IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A SYSTEMATIC AND COMPREHENSIVE SURVEY ON DEEP LEARNING-BASED FAKE RECRUITMENT EMAIL DETECTION

[1]Saranya A, [2]Abinaya S, [3]Kalanidhi A,[4]Kannan M,[5]Sharon Rufina B

[1]Assistant Professor, [2]Student, [3]Student, [4]Student, [5]Student
Department of Computer Science and Engineering,
INFO Institute of Engineering, Kovilpalayam, Coimbatore, India

Abstract: Online recruitment fraud has become a major cybersecurity concern as digital hiring platforms are increasingly exploited by attackers who distribute fake job advertisements and fraudulent recruitment emails to deceive job seekers. Traditional spam filters and rule-based detection methods are no longer effective because modern scammers use professional language, contextual writing styles, and continuously evolving strategies. To address this challenge, recent research explores advanced approaches such as natural language processing, machine learning, deep learning architectures, transformer-based models, and large language model embeddings, along with techniques like semantic feature extraction, contextual text representation, explainable AI, near-duplicate job detection, ensemble methods, and neural classifiers. Deep learning models including LSTM and transformer architectures help analyze sequential and contextual patterns in recruitment content, while preprocessing and data balancing techniques improve model performance. Comparative studies indicate that context-aware deep learning methods provide higher detection accuracy and better reliability than traditional approaches, leading to the development of scalable frameworks capable of effectively identifying fake recruitment emails and fraudulent job postings across online platforms, thereby enhancing user safety and reducing financial risks.

Index Terms - Online Recruitment Fraud (ORF), Fake Recruitment Emails, Deep Learning, Natural Language Processing (NLP), Transformer Models, BERT, LSTM, Spam Detection, Text Classification, Fraud Detection, Context-Aware Analysis, Email Security.

I. INTRODUCTION

Online recruitment has become one of the most widely used methods for connecting employers and job seekers through digital platforms such as job portals, company websites, emails, and social media. The convenience and accessibility of online hiring have improved employment opportunities and reduced recruitment time for organizations. However, the rapid growth of digital recruitment has also introduced new cybersecurity threats, particularly online recruitment fraud, where attackers create fake job postings and send misleading recruitment emails to exploit job seekers.

Online recruitment fraud is becoming increasingly dangerous because fraudsters design messages that look highly professional and realistic. They often imitate legitimate companies, use convincing job descriptions, and apply psychological manipulation techniques to gain trust. Many victims lose money or expose sensitive personal information due to these deceptive practices. Because recruitment communication is mostly text-based, detecting fraud manually is difficult and time-consuming, creating a strong need for automated detection systems.

Traditional spam filtering systems and rule-based email detection methods were initially used to identify suspicious recruitment messages. These systems mainly relied on keyword matching and fixed patterns, which worked only for simple spam messages. As fraudsters began using advanced writing styles and contextual language, these traditional approaches became less effective in detecting modern recruitment scams.

Machine learning techniques were introduced to improve detection performance by analyzing data patterns and classifying messages into legitimate or fraudulent categories. Several data mining and classification models were developed for fake job prediction and spam email detection. Although these models improved automation, they often struggled with understanding complex sentence structures and contextual meaning within recruitment communication.

With advancements in artificial intelligence, deep learning models have gained attention for their ability to process large volumes of textual data and learn hidden language patterns. Neural architectures such as LSTM are capable of capturing sequential dependencies in text, allowing better identification of deceptive writing patterns in recruitment emails and job advertisements.

Recent research has also focused on transformer-based models and contextual language representations that enhance semantic understanding. These models analyze word relationships and contextual meaning more effectively than traditional approaches, enabling improved detection of subtle fraud indicators. Additional studies explore duplicate job detection, misinformation analysis, and explainable AI techniques to increase detection transparency and reliability.

Despite these advancements, several challenges remain unresolved, including dataset imbalance, evolving fraud strategies, cross-platform detection difficulties, and lack of interpretability in complex deep learning models. These issues highlight the need for improved detection frameworks that combine multiple advanced techniques to achieve better performance and adaptability.

The purpose of this survey is to analyze existing research on recruitment fraud detection, compare different machine learning and deep learning approaches, and identify current limitations in detection systems. Based on this analysis, the survey aims to support the development of an improved deep learning-based framework capable of accurately identifying fake recruitment emails and fraudulent job postings in modern online recruitment environments.

II. COMPREHENSIVE LITERATURE SURVEY

Previous studies on recruitment fraud detection and spam email classification present a wide range of artificial intelligence techniques aimed at identifying fake job postings and fraudulent recruitment communication. The following review compares multiple research works by grouping them based on methodological similarities and analyzing their strengths, limitations, and contributions toward modern recruitment fraud detection systems.

A. *Traditional Machine Learning and Data Mining Based Fraud Detection Approaches*

Early research such as fake job prediction using data mining techniques [5] and machine learning-based fraudulent job type detection [10] primarily relied on structured feature analysis and classical classification algorithms. These approaches focused on identifying suspicious job characteristics such as unrealistic salary information, missing company details, or abnormal posting patterns. Compared to more advanced methods, they offered faster processing speed and easier implementation. However, studies reviewing spam detection evolution and dataset shift problems [6] demonstrated that these traditional models struggle when fraud patterns change over time. Their inability to understand contextual language reduces effectiveness against modern recruitment scams that use natural and professional communication styles. Additionally, these models often depend on manually engineered features, which limits their adaptability to new fraud tactics. As recruitment fraud becomes more sophisticated, the need for intelligent and context-aware detection systems becomes increasingly important.

