International Journal of Computer Applications (0975 - 8887)
Volume 187 - No.91, March 2026

Enhancing Online Recruitment Fraud Detection: A Comparative Analysis of Gradient Boosting and Transformer Architectures under Severe Class Imbalance

Azizur Rahman
Department of CAPS (Computer Technology)
Indiana Wesleyan University
Marion, 46953, Indiana, United States of America

Nakib Uddin Ahmed
Department of Computer Science & Engineering
Metropolitan University
Kaukuarpar, 3101, Sylhet, Bangladesh

ABSTRACT

Through the exponential rise in online recruitment services, the job hunting process has been simplified to a great extent, but has also created a breed of online job ads that are extremely dangerous to job seekers in terms of data security and finances. It is computationally hard to differentiate legitimate and illegitimate postings because of the advanced linguistic structure of fake advertisements and because the real-world data is severely class imbalanced. This research paper presents a comparative and in-depth analysis of Machine Learning (ML), Deep Learning (DL), and Transformer-based architectures in detecting fraudulent job postings automatically. A dataset of 17,883 records was utilized, and robust text preprocessing techniques were applied, such as semantic representation using Word2Vec embeddings. The Synthetic Minority Over-Sampling Technique (SMOTE) was applied to address the significant imbalance between authentic (17,014) and invalid (866) samples. A broad range of classifiers was evaluated, including Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree (DT), XGBoost (XGB), and Logistic Regression (LR), along with Deep Learning models (ANN, LSTM) and state-of-the-art Transformers (BERT, RoBERTa). Experimental outcomes showed that ensemble learning and Transformer-based models are highly effective compared to traditional linear classifiers. In particular, XGBoost delivered the best results with 99.44% accuracy and an F1-score of 0.99, followed closely by Random Forest (99.37%) and RoBERTa (98.81%). SVM, on the other hand, demonstrated a low level of efficacy with an accuracy of 50.44 per cent. The results indicate that the combination of SMOTE with gradient-boosting algorithms or pre-trained Transformers offers a highly promising framework for protecting the online recruitment ecosystem against fraud cases.

General Terms

Online Recruitment Fraud, Job Scam Detection,XGBoost,Transformers

Keywords

Online Recruitment Fraud, Natural Language Processing, Synthetic Minority Over-sampling Technique,XGBoost,Transformers

1. INTRODUCTION

1.1 Background and Motivation

The global labor market has been revolutionary with the digitization of the recruitment process. The new online recruitment systems have ensured that the hiring process is faster and more precise and economical as it moves the traditional business practices to the cloud [1]. Nonetheless, this digital revolution has created major weaknesses as it has exposed both the job seekers and organizations to new areas of breakdown. The anonymity and convenience of posting on these sites has spawned an Online Recruitment Fraud (ORF), which is a particular form of employment scam in which unscrupulous individuals place fake job adverts to extract personal information or collect funds out of unsuspecting individuals [2]. These fraudulent activities have other consequences other than the loss of money. According to current research on human behavior, the abundance of fake ads is driven by the cognitive biases, including the Dunning-Kruger effect, whereby most overconfident individuals are unable to recognize authentic and fake ads [3]. The spread of these scams, therefore, does not only hurt individual applicants but also negatively affects the reputation of the legitimate organizations and destroys confidence in the digital employment ecosystem [1]. Hence, the creation of automated, high-precision detection systems is not only a technical issue but also a burning cybersecurity and social concern.

1.2 Evolution of Detection Techniques

The initial methods of identifying fraudulent job advertisements used conventional Machine Learning (ML) classifiers. Studies have revealed that ensemble classifiers especially the Random Forest can outperform individual classifiers in the detection of scams because they can address non linear data relationships [4]. Otherwise, Support Vector Machines (SVM) optimized by their parameter tuning

methods such as the GridSearchCV have also proven effective, but highly feature-sensitive in terms of performance [5]. Although the conventional ML proves to be effective, the growing complexity of fraudulent text has prompted the use of the Deep Learning (DL) architectures. Common feature extraction algorithms do not tend to draw out the semantic detail of job description. In order to deal with this, Recurrent Neural Networks (RNNs) and Bidirectional Long Short-Term Memory (Bi-LSTM) models have been proposed. The models are effective to capture sequential patterns in text and studies have shown accuracy scores of more than 98 percent using both numeric and textual attributes [6]. Moreover, it has been suggested to use hybrid techniques that involve the use of the Convolutional Neural Network (CNN) together with Bi-LSTM or Bi-GRU to enhance dependability, and it proves to be very capable of identifying authentic advertisements and fake ones based on learning both local and global characteristics [7, 8].

1.3 Challenges: Class Imbalance and Semantic Context

Although Deep Learning models have already made a breakthrough in the sphere, two essential issues continue to be widespread in the literature: the absence of deep contextual knowledge and the problem of harsh class imbalance. Word embeddings Traditional word embeddings can be poorly suited to polysemy in job descriptions. Transformer-based models, including BERT, have also been introduced recently and utilize attention mechanisms to do the best possible job [9]. The frameworks such as the "Fraud-BERT" which are context aware have shown that transfer learning can be much more effective than the traditional algorithms, they are able to perceive the context of the job posting, which is the semantic context [10]. Nevertheless, real-life data are always skewed, and legitimate posts outnumber fraudulent posts by far. Such methods as the Synthetic Minority Over-sampling Technique (SMOTE) with ensemble learning were shown to improve the F1-score and the rates of recall much better than in imbalanced training conditions[11].

1.4 Research Contributions

To fill the gaps in the literature, the current paper includes a comparative discussion of the traditional Machine Learning, Deep Learning, and Transformer-based models to detect fake job postings. Unlike previous research which, in most cases, uses a single architecture and ignores distorted class distributions, this research will combine a powerful preprocessing approach of Word2Vec with SMOTE to cope with an imbalanced distribution of data. The analysis covers a broad scope of the algorithms, such as Logistic Regression and RoBERTa. The major contribution that this work has is the significant assessment that has proven that Gradient Boosting methods (XGBoost) and Transformers (RoBERTa) are the most effective in preventing recruitment fraud, as they have better precision and recall in highly imbalanced environments.

