iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81536

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 14 Issue IV Apr 2026- Available at www.ijraset.com

Advanced Fake Online Recruitment Post Detection Using Machine Learning

Y. Anusha[1], N. Thrishalini[2], P. Madhu Latha[3], T. Geetha Sai[4], M. Susanna[5]
[1]*Assistant Professor, Department of Computer Science and Engineering, Bapatla Women's Engineering College, Bapatla, AP, INDIA*
[2,3,4,5]*B.Tech, Computer Science and Engineering, Bapatla Women's Engineering College, Bapatla, AP, INDIA*

Abstract: ***The Rapid Growth of Online Recruitment Platforms has made jobs searching easier, but it has also lead to a significant increase in fraudulent job postings. This fake job advertisements often mislead jobseekers, causing financial loss, misuse of personal information, wasted time. Identifying such deceptive postings manually is challenging, as they closely resemble genuine job listing and require continuous monitoring.***
To address this issue, this project proposes an automated system for detecting fake online recruitment post using Machine Learning techniques. The system uses Classification algorithms such as Decision Tree and Random Forest to analyse job related data, including job Descriptions, Company Profiles, and requirements. Though data preprocessing and feature Extraction, the model learns patterns that distinguish real job postings from fraudulent ones. Experimental analysis shows that ensemble methods improve detection accuracy compared to single classifiers.
The proposed system provides real-time prediction of job authenticity and also offers additional company information to enhance transparency. This approach not only improves detection efficiency but also helps job seekers make safer decisions. Overall, the system demonstrates how machine learning can be effectively applied to reduce online recruitment fraud and create a more secure job such environment.

I. INTRODUCTION

The fields of computer science, data analytics, and Machine Learning have advanced rapidly, enabling the development of intelligent systems for real world problems. One major issue in today's digital era is the rise of fraudulent job postings on online recruitment platforms. As job seekers increasingly depends on these platforms, there often exposed to fake job advertisements that appear genuine but are intended to mislead and exploit users.

The main idea of the project is to develop an intelligent system that can automatically detect fake job postings on online recruitment platforms. many fraudulent job advertisements are designed to look genuine, making it difficult for users to identify them. To overcome this problem, our project uses Machine Learning techniques to analyse job related and classify whether a job post is real or fake. The system works by taking inputs such job description, company profile, recruitments and other details like URL of the post from job posting. This data is first processed using data preprocessing techniques to remove noise and convert it into a structured format. Then, feature extraction methods are applied to identify important patterns in the data. After that, Machine Learning Algorithms like Decision tree and Random Forest are used to train the model and predictions based on the learned patterns.

Finally, the system provides an output indicating whether the job post is genuine or fraudulent and display Company's linked In and official website links in the output screen. This helps users make better decisions, avoid scams, and improves the overall safety of online jobs searching. Thus, the project combines data processing and Machine Learning to provide an accurate and automated solution for fake job detection.

The system applies techniques to remove noise and convert the data into a structured format. Then, feature extraction methods are used to identify important patterns in the data.

After that, machine learning algorithms like decision tree and random forest are used to train the model and make predictions based on the learned patterns.

Finally, the system provides an output indicating whether the job post is genuine or fraudulent and displays company LinkedIn and official website links on the output screen.

This helps users make better decisions, avoid scams, and improves the overall safety of online job searching. Thus, the project combines data processing and machine learning to provide an accurate and automated solution for fake job detection.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

A comparative study by Rakesh S. Ghosh et al. (2020) analysed different Machine Learning Algorithms including Naïve Bayes, and Gradient Boosting, for fake job detection. The study of evaluated model performance using various matrix and emphasized that proper feature selection and data preprocessing significantly improve detection accuracy.[1]

Detection of fake job postings using Machine Learning has gained significant attention in recent years. Mohamed A. H. Abdelrahman et al. (2021) presented a comprehensive study on identifying fraudulent job advertisement using algorithms such as SVM, Neural Networks. There work highlights the importance of feature extraction techniques and discusses challenges in building robust models for real world applications.[2]

Furthermore, Haritha N. Bhat et al. (2022) Provided a survey on various machine learning approaches for detecting recruitment frauds. The study reviewed both supervised and unsupervised learning methods and highlighted the role of future engineering in improving model performance. It also suggested feature directions for developing more accurate and scalable fraud detection systems.