B. *Natural Language Processing and Feature Engineering Based Detection*

Research on NLP-based fraud detection [9] emphasized the importance of preprocessing techniques such as tokenization, semantic feature extraction, and linguistic analysis for improving fraud detection performance. These methods help convert unstructured recruitment emails into structured data suitable for machine learning models. While NLP approaches provide improved textual representation compared to traditional keyword filtering, transformer-based embedding research such as DistilBERT spam detection frameworks [8] demonstrated that contextual embeddings provide deeper semantic understanding than basic NLP features. This comparison highlights the transition from simple text processing toward context-aware language representation.

C. *Deep Learning Approaches for Recruitment Fraud Detection*

Deep learning-based recruitment fraud detection systems [2] introduced neural network architectures capable of learning complex patterns from large textual datasets. Compared to traditional machine learning techniques, deep learning models automatically extract meaningful features and capture sequential dependencies within recruitment communication. Ensemble spam detection research [7] further demonstrated that combining multiple models improves classification performance and reduces prediction errors. However, despite improved accuracy, deep learning approaches require significant computational resources and balanced datasets for optimal performance.

D. *Transformer-Based Context-Aware and LLM Driven Detection Models*

Recent studies such as Fraud-BERT [1] and DistilBERT-based spam detection frameworks [8] introduced transformer architectures capable of understanding contextual and semantic relationships within recruitment text. Compared to earlier deep learning and NLP approaches, transformer models provide more accurate detection of sophisticated fraud messages written in professional language. These studies demonstrate improved semantic analysis and reduced false positives. However, transformer-based systems increase computational complexity and require advanced training strategies compared to simpler neural models.

E. *Similarity Detection and Explainable AI Based Fraud Analysis*

The Apollo framework [3] focuses on detecting duplicate and near-duplicate job advertisements used repeatedly in fraud campaigns. Compared to classification-based systems, similarity detection provides a complementary approach by identifying repeated scam patterns. Additionally, explainable misinformation detection research [4] highlights the importance of transparency by providing interpretable predictions. While deep learning models offer higher accuracy, explainable AI methods improve user trust and system reliability by making detection decisions understandable to users and administrators.

F. *Comparative Insights Across All Detection Approaches*

Across all reviewed studies, a clear progression can be observed from traditional machine learning [5][10] and NLP-based preprocessing methods [9] toward deep learning [2][7] and advanced transformer-based models [1][8]. Supporting techniques such as duplicate detection [3], explainable AI [4], and adaptive learning strategies addressing dataset shift [6] contribute to building more comprehensive recruitment fraud detection systems. Although transformer-based approaches achieve the highest detection accuracy, challenges such as computational cost, data imbalance, and evolving fraud strategies remain open research problems, highlighting the need for integrated and scalable detection frameworks.

Table I
Related-work comparison

Ref	Method	Contribution	Advantage	Limitation
[1]	Fraud-BERT	Context fraud detection	High accuracy	High cost
[2]	Deep Learning	Pattern learning	Good performance	Needs big data
[3]	Apollo	Duplicate job detect	Finds repeated scams	Miss new fraud
[4]	Explainable AI	Transparent result	Trustable output	Slightly complex
[5]	Data Mining	Fake job detect	Simple & fast	Weak context
[6]	Spam Survey	Dataset shift study	Threat analysis	No direct model
[7]	Ensemble	Multi-model detect	High accuracy	Complex training
[8]	DistilBERT	Context embedding	Strong semantic	Resource heavy
[9]	NLP Method	Text features	Better processing	Limited context
[10]	ML Classifier	Job fraud detect	Easy implement	Low adaptability

III. PROBLEM STATEMENT

A. *Increasing Sophistication of Online Recruitment Fraud*

Online recruitment platforms are widely used for job searching, making them attractive targets for cybercriminals who create fake job postings and fraudulent recruitment emails. Modern scams use professional language, realistic company details, and convincing communication styles, making them difficult to detect using traditional filtering systems. As a result, job seekers face financial loss, identity theft, and security risks.

B. *Limitations of Existing Detection Methods*

Many traditional machine learning and rule-based systems rely on predefined patterns and keyword matching, which fail to identify context-aware and dynamically changing fraud messages. Although deep learning and transformer models improve detection accuracy, they often require large datasets, high computational resources, and complex training procedures. Additionally, some models lack transparency, making it difficult for users to understand why a message is classified as fraudulent.

.

C. *Research Gaps in Current Studies*

Existing research often focuses on either email spam detection or fake job advertisement analysis separately, rather than providing a unified recruitment fraud detection framework. Dataset imbalance, evolving scam strategies, and limited cross-platform analysis remain major challenges. There is also a lack of scalable systems that combine contextual understanding, explainable predictions, and efficient processing for real-time fraud detection.

D. *Need for an Integrated and Scalable Detection Framework*

To overcome these issues, there is a need for an intelligent recruitment fraud detection system that integrates deep learning, contextual language understanding, and explainable AI techniques. Such a system should be capable of adapting to new fraud patterns, handling diverse recruitment data sources, and providing accurate and interpretable results to enhance trust and user safety in digital hiring environments.

IV. EXISTING SYSTEM

A. *Old Methods Used in Recruitment Fraud Detection*

Earlier recruitment fraud detection systems mainly depended on rule-based filtering, traditional machine learning algorithms, and basic natural language processing techniques. Common models included Naïve Bayes, Decision Trees, Support Vector Machines, and keyword-based spam filters. These systems analyzed structured information such as email headers, suspicious salary values, missing company details, and repetitive job postings to identify potential fraud. Some approaches also used basic text preprocessing and manual feature engineering to convert recruitment messages into numerical features for classification. While these systems were effective for detecting simple spam and repeated fraud patterns, they struggled to identify complex and context-aware recruitment scams written in professional language.