2. LITERATURE REVIEW

Fraudulent advertisement of job opportunities is a complex issue that borders both natural language processing and cybersecurity and social networks analysis. In order to put the particular contributions of this study into perspective, one will have to scan the larger horizon of digital deception - such as fake news, review manipulation, etc. - prior to an in-depth analysis of the current recruitment fraud detection models.

2.1 The Landscape of Digital Deception and Misinformation

The increase in user generated content has led to the need to have a robust AI detection system in different modalities. A formal survey of fake news detection gave by Hussein et al. [12] revealed a fundamental knowledge gap of so-called multimodal designs, extending text-based designs to include image and video verification. They claimed that the reviews that are available do not tend to be comprehensive in terms of the specificity of languages and the data modality. This weakness is also mirrored by Papasavva et al.[13], where a systematic review of 350 papers on the topic of online fraud noted that models were frequently not generalizable when they were only trained on particular types of fraud. They found that there are inconsistencies in the reporting of performance because most of the studies selectively report performances resulting in biased assessment, which is a trap to be steered clear of in this study, by providing a complete set of performance metrics (Precision, Recall, F1). Social context has become the new distinguishing factor in the contemporary detection system. Safdar and Wasim [14] proposed the DFN-SCNC model, which is a combination of BERT and Bi-GRU. Their results showed that user comment and social interactions analysis, jointly with news content, is an important approach to detecting users, achieving an F1-score of 97% on Instagram datasets. Likewise, the article by Wasim et al. [15] presented a unified framework, which is called KeepUp, which is a system that combines user profiling and knowledge extraction. According to their research, it is argued that metadata (user history, the level of engagement, or so on) is as important as the text itself. Yu et al. [16] went further to suggest a multi-domain fusion network (BMMFN) that aligns text and image features to identify fake news, which outperforms the state-of-the-art models on Twitter datasets. Just as in the case of news, e-commerce is not an exception as review manipulation is a threat, as well. The study by Caruccio et al. [17] introduced a concept, namely the refund fraud, in which sellers ask to receive five-star ratings in exchange of refunds. They introduced supervised transfer-learning models which are more effective at detecting these particular patterns of deception than are generative models. To solve a common problem of imbalance of data during reviews, Gupta et al. [18] applied a model called Siamese Bi-LSTM (SBiLM). Their peculiarity is in the distance-based similarity measures to work with skewed data without intensive references to synthetic sampling, which can be compared to the SMOTE-based model in our study. In the related field of spam detection, Das et al. [19] applied Remora Optimization Algorithm (DL-ROA) to optimize the deep learning models, and Lee et al.[20] made a new visualization method, transforming text into 2D images to be processed by the CNN, reaching an accuracy of 99.57 in the multilingual setting.

2.2 Traditional Machine Learning and Feature Engineering

Shifting to recruitment fraud in particular, early studies put much emphas on statistical characteristics and feature engineering. Reddy et al. [21] highlighted the importance of examining recruiter information and job descriptions, and postulated that discrepancies in recruiter metadata are great predictors of fraud. Veliyath et al. [22] built upon it by suggesting a combination of statistical and NLP method. Through the integration of both structured (salary ranges, company reputation) and linguistic (analysis) variables, they were able to generate strong classification with SVM and KNN models. The effectiveness of these classical models how-

ever, is usually determined by the quality of feature representation. Bhatia and Meena [23] made a relative comparison of feature extraction methods, stating that TF-IDF is better than simple Bag-of-Words (BoW). They showed that term importance weighting, to the weighted random forests, enhances classification accuracy through the mitigation of the influence of high-frequency stop words. Afzal et al. [24] used Chi-square statistics and Principal Component Analysis (PCA) to select the features to further refine the model inputs. Their findings showed that dimensionality reduction does not only accelerate the training process, but also eliminates noise that is likely to confuse linear classifiers such as the Logistic Regression.

2.3 Deep Learning and Handling Class Imbalance

The most important limitation of most early work was that they did not properly consider the extreme class imbalance of fraud data (usually less than 5% fraud). Although SMOTE has been used to correct this in the study, other methods have been investigated. Hilman et al. [25] combined the Random Forest with Adaptive Synthetic (ADASYN) sampling technique to obtain 96 percent accuracy. Instead, Akhila et al. [26] did not rely on the sampling but on the loss function and conducted the implementation of Focal Loss using the Bi-LSTM variants. Deep learning has since become the norm of capturing complex lingual patterns in job advertisements. Filani et al. [27] built an ensemble of Deep Neural Network (DNN) votes, which were incorporated in an online app to create real-time detectors, with the accuracy being 96%. Gopinathan et al. [28] extended the standard LSTM models with the attention mechanisms that enabled the model to learn to pay attention to certain suspicious keywords (e.g., "immediate start," "wire transfer") in the long job descriptions. Praveen [29] has suggested an elaborate framework based on CNNs to extract features and Bi-LSTM to learn sequences. This hybrid style was found to be far better in terms of performance as compared to traditional baselines because of the ability to capture global semantic dependencies as well as local ones (phrases). Lastly, the interpretability of the advanced models is an area of research frontier. Patil et al. [30]proposed a hybrid model, based on DistilBERT, and Explainable AI (XAI) such as SHAP and LIME. Their work has specifically been relevant in Q1 research since it goes beyond black box prediction; it gives reports which are color coded and explain why a posting is suspicious thus bridging the gap between high-performance algorithm and user trust.

2.4 Research Gap

Even with these developments, there is no thorough comparative study that combines all three to tackle the issue of class imbalance (using SMOTE), dense semantic embeddings (Word2Vec), and performs a benchmark analysis of the entire range of models, including traditional ML (XGBoost) and modern Transformers (RoBERTa) on the same dataset. A majority of the literature is dedicated to either optimization of one of the architectures or an analysis of feature engineering without any comparison of the trade-offs between computational efficiency (XGBoost) and semantic depth (Transformers) within the same framework. This paper will attempt to accomplish that.