III. EXISTING SYSTEM

The existing system mainly relies on manual verification of job postings by users or administrators. It uses basic keyword Based filtering techniques to identify suspicious job advertisements. Detection is often based on user complaints rather than automated real time analyses, making it inefficient and less accurate and uses some in accurate algorithms to detect fraudulent jobs.

IV. PROPOSED SYSTEM

The proposed system presents an intelligent and automated framework for detecting fraudulent job postings using machine learning techniques. The architecture is designed to process job-related information, extract meaningful features, and accurately classify job advertisements as genuine or fake. Unlike traditional manual verification methods, the proposed approach leverages data-driven learning to ensure scalability, efficiency, and improved detection accuracy.

1) Fake Job Detection Workflow: The proposed fake job detection system is designed to automatically identify whether a job posting is genuine or fraudulent using advanced machine learning techniques. The system follows a structured workflow in which job-related data is collected, processed, and analysed to generate accurate predictions. Unlike traditional manual verification methods, this automated approach reduces human effort and improves detection efficiency. The workflow ensures that every stage, from input to output, is systematically handled to achieve reliable results.

2) Input Stage: In this stage, the system accepts job-related details from the user through a user-friendly interface. The input includes attributes such as job title, company name, location, Employment type, Experience, Education, Industry and job description. Thers details can be manually enter by the user which is collected from job portals the quality and completeness of input data play a crucial role in improving the accuracy of the prediction model.

3) Data Pre -Processing: The collected if often unstructured and may contain missing values, noise, and irrelevant information. Therefore, preprocessing is performed to clean and organise the data. This include removing unwanted columns, handling missing values, eliminating duplicates and correcting inconsistencies. Text normalization techniques such as converting text to lowercase, removing punctuation, stop words, and special characters are applied. This step ensures that the dataset is consistent, clean, and suitable for further processing. The whole process can be done in backend.

4) Feature Extraction: After preprocessing, the cleaned textual data is converted into a numerical representation using feature extraction techniques such as TF-IDF (Term Frequency-inverse Document Frequency). This method identifies the importance of words in a document relative to the entire dataset. By transforming text into feature vectors. the system can effectively capture patterns and relationships in job postings, enabling the machine learning modes to perform accurate classification.

5) Applying Machine Learning Algorithms: The system employs supervised machine learning algorithms such as Decision Tree and Random Forest for classification. These algorithms are capable of learning pattern from labelled data and distinguishing between real and fake job postings. Random Forest, being an method, enhances performance by multiple decision tree, there by improving accuracy and reducing overfitting.

6) Model Training: The dataset is divided into training and testing subsets to build and evaluate the model. During the training phase, the model learns the relationship between input features and output labels. The training process enables the models to generalize patterns effectively, which helps in making accurate predictions on unseen data.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

7) Prediction: Once the model is trained, it is used to predict the authenticity of new job postings. When a user submits job details, the system process enables the model to generalize patterns effectively, which helps in making accurate predictions on unseen data.
8) Evaluation: The system collection job posting data from various sources, including online job portals, datasets, and company career pages. The collected data consist of both genuine and fraudulent job posting along with relevant attributes such as job descriptions, company details, and requirements. This data forms the foundation for training the machine learning model.

A. *Data Collection*

The system collects job posting data from various sources, including online job portals, datasets, and company career pages. The collected data consists of both genuine and fraudulent job posting along with relevant attributes such as job descriptions, company details and requirements. This data forms the foundation for training the machine learning model.

B. *Data Pre -Processing*

In this phase, the collected data is cleaned and structured to improve its quality. Missing values are handled appropriately, irrelevant attributes are removed, and text data is normalized. These preprocessing steps ensure that the data is consistent and ready for feature extraction and model training.