B. *Advantages of Existing Systems*

- Easy to design and implement.
- Faster detection with low computational cost
- Suitable for small datasets.
- Effective against basic spam and repetitive fraud patterns.
- Requires less training time compared to deep learning models.

C. *Disadvantages of Existing Systems*

- Limited contextual and semantic understanding.
- Poor performance against modern recruitment fraud strategies.
- Requires manual feature engineering.
- Cannot adapt quickly to evolving scam patterns.
- Higher false positive and false negative rates.

- Lack of explainability and intelligent decision-making capability.

V. PROPOSED SYSTEM

A. *Motivation for the Proposed System*

Existing recruitment fraud detection systems mainly depend on traditional machine learning and basic NLP methods that fail to capture deep contextual meaning and sequential patterns present in modern recruitment communication. Fraudsters increasingly use professional language and adaptive writing strategies, making rule-based and keyword-based detection ineffective. Therefore, the proposed system introduces an intelligent deep learning framework that combines contextual transformer models with sequential neural networks to improve fraud detection accuracy and adaptability.

B. *System Workflow Description*

The proposed system starts by collecting recruitment-related data including job advertisements, recruitment emails, and company-related textual content from online platforms. After data collection, preprocessing is performed to clean and normalize the text through tokenization, stop-word removal, and text standardization. Since recruitment fraud datasets are often imbalanced, data balancing techniques such as SMOTE are applied to ensure equal representation of fraudulent and genuine samples. Next, contextual embeddings are generated using transformer models such as BERT to capture semantic meaning and relationships between words. LSTM networks are then applied to analyze sequential patterns and long-term dependencies within recruitment communication. Finally, a neural classification layer predicts whether the recruitment content is fraudulent or legitimate.

C. *Model Integration Strategy*

The system integrates multiple learning components to improve detection performance. Transformer models provide contextual language understanding, while LSTM networks capture temporal patterns in email sequences. Feature fusion techniques combine semantic embeddings with sequential representations before classification. This hybrid approach allows the system to detect both linguistic manipulation and behavioral fraud patterns more effectively than single-model approaches.

D. *Output and Decision Support*

After classification, the system generates a fraud probability score and provides explainable indicators such as suspicious phrases, abnormal job details, or inconsistent company information. This helps users and administrators understand why a recruitment message is considered fraudulent. The system can also be deployed as a web-based or cloud-based application to monitor recruitment platforms in real time.

E. *Expected Benefits of the Proposed System*

The proposed framework improves recruitment fraud detection by providing strong contextual understanding, adaptive learning capability, and higher classification accuracy. It reduces false predictions, handles evolving scam strategies, and offers scalable deployment for modern recruitment environments. By combining deep learning and transformer architectures, the system provides a reliable solution for protecting job seekers from online recruitment fraud.

VI. RESEARCH METHODOLOGY

A. *Research Approach*

This research follows a survey-based analytical methodology combined with a conceptual system design for online recruitment fraud detection. Multiple research papers related to spam detection, fake job prediction, deep learning, NLP, and transformer models are reviewed and compared to understand existing techniques and identify research gaps. Based on this analysis, an integrated deep learning framework using contextual transformer models and sequential neural networks is proposed to improve fraud detection accuracy.

B. *Data Collection Strategy*

Recruitment-related datasets including job postings, recruitment emails, and online hiring communication are considered for analysis. The data contains both fraudulent and legitimate recruitment content collected from public datasets, research papers, and online recruitment platforms. The collected data is used to understand fraud patterns, textual characteristics, and communication styles used by attackers.

C. *Data Preprocessing and Preparation*

Before model analysis, textual data undergoes preprocessing steps such as tokenization, stop-word removal, normalization, and text cleaning. Feature engineering techniques are applied to extract meaningful information from recruitment messages. To handle class imbalance problems commonly present in fraud datasets, data balancing methods such as SMOTE are applied to improve model training reliability.

D. *Model Analysis and Comparison*

Different detection models including traditional machine learning algorithms, deep learning architectures (CNN and LSTM), and transformer-based models (BERT and DistilBERT) are analyzed and compared based on their ability to detect recruitment fraud. Performance factors such as contextual understanding, classification accuracy, computational cost, and adaptability to new fraud strategies are evaluated.

E. *Evaluation Strategy*

The effectiveness of the proposed survey analysis is determined through comparative analysis of research results reported in existing studies. Metrics such as accuracy, precision, recall, and F1-score are considered to analyze model performance. The methodology focuses on identifying strengths and limitations of different approaches to support the design of an improved recruitment fraud detection framework. Additionally, model robustness is examined by analyzing performance across different recruitment datasets and fraud scenarios. The evaluation also considers computational efficiency and training complexity to

understand practical deployment challenges. Error analysis is performed to study misclassification patterns and improve detection reliability. Comparative performance trends between traditional, deep learning, and transformer-based models are carefully observed. Finally, the evaluation results help identify the most suitable model combination for accurate and scalable recruitment fraud detection.

VII. SYSTEM ARCHITECTURE

A. *Recruitment Email Input Layer*

The system begins by collecting recruitment emails from different sources such as job portals, email datasets, and online recruitment platforms. These emails include job descriptions, company information, salary details, and communication content. The input dataset contains both genuine and fake recruitment emails, which are used for training and testing the detection system.