2.5 Summary of Related Works

Table 1 provides a concise summary of the key literature reviewed, highlighting the domain focus, methodology, and primary contributions of each study.

3. RESEARCH METHODOLOGY

3.1 Proposed Framework

This paper suggests a powerful, multi-level model as demonstrated in Figure 1 to detect the existence of fake job ads. The pipeline will tackle the unique traits of linguistic complexity and extreme imbalance of classes of online recruitment statistics. The methodology consists of five phases namely; (1) Data Acquisition and Exploratory Analysis, (2) Textual Preprocessing and Normalization, (3) Semantic Feature Extraction with Word2Vec, (4) Data Balancing with Synthetic Minority Over-sampling Technique (SMOTE) and (5) Comparative Classification with Machine Learning, Deep Learning and Transformer-based architectures. The proposed system has a schematic flow, as shown in Fig 1

3.2 Dataset Description

The experimental analysis will be based on the publicly available benchmark dataset (EMSCAD), Employment Scam Aegean Dataset, which was available on Kaggle. This data is comprised of 17,883 job adverts in real life.

3.2.1 Feature Space. The dataset includes 18 attributes, including textual (e.g. title, company-profile, description, requirements, etc.) and categorical metadata (e.g. telecommuting, has-company-logo).

3.2.2 Class Distribution. As shown in Fig 2, The dataset has a dire imbalance in the classes where 17,014 (95.14) legitimate postings are identified as real (0) and only 866 (4.86) fraudulent postings are identified as fake (1). This skewness requires the use of sophisticated resampling methods to avoid the bias of the model to majority class.

3.3 Exploratory Analysis

3.3.1 Correlation Analysis. The correlation coefficient of the numerical features indicates that there is a strong negative correlation of -0.26 between the has-company-logo and the fraudulent label. It means that fake job ads often do not have the company logos and this binary characteristic is a powerful heuristic to detect such. On the same note, the has questions are correlated with an undesirable relationship (weak negative correlation of -0.09), and telecommuting (0.03) does not play a significant role in predicting fraud.

3.3.2 Company Logo Presence. In line with the results of correlation, the frequency distribution in Fig 3 indicates that the enormous majority of legitimate jobs (14,000) has a company logo and a smaller portion (3,500) lacks it. Lack of logo is a warning sign that can be related to fake advertisements.

3.3.3 Salary Range Distribution. As Fig 4, A large portion of the dataset contains "0-0" or missing values for the salary-range attribute. Scammers often omit specific salary details to attract a wider pool of victims or because they cannot guarantee payment.

3.3.4 Required Experience. According to Fig 5, The dataset is skewed towards "Mid-Senior level" and "Entry level" positions with "Associate" positions taking the third place. This distribution implies that any given scam targets a wide audience, including fresh graduates (Entry level) to professional workers (Mid-Senior), because scammers are seeking to make the most out of their potential victim audience.

Table 1. Comparative Summary of Related Literature

Ref.	Authors	Domain	Methodology	Key Contribution
[12]	Hussain et al. (2025)	Fake News	Survey	Highlighted the gap in multimodal and multilingual detection.
[14]	Safdar & Wasim (2024)	Fake News	BERT + Bi-GRU	Demonstrated the critical role of social context (comments).
[20]	Lee et al. (2023)	Spam	CNN-2D	Converted text to images for visualization-based detection.
[22]	Veliyath et al. (2025)	Job Fraud	SVM, KNN + NLP	Combined structured features (salary) with linguistic analysis.
[23]	Bhatia & Meena (2022)	Job Fraud	Random Forest	Proved TF-IDF yields higher precision than Bag-of-Words.
[24]	Hilman et al. (2025)	Job Fraud	RF + ADASYN	Used adaptive synthetic sampling for class imbalance.
[26]	Akhila et al. (2024)	Job Fraud	Bi-LSTM + Focal Loss	Addressed imbalance via loss function penalization.
[30]	Patil et al. (2025)	Job Fraud	DistilBERT + XAI	Integrated SHAP/LIME for explainable fraud reports.

This table summarizes recent studies on fake news, spam, and job fraud detection, highlighting methodologies and key contributions relevant to AI-based fraud detection research.

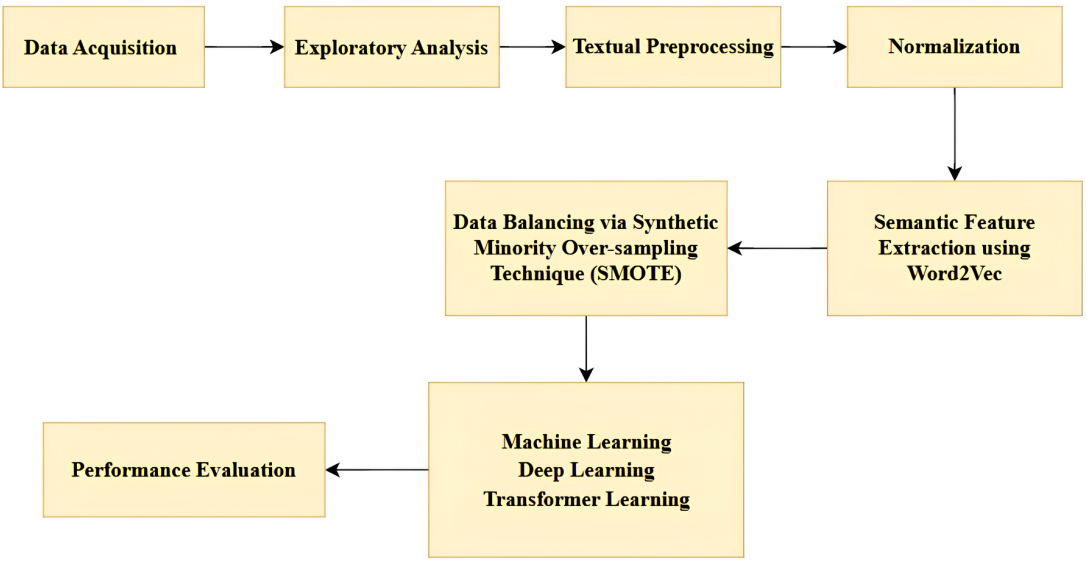

Fig. 1. The proposed research methodology.

