job postings as legitimate or fraudulent in real time. This automated approach significantly improves detection performance and helps protect job seekers from online recruitment scams.

V. SYSTEM ARCHITECTURE

The diagram illustrates a deep learning and NLP-based framework for detecting online recruitment fraud in job portals. The process begins with job portal data collection, which includes job descriptions, company information, recruiter details, and user reports. This raw data is then passed through a data preprocessing stage, where text is cleaned, tokenized, stop words are removed, and normalization is performed to prepare the data for analysis. After preprocessing, feature extraction is carried out using NLP techniques such as text embeddings along with additional metadata features like email patterns and salary information to capture both textual and contextual insights.

The processed features are then fed into machine learning algorithms like Logistic Regression, SVM, and Random Forest to establish baseline predictions, followed by more advanced deep learning models such as CNN, LSTM, and Bi-LSTM, which are capable of understanding complex language patterns and contextual meanings in job postings. These

models collectively contribute to the fraud prediction module, which classifies job posts as either legitimate or fraudulent. Based on the prediction, the system generates appropriate outputs: legitimate job postings are approved and notified, while fraudulent postings trigger alerts for further action. Overall, the framework provides an intelligent, automated, and scalable solution for identifying fraudulent job advertisements, improving platform security and protecting job seekers.

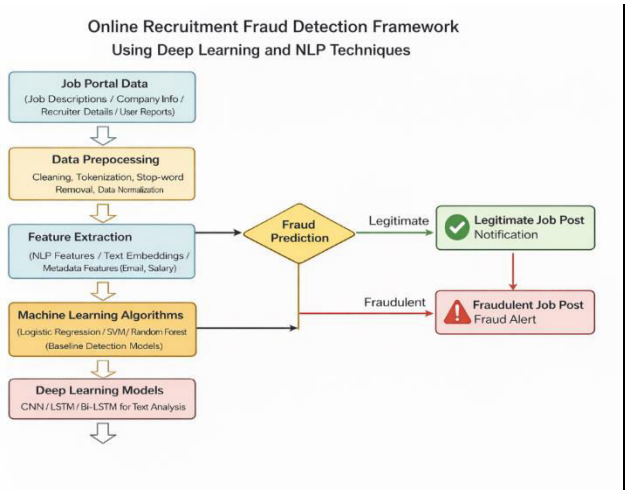

Fig 5.1: System Architecture

VI. IMPLEMENTATION

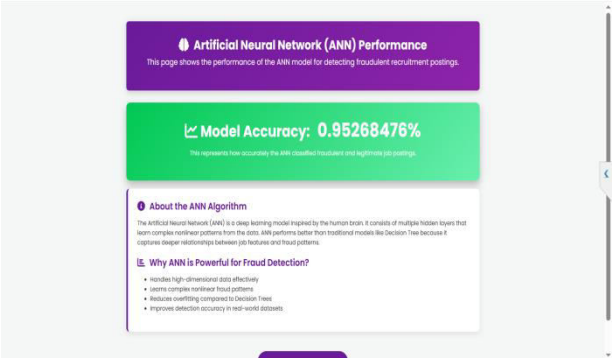

Fig 6.1: Accuracy

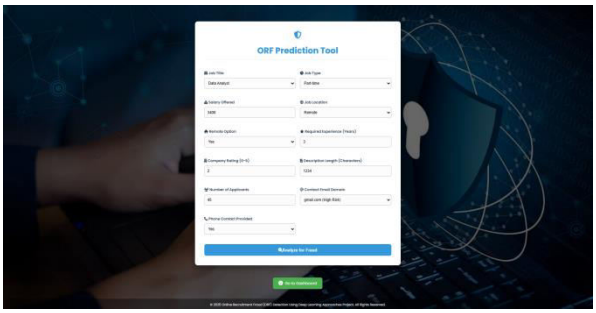

Fig 6.4: Input

VII. CONCLUSION

Online Recruitment Fraud has become a serious issue with the widespread use of online job portals. Fraudulent job advertisements exploit job seekers by requesting sensitive information or financial payments. Traditional rule-based and machine learning methods are not sufficient to detect complex and evolving fraud patterns. This research proposes a deep learning-based detection system that analyzes textual and contextual features from job postings to identify fraudulent advertisements effectively. By integrating Natural Language Processing techniques with deep neural networks, the system can automatically detect suspicious job postings with higher accuracy compared to conventional approaches. The architecture includes modules for data collection, preprocessing, feature extraction, deep learning classification, and alert generation. The proposed approach reduces manual monitoring and helps recruitment platforms maintain trust and credibility. Overall, the system contributes to improving online job portal security and

protecting job seekers from recruitment scams.

VIII. FUTURE SCOPE

The proposed Online Recruitment Fraud detection system can be further improved by incorporating advanced deep learning and artificial intelligence techniques. Future work can focus on integrating transformer-based models such as BERT or GPT-based architectures to enhance the understanding of contextual language patterns in job advertisements. These models can significantly improve the accuracy of fraud detection by analyzing semantic relationships within job descriptions. Another possible extension is the integration of real-time monitoring systems that continuously analyze newly posted job advertisements on recruitment platforms. Behavioral analysis of recruiters, including posting frequency and communication patterns, can also be incorporated to strengthen fraud detection. Additionally, the system can be expanded to include multi-modal data analysis such as images, documents, and email communications used in recruitment processes. Integration with large job portals and government employment platforms can further enhance the system's effectiveness in preventing recruitment scams and ensuring safer online job search environments.

IX. REFERENCES

[1] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, “Botometer: Evaluating social bots on Twitter,” Proceedings of the 26th International Conference on World Wide Web Companion, pp. 273–274, 2017.

[2] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini, “Online human-bot interactions: Detection, estimation, and characterization,” Proceedings of the International AAAI Conference on Web and Social Media, vol. 11, no. 1, pp. 280–289, 2017.

[3] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, “The rise of social bots,” Communications of the ACM, vol. 59, no. 7, pp. 96–104, 2016.

[4] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, “Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?” IEEE Transactions on Dependable and Secure Computing, vol. 9, no. 6, pp. 811–824, 2012.

[5] G. Stringhini, C. Kruegel, and G. Vigna, “Detecting spammers on social networks,” Proceedings of the 26th Annual Computer Security Applications Conference, pp. 1–9, 2010.

[6] K. Lee, B. Eoff, and J. Caverlee, “Seven months with the devils: A long-term study of content polluters on Twitter,” Proceedings of the International AAAI Conference on Web and Social Media, pp. 185–192, 2011.

[7] K. C. Yang, O. Varol, P. Hui, and F. Menczer, “Scalable and generalizable social bot detection through data selection,” Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 1, pp. 1096–1103, 2020.

[8] E. Ferrara, “Disinformation and social bot operations in the run up to the 2017 French presidential election,” First Monday, vol. 22, no. 8, 2017.

[9] S. Kudugunta and E. Ferrara, “Deep neural networks for bot detection,” Information Sciences, vol. 467, pp. 312–322, 2018.