C. *Applying Algorithm*

Machine learning algorithms such as Decision Tree and Random Forest are applied to classify job postings. These algorithms analyse patterns in the dataset and learn to differentiate between genuine and fake job advertisements. The use of ensemble methods further improves the robustness and accuracy of the system.

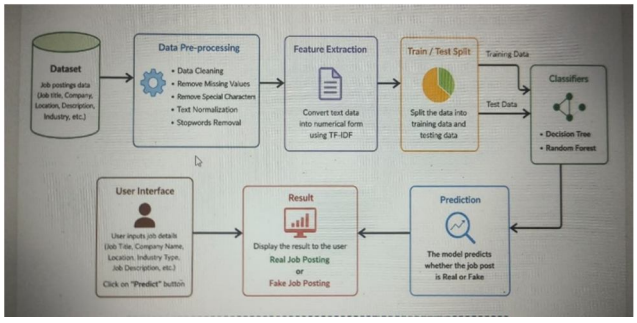
D. *Evaluation*

The system evaluates the performance of different models using metrics such as accuracy, precision, recall, and F1- score. Based on these evaluations, the best-performing model is selected for final prediction. This ensures that the system delivers accurate and reliable result in real-world scenarios.

V. MODULES

1) Input Module: Collects job Posting Details from the user such as job title, company name, and other requirement details. This acts as the initial data input for the system.
2) Data Preprocessing Module: Processes the collected data by removing unnecessary values, handlings missing data, and converting text into a suitable encoded format for machine learning.
3) 3.Feature Extraction Module: Transforms text data into numerical features using techniques like TF-IDF or bag of words, which helps the model to understand the input data.
4) Machine Learning Prediction Module: Applies Decision Tree and Random Forest algorithms to analyse the input data and classify job postings as fake or real.
5) 5.Company Information Module: Displays relevant company details and job-related information to provide better understanding to the user.
6) 6. Result Display Module: Shows the final Prediction result (Fake or Real job) clearly to the user with necessary details.