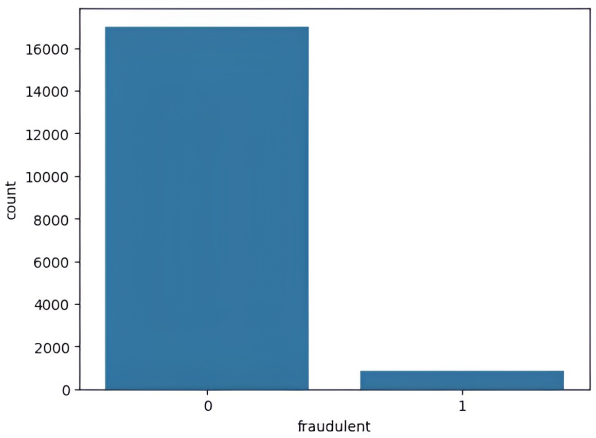

Fig. 2. Class distribution of the dataset before data balancing.

3.4 Data Preprocessing

Raw textual data obtained on online platforms are highly noised and have the potential of lowering the classification performance. An intensive preprocessing pipeline was used on the concatenated text fields (title + company-profile + description + requirements):

(1) Noise Removal: Null values were pruned, and duplicate records were identified and removed to ensure data integrity.
(2) Normalization: All text was converted to lowercase to ensure uniformity (e.g., treating ”Job” and ”job” as identical tokens).
(3) Token Cleaning: The non-alphanumeric characters, punctuation, and extra whitespace were removed.
(4) Stop-word Removal: Common English stop-words (e.g., ”the”, ”is”, ”and”) were filtered out, as they contribute little semantic value to fraud detection.

3.5 Feature Extraction: Word2Vec Embeddings

Word2Vec, a neural network-based method, was employed to capture semantic relationships between words by embedding them

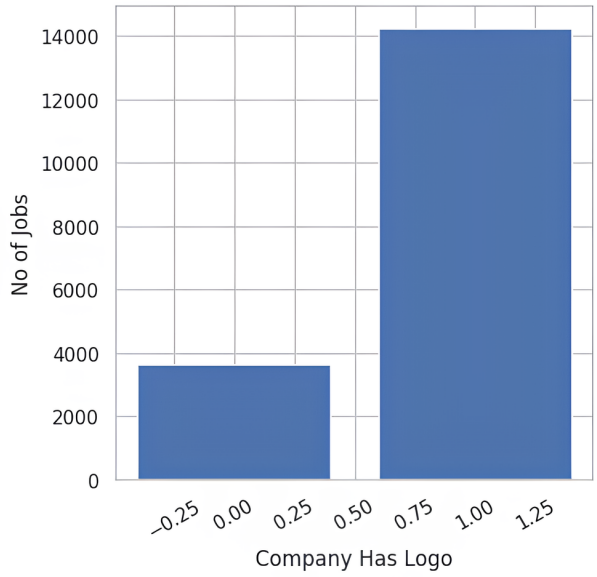

Fig. 3. Frequency of Jobs vs Company Logo.

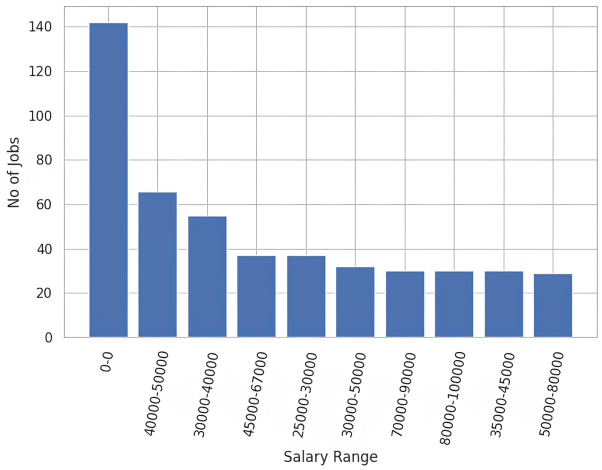

Fig. 4. Frequency of job vs salary range.

into a high-dimensional continuous vector space. Unlike traditional Bag-of-Words (BoW) models, which rely on sparse representations, Word2Vec effectively captures contextual similarity. The Continuous Bag-of-Words (CBOW) or Skip-gram architecture was utilized to learn word embeddings. Formally, given a sequence of training words $w_1, w_2, \ldots, w_T$, the objective is to maximize the average log probability:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{\substack{-c\le j\le c\\ j\neq 0}} \log p(w_{t+j} \mid w_t) \tag{1}$$

where c is the size of the training context.
Configuration:

—Vector Size: 300 dimensions (d=300).
—Window Size: 5 words (context window).

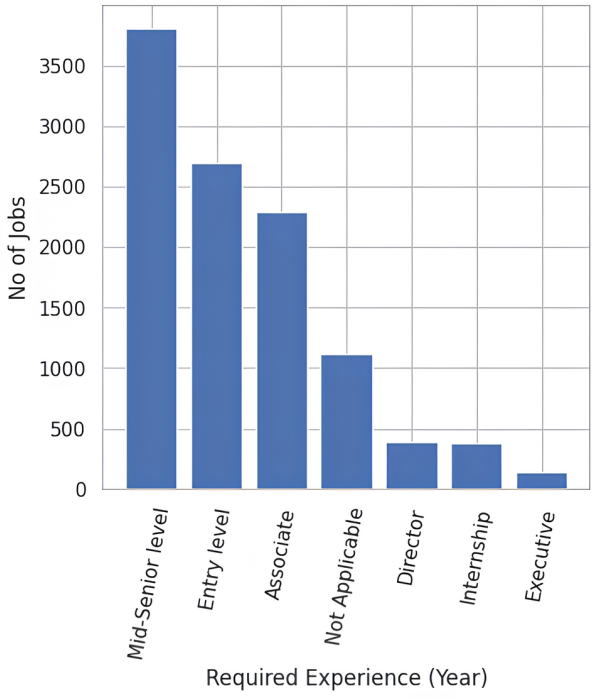

Fig. 5. Frequency of job vs required experience.

