An Effective Survey on Prediction of Fraudulent Online Job Recruitment

Revathi V[1], Balakrishnan C[2]

[1]M.E. Student, [2]Associate Professor
Department of Computer Science and Engineering,
S.A. Engineering College, Chennai, India – 600 071

Abstract - The proliferation of fraudulent job ads has become a major worry with the growing reliance on online platforms for hiring, raising potential hazards for both companies and job seekers. The survey commences by delineating the dynamic domain of virtual employment recruitment and the growing obstacles linked to the widespread dissemination of fraudulent job postings. The investigation of predictive modeling methods used to detect and counteract fraudulent recruitment is a crucial aspect of this survey. Traditional machine learning algorithms, such decision trees and support vector machines, as well as more sophisticated techniques, like ensemble methods and deep learning architectures, are covered in this overview. In order to improve the accuracy of predictive models, the survey also addresses the incorporation of natural language processing (NLP) approaches to assess textual material inside job postings. The survey also looks into how data sources are used for both training and validating predictive models in order to give a thorough understanding. To increase the models' resilience and capacity for generalization, this entails crowdsourcing data, analyzing publicly accessible datasets and working with employment platforms. The purpose of this study is to present a thorough overview of the methods and studies that have already been done in the field of identifying and preventing online fake recruitment.

IndexTerms - Fraudulent job posting, Machine Learning Techniques, Natural Language Processing, Preventing from online fake recruitment

I. INTRODUCTION

The study delves into the array of strategies devised to combat false job postings online. It encompasses a comprehensive survey of each chosen approach aimed at unearthing counterfeit job communication on the web. Ultimately, the investigation aims to identify gaps in the existing body of work dedicated to this topic. Notably, a collection of prominent papers from 2019 to 2023 is compiled, considering the commonalities across various approaches to enhance detection efficacy. This rigorous examination endeavours to determine the most effective mechanism for identifying spurious job postings.

At present, the majority of hiring is done online via sites like naukri.com, linkedin, careerbuilder.com and monster.com. On these sites, businesses submit job listings with the necessary qualifications. Job seekers and applicants can upload their resumes and skill descriptions on these websites. It has led to the misuse of private information, financial loss for job seekers, and damage to the reputation of companies [29]. Candidates can now apply to job profiles they are interested in, and companies can now check the profiles of possible applicants and get in touch with them. After the initial screening, companies get in touch with the selected individuals for further processing and hire qualified candidates. Online hiring is beneficial for both job seekers and employers. About 67 million people were registered on Naukri.com as of December 2022, and 11000 resumes were updated daily. This exemplifies the impact that these internet employment boards have on its users. Online recruitment is advantageous to both recruiters and job seekers. Internet Recruitment Fraud (IRF) is a new type of fraud that has emerged as a result of scammers entering the online recruitment market in recent years. The recruiting environment has changed significantly in the modern digital era, when technology is readily incorporated into a variety of our life. Online platforms are now the hub of talent acquisition, offering previously unheard-of levels of accessibility and ease to companies as well as job seekers. But this paradigm shift has also opened the door for evil actions, especially in the area of harmful internet recruitment, in addition to beneficial transformation.

Prospective employees receive alluring job offers from IRF spammers while their money and personal information is stolen. Not only is IRF bad for users, but it's also bad for business. Moreover, malicious online recruitment takes many forms, each with its own set of obstacles and threats. Moreover, malicious online recruitment takes many forms, each with its own set of obstacles and threats. The threat is persistent and expanding, ranging from sophisticated phishing scams that target naive applicants to fake job advertisements that confuse the desperate job seeker. Impersonation of legitimate institutions, identity theft, and misuse of personal information all add to the complexities of this dark undercurrent. Job searchers may be infected with malware via ostensibly legitimate recruitment channels, resulting in data breaches or financial losses. Scammers readily draw in job seekers by utilizing reputable firm names. As a result, it damages businesses' reputations and leaves job seekers with a negative picture of the specific business. It displays several news clips emphasising the harm brought on by the IRF situation.

In addition to being detrimental to users, ORF presents issues for businesses [10]. Consequently, it damages companies' reputations and leaves job seekers with a negative picture of that specific business. It displays a few news snippets emphasizing the harm the ORF issue has created [13]. Fake job ads are a big issue in the internet era because they hurt real-world communities by causing social unrest, distributing false information, and harming reputations. Intentional deception or inaccurate information can lead to fake job postings. These days determine if a job is legitimate or not gets harder and harder. By analyzing vast amounts of data, machine learning algorithms have demonstrated promising results in the detection of counterfeiting. These systems recognize patterns in the data and produce results based on those patterns. To identify phony job advertisements, machine learning can be used in a variety of contexts and applications. To train machine learning algorithms for the identification of phony jobs, large datasets containing both real and fake job items are required. These databases are used to train algorithms, enabling them to identify patterns in fraudulent employment recruitment. By fine-tuning a machine learning algorithm based on user feedback, one can increase its precision and accuracy. This study attempts to provide a comparative review of the existing method for fraud job detecting advertising using several machine learning algorithms.

In this survey report, opinions are shared regarding different approaches used to find job postings online as well as the degree of accuracy attained with technology.