VI. SYSTEM ARCHITECTURE



©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VII. RESULT AND ANALYSIS

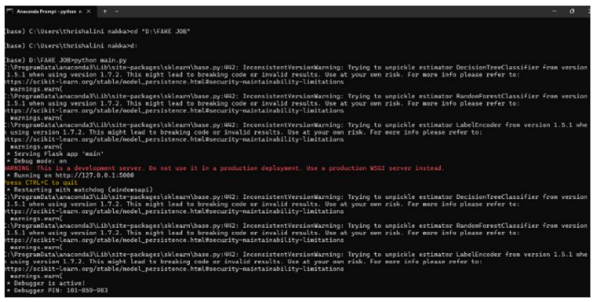


Fig 1: command prompt

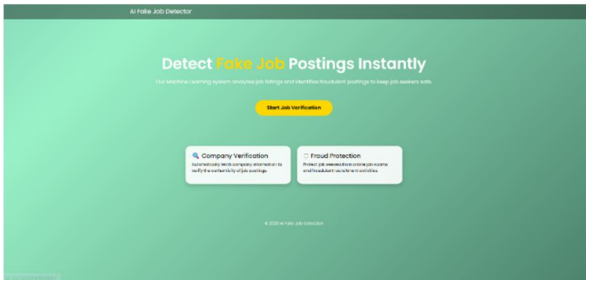


Fig 2: Home Page

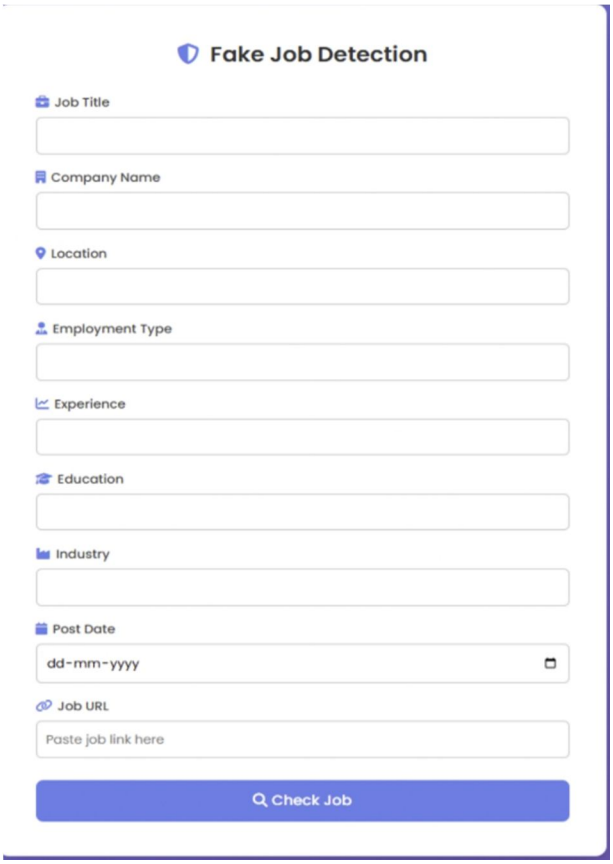


Fig 3: Input interface

Result for Real Job Posting details

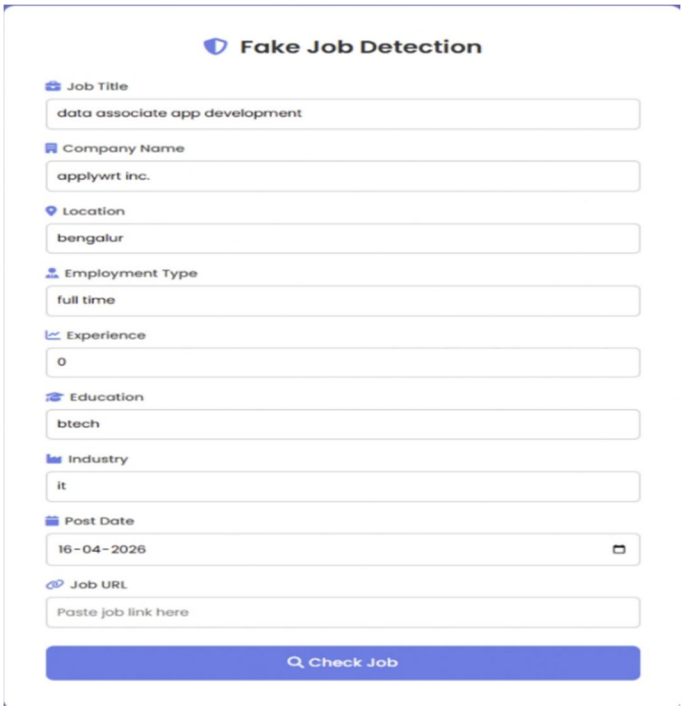


Fig 4: Input Page with job post details

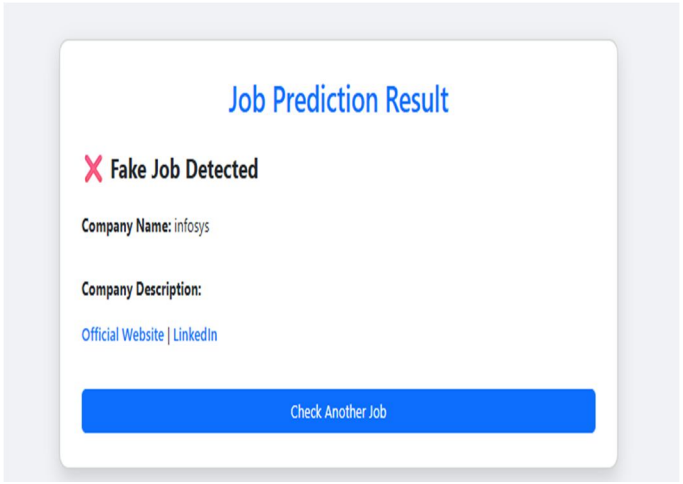


Fig 7 : Result Page

VIII. CONCLUSION

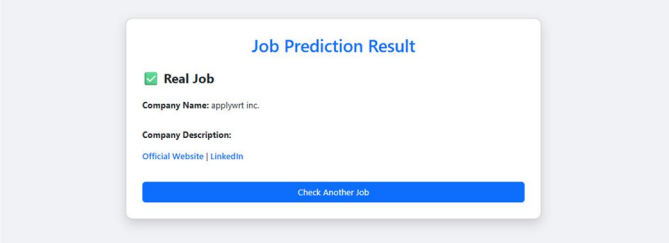


Fig 5: Result Page

Result for Fake Job Posting Details

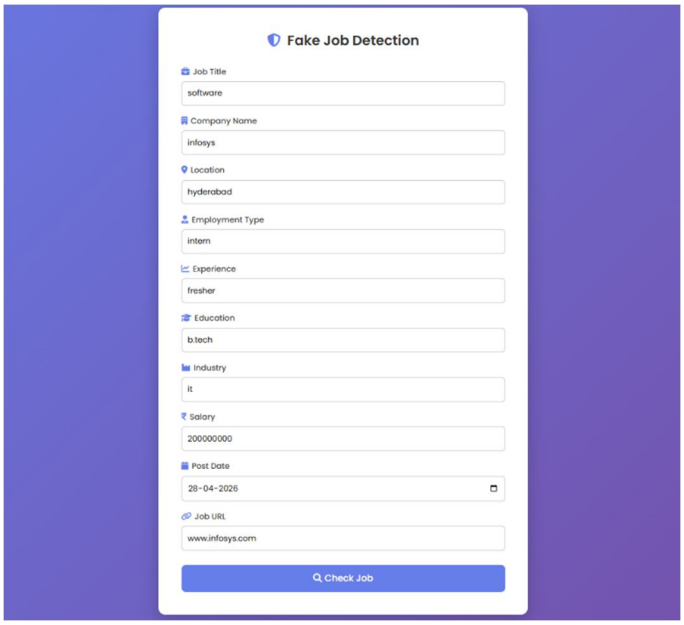


Fig 6 : Input interface with details

The fake job Detection System is developed to identify fraudulent job postings and help users avoid online job scams. In today's digital world, many fake job advertisements mislead job seekers by providing false information about companies, salaries, and opportunities. this system aims to solve that problem by using machine learning techniques. The system collects job-related data, preprocesses it, and converts it into meaningful features. By applying algorithms like decision Tree and Random Forest, the system analyses patterns in job postings and classifies them as real or fake. Among these, Random Forest generally provides better accuracy due to its ensemble learning approach. the use data preprocessing and feature extraction technique improves the efficiency and accuracy of the model. The system also provides a user -friendly interface where users can easily input job details and get instant results. This makes the solution practical and useful in real-world scenarios.