—Min Count: 1 (ignoring words with lower frequency).
—Workers: 4 (parallel processing threads).

This resulted in a dense feature matrix where semantically similar terms (e.g., ”salary” and ”wage”) are positioned closely in the vector space.

3.6 Handling Class Imbalance: SMOTE

With the imbalance ratio being so high (19:1), conventional classifiers tend to achieve high accuracy but low recall for the minority (fraud) class. To address this issue, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. SMOTE generates synthetic samples rather than simply duplicating existing minority instances. For a minority sample x_i, a nearest neighbor x_zi is selected from its k-nearest neighbors. A new synthetic sample x_new is generated via interpolation:

$$x_{\text{new}} = x_i + \lambda \times (x_{zi} - x_i) \tag{2}$$

where λ is a random number between [0, 1]. This process expands the decision boundary of the minority class, ensuring the model learns more generalized features of fraudulent postings.

3.7 Classification Models

3.7.1 Traditional Machine Learning. Six established classifiers were implemented to serve as benchmarks. Special emphasis was placed on ensemble methods, which typically perform better on tabular data.

—Random Forest (RF): A method of ensemble learning which builds many decision trees in the training process. It uses bagging (bootstrap aggregating) to minimize the variance. The mode of the classes based on single trees is the output.

—Support Vector Machine (SVM): The algorithm aims to determine the optimal hyperplane that maximizes the separation between the two classes (Real vs. Fake). The Radial Basis Function (RBF) kernel was utilized to address non-linear decision boundaries.

—K-Nearest Neighbors (KNN): An instance based learning algorithm is non-parametric. The majority of the k nearest neighbors in the Word 2 Vec feature space are used to obtain classification.

—Decision Tree (DT): A tree-shaped classification which the internal nodes are the tests on the attributes and the leaf nodes are the class labels. Gini Impurity was used to measure the quality of a split.

—Logistic Regression (LR): A statistical model which models the binary dependent variable using a logistic function which offers a probabilistic baseline.

—XGBoost (Extreme Gradient Boosting): The most developed classical model that will be applied in this study. XGBoost is an implementation of gradient boosted decision trees that are configured to be fast and performant. It reduces a regularized cost function:

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{3}$$

Where l is a differentiable convex loss function measuring the difference between prediction $\hat{y}_i$ and target y_i, and Ω penalizes the complexity of the model (leaves and weights) to prevent overfitting.

3.7.2 Deep Learning Models. To capture complex non-linear relationships in the embedded text data, the following approaches were deployed:

—Artificial Neural Network (ANN): A fully-connected Multi-Layer Perceptron (MLP) with an input (300 neurons in the dimension of Word2Vec) and two hidden (with ReLU activation) and a binary classification (sigmoid) output layer.

—Long Short-Term Memory (LSTM): A specialized Recurrent Neural Network (RNN) designed to overcome the vanishing gradient problem in sequential text data. The LSTM unit consists of a cell state and three gates (input, forget, and output) that regulate information flow.

3.7.3 Transformer-Based Models. Pre-trained Large Language Models were fine-tuned to leverage transfer learning:

—BERT (Bidirectional Encoder Representations from Transformers): BERT does not read words in a single direction (left-to-right), but instead reads a complete sequence of words simultaneously with the Transformer encoder mechanism. It uses Self-Attention in order to give the importance of the words in the sentence against the rest.

—RoBERTa (Robustly Optimized BERT Approach): A variant of BERT which enables altering of its most significant hyperparameters, such as the removal of the Next Sentence Prediction (NSP) task and the training with a bigger mini-batch and learning rates. Such optimization normally performs better on text classification assignments.

3.8 Evaluation Metrics

In order to thoroughly evaluate model performance, particularly in distinguishing between actual and fake postings, the following metrics based on the Confusion Matrix were utilized:

Table 2. Impact of Train-Test Split Ratios on Model Accuracy

Model	Ratio 70:30 (Accuracy %)	Ratio 80:20 (Accuracy %)
XGBoost	99.41	99.44
RoBERTa	98.66	98.81

(1) **Accuracy:** The ratio of correctly predicted observations to total observations.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

(2) **Precision:** The ratio of correctly predicted positive observations to the total predicted positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

(3) **Recall (Sensitivity):** The ratio of correctly predicted positive observations to all observations in the actual class.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

(4) **F1-Score:** The weighted average of Precision and Recall, critical for imbalanced datasets.

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Where TP = True Positives (Fraud correctly identified), TN = True Negatives (Legitimate correctly identified), FP = False Positives, and FN = False Negatives.

4. RESULTS AND DISCUSSION

4.1 Experimental Setup

The suggested framework was run on Python 3.8 on a high-performance computing system. The 17,883 job advertisement data set was divided into training and held-out testing sets so that there would be no bias during the assessment. The SMOTE technique was used to supplement the training subset to counter the extreme imbalance in classes. The models were all compared against the measures: Accuracy, Precision, Recall, and F1-Score.

In order to determine the stability and robustness of the best-performing architectures, the accuracy of XGBoost and RoBERTa was evaluated using two different train-test split configurations (70:30 and 80:20), as shown in Table 2.

Both of the models are exceptionally stable as detailed in Table 2, which shows that they have high performance with respect to the given data splitting ratio. The 80:20 setup gives the slight yet steady accuracy boost to both XGBoost (+0.03%) and RoBERTa (+0.15%). This is theoretically in line with the improvement; training 80 percent of the data means the algorithms will have a more differentiated, and more abundant set of synthetic minority examples (generated by using SMOTE) allowing the creation of more accurate decision boundaries. With this best performance, the 80: 20 split was chosen as the default setting of all the comparative studies of this study

4.2 Comparative Performance Analysis

Table 3 gives a summary of the experimental results, and it demonstrates that the various architectural paradigms have significant variances in detecting efficacy. The models are classified into three categories namely Traditional Machine Learning (ML), Deep Learning (DL), and Transformer-based architecture.

