

# FAKE JOB DEFECTION SYSTEM USING Llama-3 AND MACHINE LEARNING

T.Amalraj Victoire<sup>1\*</sup>, A. Ashika<sup>2\*\*</sup>

(Master of computer application, Sri Manakula Vinayagar Engineering College, Pondicherry.  
Email: [amalrajvictoire@gmail.com](mailto:amalrajvictoire@gmail.com))

(Master of computer application, Sri Manakula Vinayagar Engineering College, Pondicherry.  
Email: [ashikaanandhan16@gmail.com](mailto:ashikaanandhan16@gmail.com))

\*\*\*\*\*

## Abstract:

With the evolution of online recruiting websites, fake job postings have become rampant leading to money extortion and cyber-crimes. Traditional methods for the detection of fake job postings are unable to recognize any semantic and hidden patterns in frauds. In this paper, a smart Fake Job Detection System has been proposed using the LLaMA-3 large language model in combination with the Machine Learning algorithms and NLP techniques. This system will detect fraudulent jobs by analyzing details like the description of the job, recruiter details, company details, salary offered. The framework has been implemented and experimentally analysed.

**Keywords:** *Fake Job Detection, LLaMA-3, Machine Learning, NLP, Fraud Detection, Artificial Intelligence.*

\*\*\*\*\*

## 1. Introduction

The rise in online recruitment systems has made the hiring process easier due to the availability of job openings. The rising trend of online job boards has also led to an increase in fake job postings and job frauds. Fake recruiters use deceptive techniques such as false salary promises, registration fees, phishing links, and fabricated company details to defraud job hunters.

Traditional approaches to detecting such fraud rely on human verification or keyword search techniques which are not able to detect sophisticated scams. Advanced technology like Artificial Intelligence and Natural Language Processing can offer more effective ways of detection through the use of semantics and context analysis.

This project suggests the development of a Fake Job Detection System using LLaMA-3 transformer in combination with machine learning techniques.

## 2. Literature Review

A number of scholars have applied machine learning/deep learning to detect job fraud.

Dutta and Bandyopadhyay created a machine learning based fraud detection system using Logistic Regression and Random Forest classifiers. The model had acceptable classification performance but did not have semantics capability. Devlin et al. proposed transformers-based BERT models that could do deep contextual language processing, thus outperforming traditional NLP classification.

The recent work on transformer models using LLaMA shows enhanced semantic and contextual modelling which is useful in job fraud detection.

## 3. Proposed System

The proposed system employs the usage of LLaMA-3 and Machine Learning algorithms to spot any cases of employment fraud based on the meaning and context of the post.

### Modules Used

- User Login and Registration
- Employment Posting
- Data Preprocessing
- Feature Extraction

- Semantic Analysis with LLaMA-3
- Classification of Fraud
- Generation of Risk Scores
- Alert System

It identifies patterns that include exaggerated job salaries, manipulation through words, and strange hiring practices.

#### 4. Methodology

##### 4.1 Data Collection

The data set is composed of genuine and fraudulent job listings sourced from online hiring websites. This includes job postings, salaries, company descriptions, and employment information.

##### 4.2 Preprocessing

Preprocessing of textual data is done using different methods.

###### Methods Used:

- Tokenization
- Stop Words Removal
- Converting to Lower Case
- Noise Filtering

##### 4.3 Feature Extraction

Data is then transformed into word embeddings through the LLaMA-3 tokenizer. These embeddings include semantics between words.

##### 4.4 Semantic Analysis

Through its self-attention mechanism, the transformer conducts contextual analysis and detects concealed fraud markers.

##### 4.5 Classification

Machine learning models classify job posts as either genuine or fraudulent using extracted features and semantic analysis.

#### 5. System Architecture

The architecture of the proposed system includes the following:

- Frontend Interface
- Flask Backend

- NLP Processing Component
- LLaMA-3 Model
- Classification Engine
- Database System
- Notification Component

The system analyses job data serially and gives instantaneous results for fraud detection.

#### 6. Results and Analysis

The system has achieved success in detecting fake job postings by applying semantics and machine learning classification techniques.

##### Tests Carried Out

- Unit Tests
- Integration Tests
- Functional Tests
- Security Tests

##### Performance Analysis

Parameters	Results
Accuracy in Fraud Detection	High
Response Time	Fast
Security	Successful
Alerting	Real-Time
Stability	Stable

From the above results, we have found out that there is an improvement in the accuracy of detecting fake job postings using transformer semantic analysis.

#### 7. Merits of the System

- Reliable fraud detection
- Effective job analysis in real time
- Better recruitment security
- Cybercrime prevention
- Automatic alerts
- Fast semantic comprehension

## **8. Conclusion**

The above suggested Fake Job Detection System utilizing LLaMA-3 and Machine Learning offers a smart approach to detecting fake jobs on online recruitment sites. Semantic analysis, NLP, and machine learning have been used in the design to ensure that the classification is reliable.

This helps to improve the level of security in recruiting since it enables identification of any suspicious job posts at once and sends instant alerts to the concerned persons.

## **9. Future Enhancements**

Future developments may include:

- Multilingual capabilities
- Development of mobile applications
- Integration of Explainable AI
- Deployment on the cloud
- Web scraping in real time

Such advancements would be helpful to further enhance the scalability and efficiency of the fraud detection system.

## **10. References**

1. Dutta D, Bandyopadhyay S. "Detection of fake job postings using machine learning techniques," ICACCI, 2019.
2. Devlin J, Chang M, Lee K, Toutanova K. "BERT: Pre-training of deep bidirectional transformers for language understanding," NAACL, 2019.
3. LeCun Y, Bengio Y, Hinton G. "Deep Learning." Nature, 2015.
4. Meta AI. LLaMA: Large Language Mo