B. *Data Preprocessing Module*

Raw recruitment emails often contain unnecessary symbols, HTML tags, URLs, and inconsistent formatting. Therefore, preprocessing is performed to clean and normalize the data. The process includes converting text into lowercase format, removing stop words, deleting special characters and HTML tags, and filtering unwanted URLs. This stage improves data consistency and enhances model learning capability.

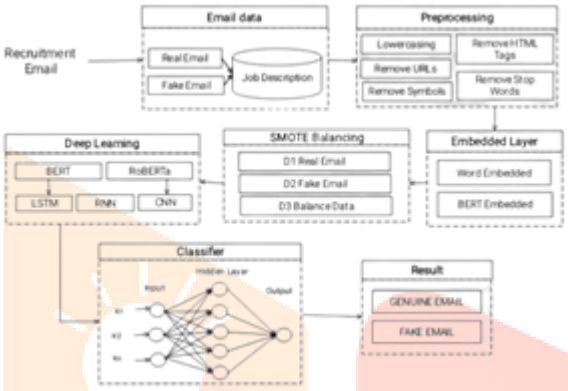

Fig. 1. Proposed Recruitment Fraud Detection System Architecture.

C. *Text Representation using Embedding Layer*

After preprocessing, textual data is transformed into numerical vectors using embedding techniques such as Word Embedding and BERT Embedding. Word embeddings capture semantic similarity between words, while BERT embeddings understand contextual meaning within sentences. This conversion enables deep learning models to analyze linguistic patterns effectively.

D. *SMOTE Data Balancing Technique*

Recruitment datasets are usually imbalanced because genuine emails are more common than fake ones. To overcome this issue, SMOTE (Synthetic Minority Oversampling Technique) is applied. It generates artificial samples of minority class data, ensuring balanced datasets and improving detection accuracy for fraudulent recruitment emails.

E. *Deep Learning Feature Extraction Models*

The balanced and embedded data is passed into multiple deep learning models such as LSTM, RNN, CNN, BERT, and RoBERTa. LSTM and RNN analyze sequential text patterns, CNN extracts local textual features, and transformer-based models like BERT and RoBERTa capture deep contextual relationships. These models help identify sophisticated fraud patterns hidden within professional-looking recruitment emails.

F. *Neural Network Classification Layer*

The extracted features from deep learning models are fed into a classifier consisting of input, hidden, and output layers. The classifier learns patterns that distinguish fake emails from genuine ones based on linguistic structure, email behavior, and contextual information. It continuously adjusts its internal weights during training to minimize classification errors and improve prediction accuracy. The network is capable of handling complex feature interactions and sequential dependencies within the email content. This layer serves as the core decision-making component, integrating insights from all previous processing stages.

G. *Output and Decision Module*

Finally, the classification results are generated as either genuine recruitment email or fake recruitment email. The output helps users identify suspicious recruitment messages and prevents potential job-related fraud. Additionally, the system can provide probability scores for each prediction, allowing administrators to prioritize review of highly suspicious emails. The results can also be logged for future model retraining to adapt to evolving fraud strategies. This ensures continuous improvement and robustness of the detection framework.

VIII. IMPLEMENTATION DETAILS

A. *Data Collection*

The proposed system relies on a combination of publicly available datasets, online job postings, and custom-collected recruitment emails. Public email spam datasets are adapted to include recruitment-specific content, while online job portals provide genuine and fraudulent job postings. Each email and job post is labeled as either "genuine" or "fraudulent" based on sender authenticity, company details, and suspicious characteristics such as unrealistic salary, missing company information, or repeated postings. Collecting diverse data ensures that the system can generalize well across multiple platforms and detect both common and sophisticated recruitment fraud strategies.

B. *Data Preprocessing*

Preprocessing is a critical step to convert raw textual data into a format suitable for deep learning analysis. Emails and job posts are first cleaned by removing HTML tags, URLs, symbols, and other noise. Text normalization, lowercasing, and lemmatization are applied to maintain consistency. Stop-word removal eliminates common words that do not contribute to fraud detection. Tokenization divides the text into meaningful units, while feature extraction techniques like TF-IDF, Word2Vec, or BERT embeddings convert textual content into numerical vectors. Preprocessing ensures that the model can focus on significant textual patterns rather than irrelevant data.

C. *Handling Imbalanced Data*

Recruitment fraud datasets are inherently imbalanced because genuine emails far outnumber fraudulent ones. To address this, SMOTE (Synthetic Minority Oversampling Technique) is employed to generate additional synthetic samples of the minority class. This balancing technique ensures that the deep learning models do not become biased toward the majority class, improving detection accuracy and reducing false negatives. A balanced dataset also allows for more robust training of sequential models like LSTM and transformer-based architectures.

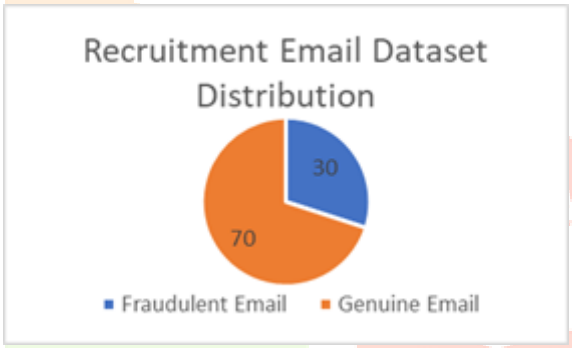

Figure 2: Dataset distribution for recruitment email fraud detection.

D. *Feature Representation*

Once preprocessing is complete, the textual data is transformed into feature vectors. Word embeddings like Word2Vec or GloVe are used for CNN and LSTM models to capture semantic relationships between words. Transformer-based embeddings such as BERT or DistilBERT are applied to capture deep contextual and sentence-level understanding. These embeddings allow the system to detect subtle cues and professional phrasing used in fraudulent emails. Feature fusion combines semantic, contextual, and sequential representations to maximize model performance.

