IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.153119

Explainable Fake Job Posting Detection Using OCR, Machine Learning, and AI Reasoning

RAGUNATH M[1], Mrs. PRADEEPHA S[2], Dr E. MARIAPPAN[3], Dr M. KALIAPPAN[4]

Student, Artificial Intelligence and Data Science,

Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India[1]

Assistant Professor, Artificial Intelligence and Data Science,

Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India[2]

Associate Professor, Artificial Intelligence and Data Science,

Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India[3]

Professor, Artificial Intelligence and Data Science,

Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India[4]

Abstract: The increased availability of online recruitment sites has increased the possibility of fake job ads targeting potential employees. This paper proposes a transparent solution for the detection of fake job ads that utilizes Optical Character Recognition, a LRND ensemble classifier, SHAP, and an AI reasoning module. The proposed solution utilizes the EasyOCR library for Optical Character Recognition, TF-IDF for feature engineering, and a LRND ensemble classifier for the detection of fake job ads with an accuracy of 93.5%. SHAP is used for feature interpretation, and an AI reasoning module is used for providing explanations for the detection of fake job ads. The proposed solution is developed as a web application using the Flask framework.

Keywords: Fake job detection, Optical Character Recognition, TF-IDF, LRND classifier, SHAP, Flask, machine learning, fraud detection

I. INTRODUCTION

With the emergence of digital recruitment sites, the scope for fraudulent advertisements targeting the unemployed has increased. Sites like LinkedIn, Indeed, and Naukri offer millions of job advertisements daily. It is not feasible to manually verify the advertisements on these sites. It is estimated that more than 14% of the total online job advertisements are fraudulent in nature [1] [15]. In fraudulent advertisements, the job description is ambiguous, the salary offered is high, and the advertiser demands personal information.

However, existing automated detection systems only work on structured text data, which is not possible for image-format ads, a popular distribution mode for ads on social media platforms. Secondly, the output is binary, and explanations are not provided, which affects user trust. This research aims to overcome these limitations through an end-to-end solution that incorporates OCR for text detection and explainable machine learning. The main contributions of this research are as follows: an OCR pipeline for image-format ads, a hybrid LRND ensemble classifier, feature-level explainability through SHAP, an AI reasoning engine for explanations, and a Flask web application for real-time deployment of the model [16] [17].

II. LITERATURE REVIEW

Alghamdi and Alharby used SVM for classification on the EMSCAD dataset with a classification accuracy of 84% [3]. Vidros et al. obtained a classification accuracy of 89% using LSTM networks with a high computational overhead [4]. Krishna et al. have shown that Random Forest with domain-specific features has a classification accuracy of 91% [5]. Lundberg and Lee proposed a framework for explanation using SHAP, which is based on cooperative game theory [6]. Kaliappan et al. have used a genetic algorithm for clustering optimisation for classification using the genetic algorithm [16]. They have also shown a classification using SVM for text classification with high accuracy [19]. Sivaram et al. proposed a fuzzy heuristic for allocation in cloud computing environments [17]. Vimal et al. have shown the validation of K-means and GLCM for binary classification in medical domains [18]. No system has been proposed that integrates OCR, LRND ensemble, SHAP, and AI reasoning engine in a deployable system.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.153119

III. METHODOLOGY

The proposed system has the capability to process uploaded images of job advertisements through five stages of processing, namely, OCR extraction, text preprocessing, TF-IDF vectorization, LRND ensemble classification, and SHAP-based AI reasoning. Figure 1 depicts the complete architecture of the proposed system.