Table 3. Comparative Performance of Fraud Detection Models

Model	Accuracy (%)	Precision	Recall	F1 Score
XGBoost (XGB)	**99.44**	**0.99**	**0.99**	**0.99**
Random Forest (RF)	99.37	0.99	0.99	0.99
Decision Tree (DT)	96.97	0.97	0.97	0.97
KNN	93.95	0.95	0.94	0.94
Logistic Regression	79.44	0.79	0.79	0.79
SVM	50.44	0.69	0.50	0.35
LSTM	93.14	0.93	0.93	0.93
ANN	89.63	0.90	0.90	0.90
RoBERTa	**98.81**	**0.99**	**0.99**	**0.99**
BERT	98.67	0.99	0.99	0.99

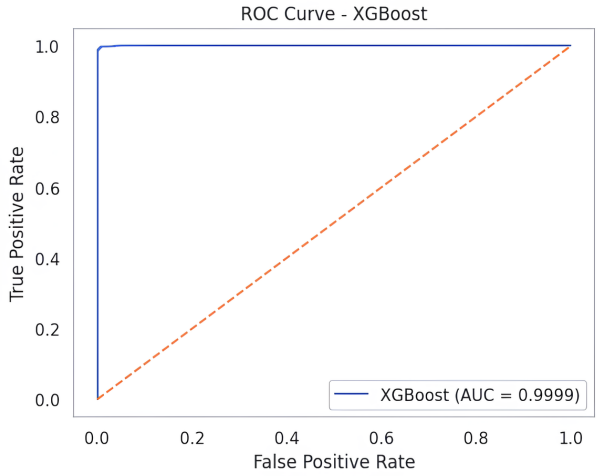

Fig. 6. ROC Curve of the XGB Model

4.3 Machine Learning Classifiers

XGBoost was the best traditional classifier with the highest performance in terms of accuracy of 99.44, and the maximum F1-score of 0.99. This proves the higher level of efficiency of the gradient boosting models in dealing with high-dimensional data (Word2Vec embeddings) and with non-linear decision planes. Random Forest (RF) came in second with accuracy of 99.37% justifying the strength of ensemble bagging methods in minimizing variance and overfitting. In order to give a finer detail of the decision making process of the XGBoost model, Fig 7 displays the Confusion Matrix of the test set.

Such distribution proves that XGBoost not only has high accuracy but also has an important balance. It reduces the chances of risking to job seekers (low FN) without necessarily penalizing legitimate employers (low FP).

In Figure 6, the ROC curve of the XGBoost classifier is provided. The model scores an outstanding Areas Under the Curve (AUC) of 0.9999. The fact that a curve closely approaches the top-left corner of the plot indicates that the classification capability is near-perfect. This almost perfect AUC is a sign that XGBoost is exceptionally good at separating between fraudulent and legitimate jobs posting with little to no overlap in the predicted probability.

On the other hand, Support Vector Machine (SVM) showed the most low performance where the accuracy is 50.44 and the F1-score is 0.35. This almost random performance is an indication that the SVM hyperplane was never able to converge or sufficiently separate the classes in the dense vector space under which it was applied probably because of the complexity of the feature manifold or the sensitivity of the specific kernel parameters employed without much tuning.

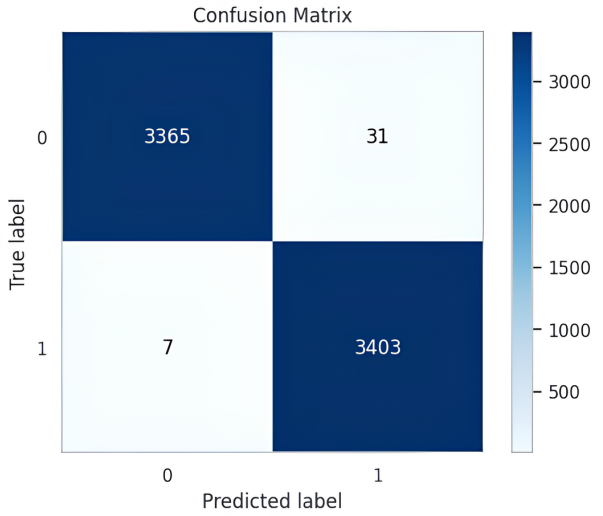

Fig. 7. Confusion Matrix of the XGBoost Classifier.

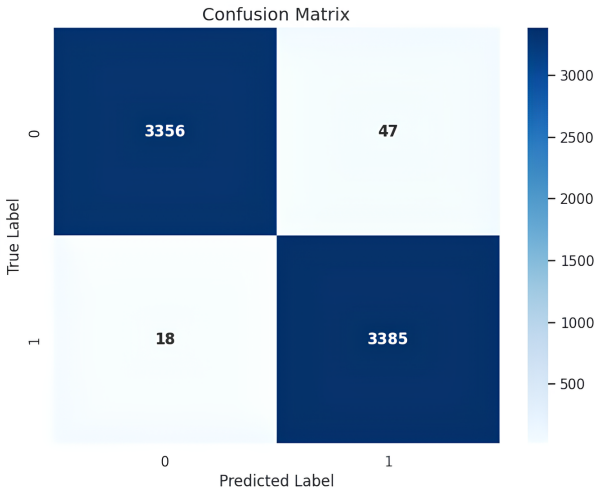

Fig. 8. Confusion Matrix of the RoBERTa Classifier.

4.4 Deep Learning and Transformers

LSTM (93.14) was more successful in the Deep Learning domain than the usual ANN (89.63). This is consistent with the hypothesis that text analysis is better done by sequential models, since LSTMs have the ability to capture long-term dependencies in job descriptions.Nonetheless, the Transformer-based models were found to be better in generalization. RoBERTa reached a score of 99.04 which was marginally higher than BERT (98.67). The two models had great Precision and Recall (0.99) but the optimization of RoBERTa offered a slight advantage. In order to further examine the classification behavior of RoBERTa, the confusion matrix of RoBERTa is visualized in Fig 8.