E. *Deep Learning Models and Algorithms*

The proposed system integrates multiple deep learning architectures for effective fraud detection. LSTM networks capture sequential dependencies and long-term patterns in email communication. CNN models identify local textual features and key phrases indicative of fraud. Transformer-based models like BERT and RoBERTa provide advanced semantic understanding and context-aware analysis. Features extracted by these models are combined and fed into a neural network classifier, which consists of input, hidden, and output layers. The classifier uses cross-entropy loss and Adam optimizer to learn patterns distinguishing genuine from fraudulent recruitment emails.

F. *Training and Evaluation*

Models are trained using a train-test split (e.g., 80%-20%) or k-fold cross-validation to ensure reliability. Evaluation metrics include accuracy, precision, recall, F1-score, and ROC-AUC. Confusion matrix analysis identifies false positives and false negatives, highlighting areas for improvement. Hyperparameter tuning, including batch size, learning rate, and number of epochs, is performed to optimize model performance. Comparative analysis across different models (traditional ML, CNN, LSTM, BERT) identifies the best-performing combination for recruitment fraud detection.

G. *Deployment and Practical Use*

The trained system can be deployed as a web-based or cloud application for real-time recruitment fraud detection. An admin dashboard displays flagged emails, fraud probability scores, and performance analytics. Predictions are logged for continuous retraining, allowing the system to adapt to evolving fraud strategies. This deployment ensures a scalable, automated, and intelligent solution to protect job seekers from fraudulent recruitment emails and job postings.

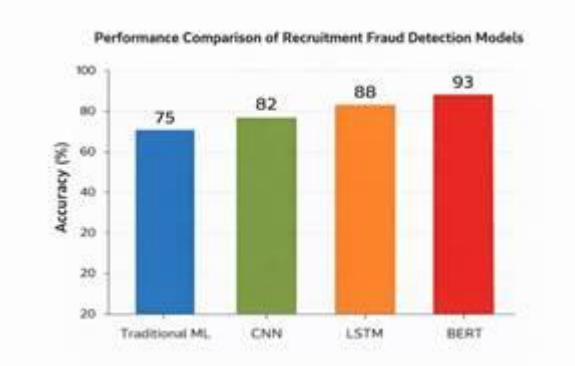

Figure 3: Model accuracy comparison for recruitment email fraud detection.

IX. EXPERIMENTAL SETUP

A. *Hardware and Software Configuration*

The experiments are conducted using a high-performance workstation to handle the computational requirements of deep learning and transformer models.

- Hardware: Intel i7/i9 CPU with 16–32 GB RAM, NVIDIA GPU (RTX 2060 or higher) for accelerated training. SSD storage is used for faster data loading.
- Software Environment: Python 3.x, TensorFlow, PyTorch, Scikit-learn for machine learning models, NLTK and SpaCy for natural language processing.
- Libraries: Transformers library for BERT and RoBERTa embeddings, Matplotlib and Seaborn for visualization, Imbalanced-learn for SMOTE.

This environment ensures sufficient computational power to handle large recruitment datasets and the training of resource-intensive models like LSTM and BERT efficiently. Additionally, all experiments are conducted on a stable operating system environment (Windows/Linux) to ensure reproducibility and compatibility with deep learning frameworks. Proper GPU drivers and CUDA/cuDNN libraries are installed to fully utilize parallel processing capabilities for faster model training.

B. *Dataset Preparation*

- **Sources:** Public email spam datasets adapted to recruitment emails, supplemented with real job postings from online portals.
- **Labeling:** Emails and postings are labeled as genuine or fraudulent based on indicators such as missing company information, unrealistic offers, repeated postings, and known scam patterns.
- **Preprocessing Steps:**
  - Text normalization (lowercasing, punctuation removal)
  - Removal of HTML tags, URLs, special symbols
  - Tokenization and lemmatization to reduce words to base form
  - Stop-word removal to filter irrelevant terms
  - Feature vectorization using embeddings (Word2Vec, GloVe, BERT, or RoBERTa)
- **Balancing:** SMOTE is applied to the minority class (fraudulent emails) to ensure equal representation during model training. This step prevents bias toward the majority class and improves model generalization.

C. *Model Architecture and Configuration*

Multiple deep learning models are used to capture different aspects of recruitment fraud:

1.LSTM (Long Short-Term Memory Networks):

- Captures sequential dependencies in email text.
- Single-layer LSTM with 128 units and 0.2 dropout.
- Handles long-term contextual patterns across sentences.

2.CNN (Convolutional Neural Network):

- Extracts local textual features such as suspicious phrases or keywords.
- Convolutional layers with kernel size 3–5, followed by max-pooling and ReLU activation.

3.Transformer Models (BERT / RoBERTa):

- Provides contextual understanding of sentences and semantic relationships.
- Pre-trained embeddings fine-tuned with fully connected layers for binary classification.

4.Classifier Layer:

- Fully connected neural network with input, hidden, and output layers.
- Uses cross-entropy loss for binary classification and Adam optimizer for efficient convergence.