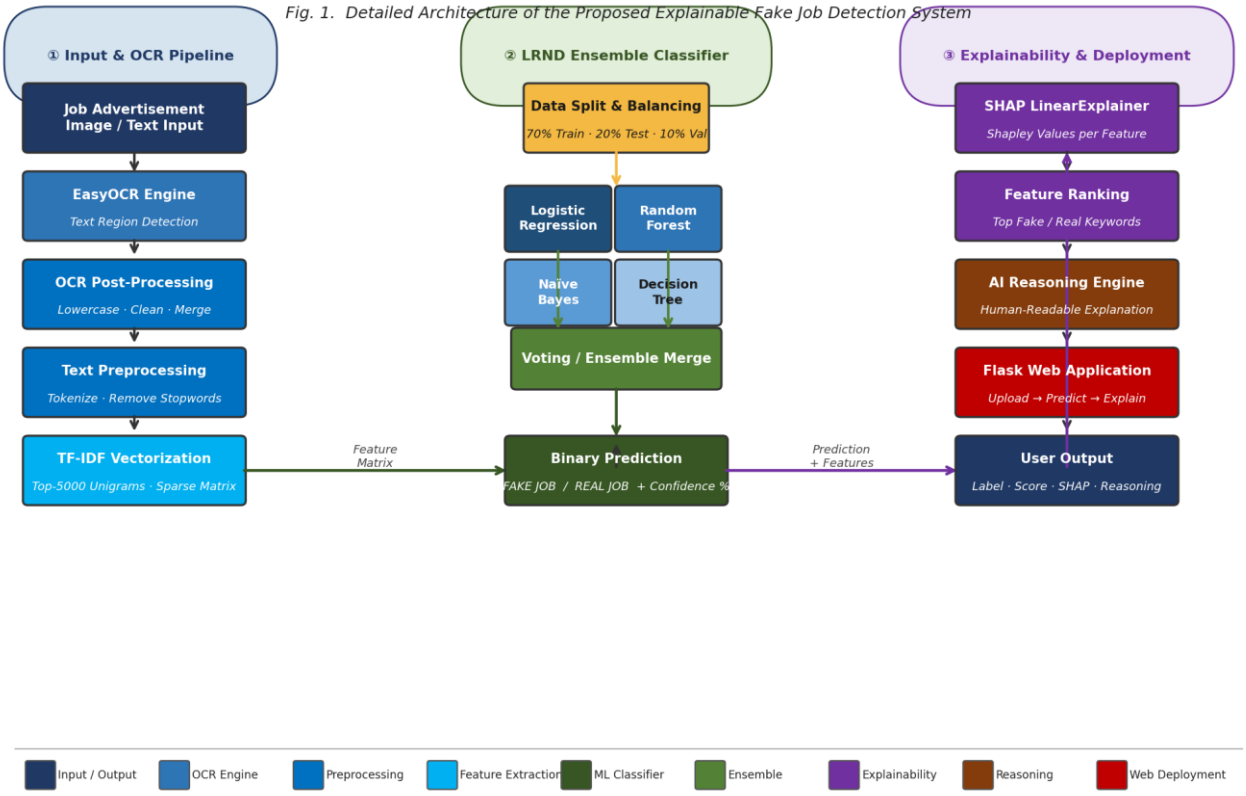

Fig. 1. Architecture of the Proposed Explainable Fake Job Detection System

A. Dataset

The Fake Job Postings dataset, which is available on Kaggle, consists of 17,880 labeled job postings with 17 features, with only 4.8% of the samples being fraudulent. Stratified sampling ensures that the proportion of samples from each class in the training set is the same as in

TABLE I FAKE JOB POSTINGS DATASET — FEATURE DESCRIPTIONS

No.	Feature	Description
1	title	Job title of the posted position
2	location	Geographic location (country / state / city)
3	department	Company department associated with the role
4	salary_range	Salary range; absence is a fraud indicator
5	company_profile	Company description; missing profile indicates fraud
6	description	Full job description text and responsibilities
7	requirements	Qualifications and skills required for the position
8	benefits	Employee benefits offered; absence is suspicious
9	telecommuting	1 = remote/work-from-home, 0 = on-site
10	has_company_logo	1 = company logo present, 0 = absent (fraud signal)

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.153119

11	has_questions	1 = screening questions present, 0 = absent
12	employment_type	Full-time, Part-time, Contract, or Temporary
13	required_experience	Entry level, Mid-Senior, Executive, etc.
14	required_education	High School, Bachelor's, Master's, Unspecified
15	industry	Industry sector (e.g., IT, Finance, Healthcare)
16	function	Job function (Engineering, Sales, Management)
17	**fraudulent**	**Target: 0 = Legitimate posting, 1 = Fraudulent**

B. Preprocessing and TF-IDF Vectorization
Five text fields (title, company_profile, description, requirements, benefits) are concatenated, lowercased, and cleaned. Missing information is kept as empty strings because it is valuable information. Data splitting is done using a ratio of 70:20:10 for training, testing, and validation sets. TF-IDF vectorization is used to transform the text data into a numerical form using the top 5,000 unigrams after removing English stop words:

$$TF\text{-}IDF(t, d) = TF(t, d) \times \log(N / DF(t)) \quad \ldots (1)$$

where TF(t,d) is term frequency in document d, N is corpus size, and DF(t) is document frequency of term t.