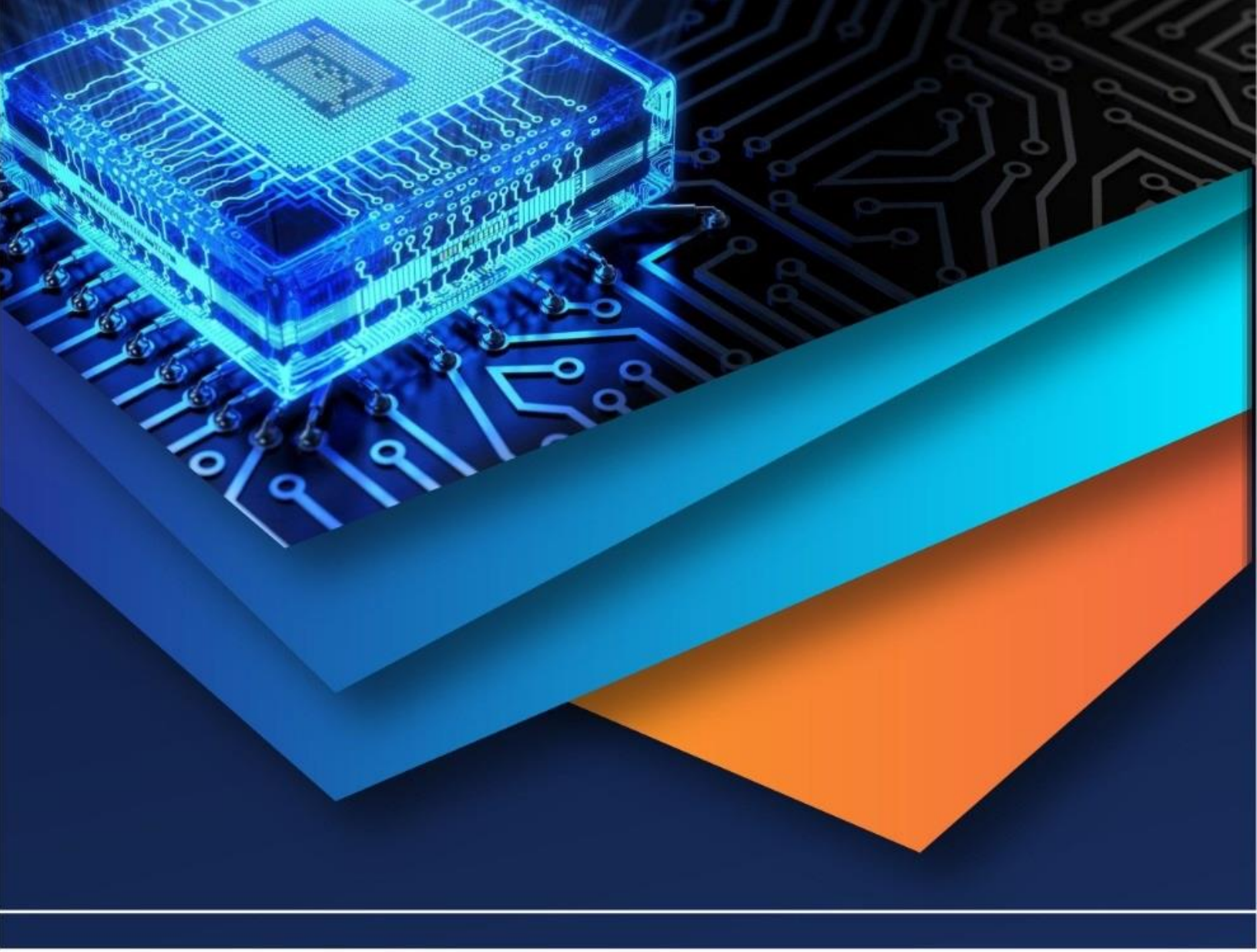
Overall, the proposed system helps in increasing awareness among job seekers and reduces the risk of falling into fraudulent job traps. It saves time, ensures safety, and builds trust in online job platforms. The system can be further enhanced by using advanced deep learning techniques and larger datasets to improve accuracy and performance.

IX. FUTURE SCOPE

The system can be further enhanced by using advanced machine learning and deep learning techniques to improve prediction accuracy. Training the model with larger and updated datasets can help in identifying more complex and new fraud patterns.In the future, this Systems can be integrated with real-time job portals for automatic detection of fake job postings. Additional
Features like user alerts, mobile application support, and multi-language options can improve usability and accessibility.
These improvements will make the system more efficient, reliable, and widely usable for job seekers across different platforms

REFERENCES

[1] Abdelrahman, M. A. H., El-Khamlichi, W. M., and EL-Gohary, A. F., "Detection of Fake Job Postings using Machine Learning Techniques", International journal of computer science and applications, 2021.
[2] Ghosh, R. S., Roy, S. K., and Sharma, A., "Fake job advertisements Detection: A Comparative study of Machine Learning Algorithms," International Journal of advanced research in computer science, 2020.
[3] Rao, S. S., Tiwari, N., ad Joshi, P., "Leveraging Text mining and machine learning for identifying fake job posts," international journal of data mining and knowledge engineering, 2019.
[4] Bhat, H. N., Mallya, V. S., and Kumar, R. T., "A Survey of Machine Learning Approaches for detecting job recruitments frauds," international journal of artificial intelligence research, 2022.



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)