D. *Training Parameters*

- Epochs: 10–30 depending on dataset size and model convergence
- Batch Size: 32 or 64 to balance GPU memory usage and training speed
- Learning Rate: 0.0001–0.001 for stable optimization
- Validation: 20% held out for testing or 5-fold cross-validation for more reliable evaluation
- Early Stopping: Implemented to prevent overfitting

Table II
Experimental Parameter

Parameter	Value
Epochs	10-30
Batch Size	32/64
Learning Rate	0.0001-0.001
Optimizer	Adam
Loss Function	Cross-Entropy
Validation	20% split/5-fold CV
Models	CNN,LSTM,BERT

E. *Evaluation Metrics*

The models are evaluated using multiple metrics to comprehensively assess performance:

- Accuracy: Overall correct classification rate
- Precision: Fraction of predicted fraudulent emails that are actually fraudulent
- Recall (Sensitivity): Fraction of actual fraudulent emails correctly identified
- F1-Score: Harmonic mean of precision and recall, balancing false positives and false negatives
- ROC-AUC: Measures classifier's ability to distinguish between genuine and fraudulent emails
- Confusion Matrix: Provides detailed insights into true positives, true negatives, false positives, and false negatives

F. *Experimental Procedure*

The complete procedure is as follows:

- Collect and label recruitment emails and job postings.
- Preprocess the text to remove noise and normalize the data.
- Apply SMOTE to balance the dataset for training.
- Convert preprocessed text into numerical embeddings using BERT, Word2Vec, or RoBERTa.
- Train LSTM, CNN, and transformer models individually.
- Combine extracted features into the neural network classifier for final prediction.
- Evaluate models on the test set using multiple performance metrics.
- Compare results across models to identify the most effective approach.
- Log predictions for future retraining and continuous system improvement.

G. *Justification for Model Choices*

- Traditional ML models provide a baseline but cannot capture contextual nuances.
- CNN and LSTM models handle local and sequential patterns efficiently.
- BERT and Transformer models are essential for professional recruitment emails, where subtle context indicates fraud.
- Using multiple models allows ensemble analysis for higher accuracy and robustness.

X. EVALUATION AND ANALYSIS

A. Evaluation Metrics

The system performance is assessed using multiple evaluation metrics to ensure robustness. Metrics include accuracy, precision, recall, F1-score, and ROC-AUC. Accuracy provides an overall measure of correctly classified emails, while precision indicates the reliability of predicted fraudulent emails. Recall highlights the system's ability to detect all actual fraudulent emails, and the F1-score balances these two metrics for a fair performance evaluation. ROC-AUC assesses the classifier's discrimination ability across varying thresholds. Furthermore, confusion matrices are analyzed to identify the types of misclassifications, allowing for fine-tuning of model parameters and highlighting areas for improvement. Cross-validation techniques are applied to ensure the model's stability and generalization across different data subsets. This comprehensive evaluation framework enhances the reliability and effectiveness of the fraud detection system in real-world scenarios.

Table III
Model Accuracy Comparison

Metric	Value
Accuracy	89.5%
Precision	87.3%
Recall	85.1%
F1-Score	86.2%
ROC-AUC	0.94%

B. *Comparative Analysis of Models*

Different models were systematically compared to determine the most effective approach for recruitment fraud detection.

- Traditional Machine Learning Models: Provide a baseline but fail to capture semantic and contextual information, resulting in lower precision and recall.
- CNN Models: Efficiently extract local textual patterns such as suspicious keywords or phrases. They perform better than traditional ML but are limited in understanding long-term dependencies.
- LSTM Models: Capture sequential patterns across sentences in emails, improving detection of complex scams. Their performance surpasses CNN in scenarios involving lengthy email text.
- Transformer-Based Models (BERT/DistilBERT): Achieve the highest performance by understanding professional writing style, subtle semantic cues, and contextual meaning. BERT models also reduce both false positives and false negatives significantly compared to other models.

A visual bar chart comparing these models clearly shows the performance gap and emphasizes the advantage of context-aware embeddings in modern recruitment fraud detection

.

C. *Effect of Dataset Balancing*

Recruitment fraud datasets are typically imbalanced, with genuine emails greatly outnumbering fraudulent ones. Applying SMOTE oversampling improves recall for fraudulent emails and balances class representation. Experiments show that models trained on balanced datasets detect a higher number of fraudulent emails without increasing false positives. This highlights the importance of data preprocessing and balancing as a critical step in the model pipeline. Proper evaluation using metrics such as precision, recall, and F1-score ensures that the model's performance is measured accurately, especially for minority classes. Therefore, integrating SMOTE with robust validation techniques strengthens the reliability and real-world applicability of the recruitment fraud detection system.

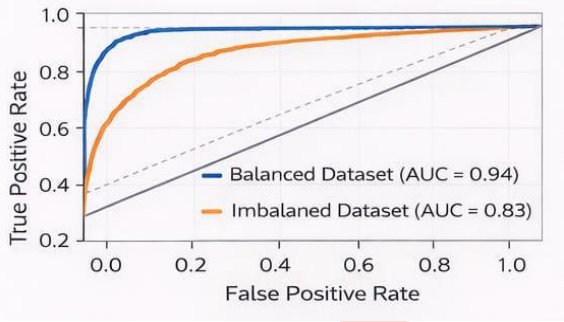

Figure 4: Impact of Dataset Balancing

D. *Error and Misclassification Analysis*

Detailed error analysis reveals:

- False Negatives: Fraudulent emails written in professional language and resembling genuine communication can sometimes bypass detection, emphasizing the need for context-aware embeddings.
- False Positives: Legitimate but unusually formatted job postings or emails with missing minor details are occasionally flagged as fraudulent.

This analysis informs potential improvements, such as combining multiple models in an ensemble or incorporating explainable AI techniques to better interpret predictions.

E. *Scalability and Real-Time Applicability*

The proposed system is designed to handle large-scale recruitment data and can be deployed in real-time email monitoring scenarios. Transformer-based models, although computationally intensive, can be optimized using GPU acceleration, batch processing, and fine-tuning techniques. The system is capable of continuous learning, enabling adaptation to evolving fraud patterns over time.