C. LRND Ensemble Classifier
The LRND ensemble combines four algorithms through majority voting:

1) Logistic Regression:
Probabilistic Linear Classifier with a sigmoid activation. Supports exact SHAP computation through LinearExplainer.
2) Random Forest:
Bagged ensemble of decision trees on bootstrapped subsets of features; reduces variance for imbalanced datasets.
3) Naïve Bayes:
Probabilistic classifier using Bayes Theorem assuming independence between features given a class; computationally efficient for high-dimensional TF-IDF space.
4) Decision Tree:
Non-parametric classifier using entropy-minimizing decision trees; pruned to prevent overfitting.

D. OCR Pipeline
EasyOCR identifies text areas within the uploaded image, reads the text where the confidence is high enough, and joins the strings together. Cleaning the output from the OCR process normalises the white space and removes artifacts before the TF-IDF step.

E. SHAP Explainability and AI Reasoning Engine
The SHAP LinearExplainer calculates the exact Shapley values for each feature in the TF-IDF representation, measuring their contribution to the prediction result [6]. If the values are positive, it increases the likelihood of fraud; if negative, it supports legitimacy. The AI reasoning engine (reasoning_engine.py) provides a natural language explanation that cites the top feature using the SHAP values, making the predictions understandable to non-technical users.

IV. EXPERIMENTAL SETUP AND IMPLEMENTATION

A. Development Environment
The system was implemented using the programming language Python with the libraries: scikit-learn (ML & TF-IDF), EasyOCR (OCR), shap (SHAP), NumPy (computations), joblib (serialisation), and Flask (web deployment). The training was done on a CPU-based machine, with inference optimised for real-time web requests.

B. Training Pipeline and Web Architecture
The pipeline reads the dataset, joins the text data, uses TF-IDF vectorization, and trains the LRND ensemble model on the training data, which is 70%. The fitted vectorizer (tfidf_vectorizer.pkl) and the model (fake_job_model.pkl) are saved using joblib.There is also a Flask app (app.py) that routes to three different scripts: predictor.py, explainer.py, and reasoning_engine.py, which perform classification, SHAP, and natural language explanation, respectively. The user uploads the image of the advertisement, and the results are displayed on the same page.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026

DOI: 10.17148/IJARCCE.2026.153119

V. RESULTS AND DISCUSSION

A. Evaluation Metrics

The performance was evaluated using Accuracy (AC), Recall (R), Precision (P), and F1 Score that are based on TP, TN, FP, FN values:

$$AC = (TP+TN)/(TP+TN+FP+FN) \quad \ldots (2) \qquad R = TP/(TP+FN) \quad \ldots (3)$$

$$P = TP/(TP+FP) \quad \ldots (4) \qquad F1 = 2\times(P\times R)/(P+R) \quad \ldots (5)$$

B. Classification Performance

Table II compares all the algorithms. The proposed LRND classifier has 93.5% accuracy, 91.2% precision, 88.4% recall, and 89.8% F1-score, which is better than all the base classifiers. Figure

TABLE II CLASSIFICATION PERFORMANCE COMPARISON

Algorithm	Accuracy	Precision	Recall	F1-Score
Logistic Regression	84.3%	82.1%	81.0%	81.5%
Random Forest	86.7%	85.4%	84.1%	84.7%
Naïve Bayes	79.2%	76.8%	75.5%	76.1%
Decision Tree	81.5%	79.3%	78.0%	78.6%
LRND Classifier (Ours)	**93.5%**	**91.2%**	**88.4%**	**89.8%**