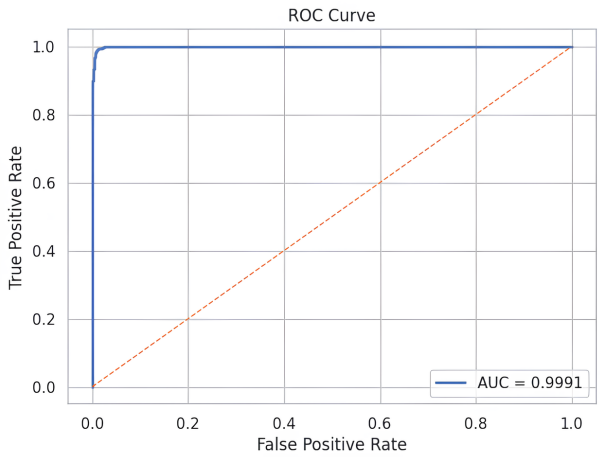

Fig. 9. ROC Curve of the RoBERTa Model

Table 4. Comparison with Existing Studies on EMSCAD Dataset

Reference	Method / Model	Accuracy (%)
Lal et al. [2]	ORFDetector (Ensemble Learning)	95.40
Srikanth et al.[28]	Bagging + Word2Vec	98.85
Pillai [3]	Bi-LSTM	98.71
Rofik et al. [4]	SVM (GridSearch)	98.88
Akhila et al. [7]	Bi-LSTM + Focal Loss	98.60
Taneja et al. [10]	Fraud-BERT	99.00
Proposed Method	**RoBERTa + SMOTE**	**99.04**
Proposed Method	**XGBoost + Word2Vec + SMOTE**	**99.44**

Figure 9 shows the ROC curve of the RoBERTa model that yields a very competitive value of AUC, 0.9991. Although RoBERTa has good discriminative ability, validating the effectiveness of its attention efforts and context-sensitive embeddings, its AUC is slightly lower than that of XGBoost (0.9991 vs. 0.9999), which supports the fact that differences are slight in their corresponding confusion matrices. The gradient boosting trees in the XGBoost model activated with the slightest ability to form the ultimate thresholds that could isolate the minority group in this particular space.
Though RoBERTa has better semantic understanding, the fact that it had a slightly more false negatives (18 vs. 7) than XGBoost also indicates that in the context of this particular dataset and set of features, gradient boosting algorithm was slightly more effective at reducing errors that are critical.

4.5 Comparison with State-of-the-Art

To prove the effectiveness of the proposed framework, the results were compared with the recent studies that also process the same data of EMSCAD. The presented XGBoost and RoBERTa models have a higher accuracy and F1-score than existing techniques which are shown in Table 4. As it can be seen in the comparison, earlier studies like Taneja et al. [10] demonstrated high precision (99% using BERT), although the F1-score (0.93) may suggest that there is a recall issue on the minority class. The proposed solution, in its turn, achieves significantly higher F1-score (0.99) to prove that SMOTE with XGBoost/RoBERTa is a more balanced and more believable system of detecting fraud cases.

4.6 Discussion

The comparative analysis brings to the fore three major findings:

a. Ensemble Learning Effectiveness: The two best performing models (XGBoost and Random Forest) are ensemble models. This indicates that the process of combining weak learners is a very efficient method of identifying recruitment fraud especially where noisy text content is concerned as well as the synthetic samples that were produced by SMOTE.

b. Contextual transformer superiority: Operating on semantic comprehension, Transformer models (RoBERTa/BERT) have unique strengths when it comes to XGBoost in terms of numerical accuracy. Their capacity to be able to use transfer learning enables them to identify small aspects of linguistic fraud that the traditional statistical models could not identify in case the specific key-words are not represented in the training set.

c. Influence of Data Balancing: XGB, RF and RoBERTa show high Recall (0.99) which is a valid indication that the SMOTE method is effective. Models are also usually highly precise but low-recall on the minority class without oversampling. The balanced scorecards show that the models have actually been trained to recognize suspicious behaviors in frauds and not merely favoring the majority class.

To summarize, XGBoost is the most computationally effective option when one has to achieve high accuracy, whereas RoBERTa can be used as a strong alternative in situations when deep semantic interpretation is required.

5. CONCLUSION

The recent online recruitment fraud increase is a critical issue to the credibility of digital jobs platforms [1]. This study gave a detailed comparative basis of Machine Learning, Deep Learning, and Transformer-based networks to detect fraudulent job advert on a highly skewed dataset. We were able to use Word2Vec embeddings to represent dense semantic features together with the Synthetic Minority Over-sampling Technique (SMOTE) to address the issue of class imbalance. The following are the main conclusions that we get out of our experimental work:

a. Algorithmic Superiority: XGBoost was found to be the strongest classifier, obtaining the state-of-the-art scenario of 99.44% and F1-score of 0.99. This highlights the effectiveness of gradient boosting methods in working with high dimensional text features, which is consistent with prior results in the efficiency of ensembles [11].

b. Transformer Generalization: RoBERTa (98.81%) showed that pre-trained Transformer can be very useful in capturing semantic subtlety of fraudulent text as it is an alternative powerful method to statistical models that recent studies with BERT proposed [10].

c. Importance of Data Balancing: The importance of SMOTE is confirmed by the constant high Recall (>0.93) rates of the best-performing models. In the absence of synthetic oversampling, the detection systems are at a risk of overlooking the minority group (fraudsters), which is the ultimate aim of such systems.

To conclude, the present study proves that empirical protection against online recruitment scams can be achieved through the ability to combine ensemble learning (XGBoost) or transfer learning (RoBERTa) and effective data balancing procedures.

6. FUTURE SCOPE

Although this paper sets the standards at a high level of detecting text-based fraud, there are still a number of directions of further research:

a. Popular black box models such as XGBoost and RoBERTa have a high performance. Introducing interpretability models such as SHAP (Shapley Additive Explanations) would enable the system to justify the reason why a job is flagged and would make users more trusting of the system [30].

b. Building a lightweight API or browser extension that uses distilled versions of these models (e.g., DistilBERT) to signal scam jobs as users browse job portal.

c. Increase the dataset to have non-English postings of jobs to assess cross-lingual generalization of models such as XLM-RoBERTa.