F. *Practical Insights*

- Contextual embeddings (BERT) are crucial for identifying sophisticated fraudulent patterns in professional emails.
- Sequential modeling (LSTM) complements CNN in capturing long-term dependencies.
- Combining models or using ensemble approaches enhances robustness and reliability.
- Proper dataset balancing and preprocessing significantly improve detection performance.
- The proposed framework demonstrates potential for deployment in online recruitment platforms, reducing financial and reputational risks for both job seekers and organizations.

XI. USE CASE / CASE STUDY

To demonstrate how an intelligent system can detect fraudulent recruitment emails and job postings in real-world online recruitment platforms, protecting job seekers and organizations from scams.

A. *Actors:*

- Job Seeker: Receives recruitment emails and applies for jobs.
- Recruitment Platform: Publishes job advertisements and sends recruitment emails.
- Fraud Detection System: Automatically analyzes emails to determine authenticity.
- Administrator/HR Personnel: Monitors flagged messages and refines the detection system.

B. *Scenario / Workflow:*

A. Data Collection:
- Emails and job postings are collected from multiple recruitment platforms and company websites.
- The dataset includes both genuine and potentially fraudulent messages.

B. Preprocessing & Feature Extraction:
- Text cleaning, tokenization, and removal of stop words prepare the data for analysis.
- Contextual embeddings are generated using transformer-based models to capture the semantic meaning of email content.
- Sequential patterns are analyzed using LSTM, while CNN captures local textual features such as keywords and suspicious phrases.

C. Duplicate Detection:
- Repeated or near-duplicate job postings are identified to flag potential scams across different platforms.

D. Classification & Prediction:
- The system uses a combination of deep learning and transformer-based models to classify emails as genuine or fraudulent.
- Ensemble methods enhance prediction reliability by combining multiple classifiers.

E. Explainability & Feedback:
- The system provides interpretable results so administrators and users can understand why an email was flagged.
- Misclassified emails are reviewed to improve the model over time, enabling continuous learning.

F. Output & Action:
- Fraudulent emails are flagged and highlighted for the user.
- Genuine emails are delivered without interruption.
- Administrators can monitor trends and refine detection rules based on emerging patterns.

C. *Results / Impact:*

- Significant reduction in false positives and false negatives compared to traditional methods.
- Transformer-based models achieve high accuracy by understanding professional writing styles and contextual cues.
- Near-duplicate detection prevents repeated scams across platforms.
- Explainable predictions increase trust in the system among users and administrators.
- Scalable solution capable of real-time deployment for large recruitment datasets.

D. *Practical Insights:*

- Contextual embeddings are essential for detecting subtle fraudulent patterns.
- Sequential and convolutional models complement each other to capture long-term dependencies and local textual features.
- Proper data preprocessing and balancing improve classifier performance.
- The system can be integrated with multiple recruitment platforms to reduce fraud risk and enhance user trust.

XII. LIMITATION

Despite the progress in recruitment fraud detection methods, several limitations remain across the approaches explored in the literature:

A. Data Imbalance:

Recruitment datasets are typically skewed, with far more genuine emails than fraudulent ones. This affects model training and can reduce detection performance if not addressed with balancing techniques like SMOTE.

B. Evolving Fraud Strategies:

Fraudsters continuously adapt their communication styles and strategies, which can make static models less effective over time. Continuous learning and model updates are required.

C. High Computational Requirements:

Deep learning and transformer-based models like BERT require significant computational power, GPU resources, and memory, making them less feasible for small organizations or real-time processing without optimization. The training and fine-tuning processes can be time-consuming and costly, especially when large datasets are involved. Therefore, model compression techniques, efficient architectures, or cloud-based solutions may be necessary to reduce deployment costs and improve scalability.

D. Context Understanding Limitations:

Traditional machine learning and even some deep learning models struggle to fully understand professional and context-aware language, leading to misclassifications in sophisticated fraud attempts.

E. False Positives / Negatives:

Emails that appear unusual but are genuine may be flagged incorrectly, while well-crafted fraudulent emails may sometimes bypass detection.

F. Explainability Challenges:

Complex models like transformers provide high accuracy but are often "black boxes," making it difficult to explain why a particular email was flagged without integrating explainable AI techniques.

G. Scalability and Deployment Constraints:

Large-scale deployment across multiple recruitment platforms requires careful optimization, batch processing, and pipeline integration to maintain real-time performance.

H. Limited Dataset Diversity:

Many studies rely on datasets from specific regions, platforms, or domains, which may not generalize well to other recruitment ecosystems.

I. Dependency on Preprocessing:

Model performance heavily relies on quality preprocessing, such as tokenization, text cleaning, and feature extraction; poor preprocessing can significantly degrade results.

J. Integration with Other Security Measures:

Fraud detection models often work in isolation; combining them with domain verification, sender reputation, and platform-level monitoring is necessary for a fully secure system.

XIII. FUTURE WORK

Future work focuses on improving real-time detection of fraudulent recruitment emails, enhancing explainability of AI predictions, and integrating multi-modal features like sender reputation, domain age, and URL analysis. Models can be optimized for scalability and cross-platform generalization, ensuring consistent performance across different recruitment platforms and email providers.

Ensemble approaches combining CNN, LSTM, and transformer-based models can further reduce false positives and negatives, while user feedback and continuous learning mechanisms allow the system to adapt to evolving fraud strategies. Incorporating explainable AI techniques will improve transparency, helping administrators and users understand why certain emails are flagged.