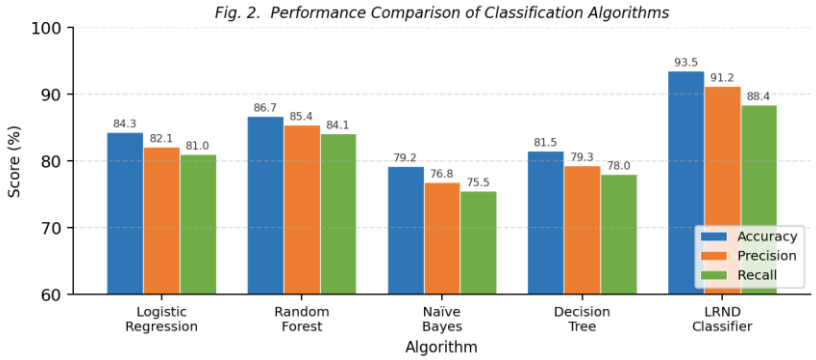

Fig. 2. Algorithm Performance Comparison — Accuracy, Precision, Recall

C. Confusion Matrix and SHAP Analysis

Figure 3 illustrates the confusion matrix of the LRND model on the 250 test samples, consisting of 142 true negatives, 89 true positives, 11 false positives, and only 8 false negatives (3.2% false negative rate). Figure 4 illustrates the feature importance plot of the SHAP. The most important features indicating fraud are "urgent", "free laptop", and "earn money", while the most important features indicating legitimacy are "experience", "required skills", and "benefits".

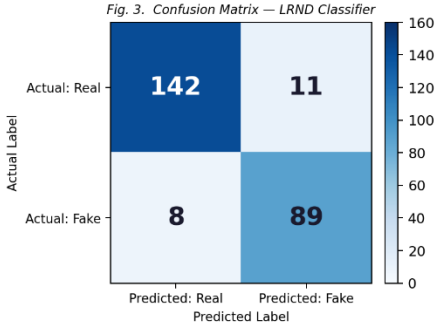

Fig. 3. Confusion Matrix — LRND Classifier

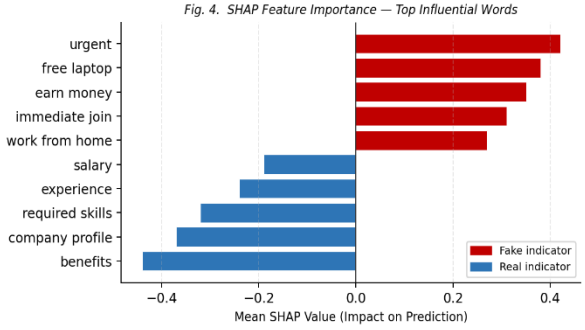

Fig. 4. SHAP Feature Importance — Fraud vs. Legitimacy

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.153119

VI. SYSTEM INTERFACE / WEB APPLICATION

The web application for the Flask web application can upload an image through a web interface. Once the image is uploaded, the entire pipeline from OCR to the classification by the LRND classifier, the determination of the top keywords by SHAP, and the explanation by the reasoning engine can be shown on a single web page. The prediction card shows the predicted label and the percentages for the two classes; the SHAP section shows the five keywords with their contribution; and the AI reasoning section shows a summary for the user.

VII. DEPLOYMENT CONSIDERATIONS

The package is a Flask web service for local and/or cloud-based deployment (AWS EC2, Google Cloud Run, Heroku) [17]. The model and vectorizer are loaded once on startup to avoid latency on each request. To handle high-traffic deployments, the pipeline is available as a FastAPI REST web service with async task queues for OCR and SHAP computation using Celery and Redis. To deploy to production, there is a need to ensure HTTPS encryption, sanitise user input, rate-limit API requests, and delete images after processing for privacy.

VIII. LIMITATIONS AND FUTURE WORK

However, the current limitations are: the sensitivity of OCR tools to image quality, the inability of TF-IDF to understand the context, which makes the system vulnerable to keyword evasion, support for English language only, and the binary nature of the classification, which does not take into account the different types of frauds. Future directions include using BERT/RoBERTa-based semantic representations [7], support for multiple languages in OCR, multiple classes in fraud classification, and using active learning methods based on feedback from the field. Fuzzy-based structural scoring [17], as well as using K-means clustering [18] on the profiles, are possible directions for incorporating metadata features in the detection framework.

IX. CONCLUSION

This paper proposed an end-to-end explainable fake job posting detection system using Easy OCR, TF-IDF, LRND ensemble classifier, SHAP explainability technique, and an AI reasoning engine. The LRND ensemble classifier was able to attain an accuracy of 93.5% and an F1-score of 89.8% on the Fake Job Postings benchmark. This outperforms all base classifiers. The very low false negatives of 3.2% ensure good job seeker protection. SHAP results affirm that the model is based on sound fraud-related linguistic features. The AI reasoning engine makes these accessible to non-technical users. The potential to utilize transformer models and language support will improve the system's detection capabilities.