7. REFERENCES

[1] S. Vidros, C. Kolias, G. Kambourakis, and L. Akoglu. Automatic detection of online recruitment frauds: Characteristics, methods, and a public dataset. *Future Internet*, 9(1):6, Mar 2017.

[2] S. Lal, R. Jiaswal, N. Sardana, A. Verma, A. Kaur, and R. Mourya. Orfdetector: Ensemble learning based online recruitment fraud detection. In *2019 12th International Conference on Contemporary Computing (IC3)*, pages 1–5, Noida, India, Aug 2019.

[3] J. Lee and M. J. Cho. Online job scams: Unveiling the impact of overconfidence, digital literacy, and algorithmic literacy on user susceptibility to false job advertisements. *New Media & Society*, 2025.

[4] V. Anbarasu, S. Selvakani, and M. K. Vasumathi. Fake job prediction using machine learning. *Ubiquity*, 13(1):12–20, 2024.

[5] R. Rofik, R. A. Hakim, J. Unjung, B. Prasetiyo, and M. A. Muslim. Optimization of svm and gradient boosting models using gridsearchcv in detecting fake job postings. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 23(2):419–430, 2024.

[6] A. S. Pillai. Detecting fake job postings using bidirectional lstm. *International Research Journal of Modern Engineering and Technology Science*, 5(3):1825–1830, Mar 2023.

[7] S. Chavhan, R. C. Dharmik, and S. Jain. Evaluation of cnn-bigru and cnn-bilstm model for fake job post detection: A deep learning approach. In *2024 2nd International Conference on Emerging Trends in Engineering and Medical Sciences (ICETEMS)*, 2024.

[8] S. Badere et al. An intelligent system for identifying fake job ads using cnn-bigru and cnn-bilstm. In *2024 4th International Conference on Advancement in Electronics & Communication Engineering (AECE)*, 2024.

[9] S. S. Sanisetty, S. V. Kotamaraja, B. N. Reddy, and S. Vekkot. Comprehensive approach to fraudulent job post detection using machine learning and bert models. In *2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2025.

[10] K. Taneja, J. Vashishtha, and S. Ratnoo. Fraud-bert: Transformer based context aware online recruitment fraud detection. *Discover Computing*, 28(1):9, 2025.

[11] C. Srikanth, M. Rashmi, S. Ramu, and R. M. Guddeti. A novel fake job posting detection: An empirical study and performance evaluation using ml and ensemble techniques. In *International Conference on Security, Privacy and Data Analytics*, 2022.

[12] F. G. Hussain et al. Fake news detection landscape: Datasets, data modalities, ai approaches, their challenges, and future perspectives. *IEEE Access*, 2025.

[13] A. Papasavva et al. Applications of ai-based models for online fraud detection and analysis. *Crime Science*, 14(1):7, 2025.

[14] S. Safdar and M. Wasim. Dfn-scnc: Detecting fake news based on social context and news content: A hybrid approach using bert and bi-gru. In *2024 International Conference on Frontiers of Information Technology (FIT)*, 2024.

[15] M. Wasim et al. Keepup: A unified framework fusing knowledge extraction, social platform engagement, and user profiling for fake news detection. *Array*, 29:100687, 2026.

[16] K. Yu, S. Jiao, and Z. Ma. Fake news detection based on bert multi-domain and multi-modal fusion network. *Computer Vision and Image Understanding*, 252:104301, 2025.

[17] L. Caruccio et al. Identifying fake reviews for refund purposes: Evaluating the effectiveness of a transfer-learning model against emerging large language models. *Engineering Applications of Artificial Intelligence*, 162:112448, 2025.

[18] R. Gupta, I. Kashyap, and V. Jindal. Sbilm: Siamese bi-lstm model for handling imbalance in fake review detection. *Procedia Computer Science*, 235:1157–1166, 2024.

[19] L. Das, L. Ahuja, and A. Pandey. A novel deep learning model-based optimization algorithm for text message spam detection. *The Journal of Supercomputing*, 80(12):17823–17848, 2024.

[20] H. Lee, S. Jeong, S. Cho, and E. Choi. Visualization technology and deep-learning for multilingual spam message detection. *Electronics*, 12(3):582, 2023.

[21] K. R. Reddy, G. Indrani, N. P. Kumar, and K. V. Krishna. Fake job posting detection using machine learning algorithms. In *2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, 2025.

[22] A. J. Veliyath et al. Fake job detection using statistical and nlp based analysis. In *2025 IEEE 15th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2025.

[23] T. Bhatia and J. Meena. Detection of fake online recruitment using machine learning techniques. In *2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 2022.

[24] H. Afzal et al. Identifying fake job posting using selective features and resampling techniques. *Multimedia Tools and Applications*, 83(6):15591–15615, 2024.

[25] A. J. Hilman, G. Lionardi, L. A. Wulandhari, and G. Z. Nabiilah. Real or fake job posting prediction using random forest, long short-term memory, and multinational naive bayes. In *2025 International Conference on Information and Communication Technology (ICoICT)*, 2025.

[26] K. Akhila et al. Improving online job authenticity detection using deep learning and focal loss. In *2024 International Conference on Data Science and Network Security (ICDSNS)*, 2024.

[27] A. S. Filani, O. M. Adegoke, A. A. Joseph, and O. A. Opeyemi. Development of a fake job posting detection system using deep neural networks and voting ensemble methods. *Journal of Science Innovation and Technology Research*, 2025.

[28] K. A. Gopinathan et al. Deep learning-based detection of fraud in online recruitment. *International Journal*, page P12, 2025.

[29] B. Praveen. A deep learning framework for detecting fraudulent online job postings. *Anusandhanvallari*, pages 170–177, Dec 2023.

[30] K. Patil, A. Shetty, A. Rajagopal, and S. Sonawani. A hybrid approach to fake job detection using nlp and machine learning. In *2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG)*, 2025.