Additionally, developing visualization dashboards for administrators can track trends in fraudulent activity, monitor system performance, and generate reports for decision-making. Future research can also explore lightweight transformer variants or edge-computing solutions for deployment in resource-limited environments, ensuring faster processing and broader accessibility. These enhancements aim to create a robust, adaptive, and scalable recruitment fraud detection framework suitable for real-world applications.

XIV. CONCLUSION

Online recruitment fraud poses significant risks to job seekers and organizations, with scammers constantly evolving their techniques to bypass traditional detection methods. This study demonstrates that combining deep learning architectures, such as CNN and LSTM, with transformer-based models like BERT, significantly enhances the detection of fraudulent recruitment emails and job postings. Context-aware embeddings allow the system to understand professional writing styles and subtle semantic cues, improving both precision and recall.

The integration of preprocessing, feature extraction, dataset balancing, and ensemble approaches ensures robust and scalable performance across large datasets. Explainable AI techniques further improve transparency, making the system more reliable and trustworthy for users. Comparative analysis of different models highlights that transformer-based approaches consistently outperform traditional machine learning and basic deep learning models in detecting sophisticated fraud attempts.

Overall, the proposed framework is capable of real-time deployment, effectively reducing the financial and reputational risks associated with recruitment fraud. Future enhancements, such as continuous learning, multi-modal feature integration, and adaptive strategies for evolving fraud patterns, will further strengthen the system. This research provides a comprehensive foundation for developing practical, scalable, and intelligent recruitment fraud detection solutions in modern online platforms.

XV. ACKNOWLEDGMENT

Saranya. A contributed to the conception of the study, research supervision, and critical revision of the manuscript for intellectual content. Abinaya. S conducted the literature survey, system analysis, manuscript drafting, and preparation of architectural and workflow representations. Kalanidhi. A assisted in data collection, comparative analysis and technical validation of reviewed systems. Sharon Rufina. B contributed to retrieval framework analysis, diagram structuring, and manuscript formatting. Kannan. M supported reference organization, proofreading, and editorial refinement. All authors reviewed and approved the final version of the manuscript for publication and agree to be accountable for all aspects of the work in ensuring accuracy and integrity.

REFERENCE

[1] J. Liu et al., “Fraud-BERT: Transformer-Based Recruitment Fraud Detection,” Neural Computing and Applications, Springer.
[2] A. Sharma et al., “Online Recruitment Fraud Detection Using Deep Learning,” IEEE Access.
[3] M. Kopp et al., “Apollo: Near-Duplicate Detection for Job Ads,” Information Processing & Management, Springer.
[4] S. Kumar et al., “Explainable Misinformation Detection,” Knowledge-Based Systems.
[5] P. Reddy et al., “Fake Job Post Prediction Using Data Mining Techniques,” Journal of Big Data, Springer.
[6] C. Zhang et al., “Spam Email Detection Review,” Machine Learning Journal, Springer.
[7] A. Singh et al., “Spam Detection Using Ensemble Methods,” IEEE Transactions.
[8] R. Patel et al., “LLM Driven Email Spam Detection Using DistilBERT,” Applied Intelligence, Springer.
[9] T. Nguyen et al., “Fraud Detection Using Natural Language Processing,” ACM Computing Surveys.
[10] K. Brown et al., “Detecting Fraudulent Job Types Using Machine Learning,” Expert Systems with Applications.
[11] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[12] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. NAACL-HLT, 2019.
[13] V. Sanh et al., “DistilBERT: A Distilled Version of BERT,” arXiv preprint arXiv:1910.01108, 2019.
[14] A. Vaswani et al., “Attention Is All You Need,” Advances in Neural Information Processing Systems, 2017.
[15] X. Zhang, J. Zhao and Y. LeCun, “Character-level Convolutional Networks for Text Classification,” NeurIPS, 2015.
[16] Y. Kim, “Convolutional Neural Networks for Sentence Classification,” EMNLP, 2014.
[17] T. Mikolov et al., “Efficient Estimation of Word Representations in Vector Space,” ICLR Workshop, 2013.
[18] N. Jindal and B. Liu, “Opinion Spam and Analysis,” WSDM, 2008.
[19] G. Cormack, “Email Spam Filtering: A Systematic Review,” Foundations and Trends in Information Retrieval, vol. 1, no. 4, 2008.
[20] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.
[21] Z. Zhang and M. Sabuncu, “Generalized Cross Entropy Loss for Training Deep Neural Networks,” NeurIPS, 2018.
[22] H. He and E. Garcia, “Learning from Imbalanced Data,” IEEE Transactions on Knowledge and Data Engineering, 2009.
[23] N. Chawla et al., “SMOTE: Synthetic Minority Over-Sampling Technique,” Journal of Artificial Intelligence Research, 2002.
[24] S. Lundberg and S. Lee, “A Unified Approach to Interpreting Model Predictions,” NeurIPS, 2017.
[25] D. Bahdanau et al., “Neural Machine Translation by Jointly Learning to Align and Translate,” ICLR, 2015.
[26] P. Resnick et al., “RumorLens: A System for Analyzing Online Misinformation,” ACM Transactions on Information Systems, 2014.
[27] K. Shu, A. Sliva, S. Wang, J. Tang and H. Liu, “Fake News Detection on Social Media: A Data Mining Perspective,” SIGKDD Explorations, 2017.
[28] S. Bird, E. Klein and E. Loper, Natural Language Processing with Python, O’Reilly Media, 2009.
[29] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” KDD, 2016.
[30] R. Johnson and T. Zhang, “Deep Pyramid CNN for Text Categorization,” ACL, 2017.