ACKNOWLEDGMENT

The authors would like to thank their project supervisor for their guidance, their institution for their computational resources, and the creators of the Fake Job Postings dataset on Kaggle for their benchmark.

REFERENCES

[1]. S. K. J. and G. S., "Prediction of Heart Disease Using Machine Learning Algorithms," 2019 1st Int. Conf. ICIICT, Chennai, India, 2019, pp. 1–5.
[2]. S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," IEEE Access, vol. 7, pp. 81542–81554, 2019.
[3]. B. Alghamdi and F. Alharby, "An intelligent model for online recruitment fraud detection," Journal of Information Security, vol. 10, no. 3, pp. 155–176, 2019.
[4]. S. Vidros, C. Kolias, G. Kambourakis, and L. Akoglu, "Automatic detection of online recruitment frauds: Characteristics, methods, and a public dataset," Future Internet, vol. 9, no. 1, p. 6, 2017.
[5]. V. Krishna, S. Ravi, M. Soora, and A. Sethuraman, "Fake job recruitment detection using machine learning," Int. J. Innovative Technology and Exploring Engineering, vol. 8, no. 10, 2019.
[6]. S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in Proc. 31st Int. Conf. NIPS, Long Beach, CA, USA, 2017, pp. 4768–4777.
[7]. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers," in Proc. NAACL-HLT, Minneapolis, MN, 2019, pp. 4171–4186.
[8]. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," JMLR, vol. 12, pp. 2825–2830, 2011.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 3, March 2026
DOI: 10.17148/IJARCCE.2026.153119

[9]. A. Hazra et al., "Heart disease diagnosis and prediction using ML and data mining: A review," Advances in Computational Sciences and Technology, vol. 10, no. 7, pp. 2137–2159, 2017.
[10]. A. Kaur and J. Arora, "Heart diseases prediction using data mining techniques: A survey," IJARCS, 2015.
[11]. M. B. T. Noor et al., "Application of deep learning in detecting neurological disorders from MRI images," Brain Informatics, vol. 7, no. 11, 2020.
[12]. G. N. Ahmad et al., "Mixed machine learning for efficient prediction of human heart disease," Applied Sciences, vol. 12, no. 15, p. 7449, 2022.
[13]. A. Lombardi et al., "A robust framework for explainable AI markers of MCI and Alzheimer's disease," Brain Informatics, vol. 9, no. 1, 2022.
[14]. P. S. Kohli and S. Arora, "Application of machine learning in disease prediction," in Proc. 4th Int. Conf. ICCCA, Greater Noida, India, 2018, pp. 1–4.
[15]. H. Sharma and M. A. Rizvi, "Prediction of heart disease using ML algorithms: A survey," Int. J. Recent and Innovation Trends in Computing and Communication, vol. 5, no. 8, pp. 99–104, 2017.
[16]. M. Kaliappan, E. Mariappan, M. V. Prakash, and B. Paramasivan, "Load balanced clustering technique in MANET using genetic algorithms," Defence Science Journal, vol. 66, no. 3, pp. 251–258.
[17]. M. Sivaram, M. Kaliappan, S. J. Shobana, Prakash, and V. Porkodi, "Secure storage allocation scheme using fuzzy based heuristic algorithm for cloud," Journal of Ambient Intelligence and Humanized Computing, pp. 1–9.
[18]. S. Vimal, Y. H. Robinson, M. Kaliappan, K. Vijayalakshmi, and S. Seo, "A method of progression detection for glaucoma using K-means and the GLCM algorithm," The Journal of Supercomputing, vol. 77, no. 1, pp. 1–17, 2021. https://doi.org/10.1007/s11227-020-03268-0
[19]. M. Kaliappan, B. Guruprakash, J. Rajalakshmi, T. Blessing Karunya, E. Mariappan, M. Ramnath, and R. Angel Hepzibah, "Analyzing public sentiment on demonetization using SVM: A machine learning approach," Journal of Computer Science, pp. 2482–2487, Dec. 2025.