II. RESEARCH METHODOLOGY

The hierarchical structure of research methodology is given in fig.1

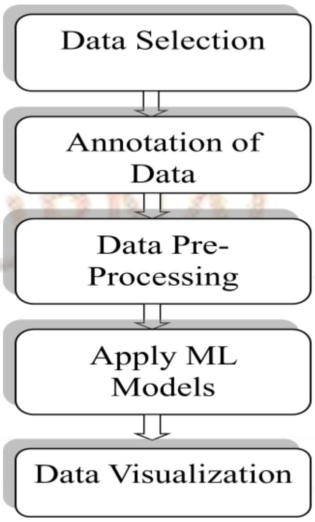

Fig.1 Data Flow Architecture

(1) Data Selection

In this case, try loading the data set from the Kaggle repository. Once the dataset has been downloaded, carry out further operations.

(2) Annotation of Data

Typically, study aims focus on the different sorts of fraudulent employment. To this end, all 866 fraudulent jobs from the EMSCAD dataset were separated out and given unique annotations. The following standards were created in consultation with a Human Resource Management (HRM) industry specialist to identify which category a particular job advertisement falls into. These standards were then used to annotate the position's corporate profile, description, prerequisites, and benefits text fields.

(3) Data Preprocessing

Data pre-processing is a method for creating a clean dataset from raw data. Since the data is in raw format and collected from many sources, analysis is not possible.

(4) Apply ML Algorithms

After the data has been separated into test and train folders, we can apply well-known machine learning algorithms to the training data. We can then evaluate the effectiveness of each method to forecast the hiring of fictitious employees and determine which algorithm produces the most accurate and efficient results.

(5) Data Visualization

The used data set is further divided into two sets, with one third designated as a testing set and the other as a training set. In order to forecast the phony job posts, here a variety of machine learning algorithms are used, including Naïve Bayes, SVM, Random Forest, and Logistic Regression. Ultimately, it determined which method was the best at predicting the messages contained in the fake job postings.

III. LITERATURE OVERVIEW

In this context, it will try to identify a set of several models or strategies that are discussed in connection with the work that is being proposed and how to steer clear of online job recruitment scams. It has put together a selection of research articles that cover various strategies or tactics for recruiting people for phony jobs.

Krishnadas Nanath et al (2023) [1] This study examines a variety of algorithms based on machine learning to address a particular instance of online recruiting fraud (ORF).In this method, Five supervised machine learning (ML) techniques are used to test a model with job posting content attributes such as Naive Bayes, Generalized Linear model, Logistic regression, Decision Trees and Random Forests. Also it delves into several techniques of crowd sourcing. They are, The direct approach to fake content detection (CSM1), The net promoter score approach (CSM2), The fuzzy logic approach (CSM3), Engagement-based approach (CSM4). That could improve accuracy of predictions and bring human knowledge to the field of machine learning algorithms. In this method, they were used EMSCAD (Employment Scam Aegean Dataset). The previously reported work performs as crucial criteria for evaluating the stability of algorithm. Four other parameters—recall, accuracy, F measure, and AUC—were taken into consideration while evaluating performance. Each crowd sourcing strategy had been evaluated across the identical ML algorithms to find out its efficacy in predicting bogus job posts. The two distinct methods of machine learning and crowd sourcing inputs were compared during the testing. One of the major limitations in this study is Large crowdsourcing would require more time and resources, which is not provide in this work, hence it could not be carried out. However expanding the amount of crowdsourced data could aid in gaining deeper understanding of big data analytics.

Cheekati Srikanth et al (2023) [2] In this proposed work, researchers tried to reduce the frequency of such frauds by identifying bogus job ads from genuine ones utilizing ensemble approaches like Bagging, Random Forest, XgBoost, Gradient Boost, AdaBoost, and Stacking classifier. Among the highlights suggested in this article are response coding with Laplace smoothing, Average Word2vec, term frequency-inverse document frequency, and weighted Word2vec. On the publicly available EMSCAD dataset, the authors used accuracy and F1-score to evaluate the performance of ensemble approaches with machine learning (ML) algorithms. On an imbalanced dataset, the stowing classifier beats all other models with an exactness of 98.85% and an F1-score of 0.88. XgBoost achieved 97.89% accuracy and 0.98 F1-score on a balanced dataset. According to the experimental results, a combination of ensemble and featurization methods combining Laplace smoothed Response coding and the BoW technique (Bag-of-Words) outperforms most advanced research on bogus job posting identification.

P. Santhiya et al (2023) [3] This study discusses the subject of identifying recruiting fraud and scams. An efficient model for detecting recruitment fraud is built using three machine learning models and covers relevant managerial, job posting, and kind of compensation data. To ascertain whether a job ad is true or not, the suggested method employs three distinct machine learning techniques: Support Vector Machine, Random Forest, and Naive Bayes Classifier. Two methods were used to extract characteristics from the data: Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF). Consequently, three models show respectable performance. An ensemble model is built by training three independent machine learning models using different segments of samples. The final predictions are then determined by a simple majority vote among the three models. An accuracy of more than 98.18% has been attained using the Random Forest Model.

Ishrat Jahan Mouri et al (2023) [4] The researchers in this study put up a method for identifying online recruitment fraud (ORF). They used the four classification models such as decision trees, random forests, Naive Bayes, and logistic regression techniques and find out which one works best for their proposed model.They calculated and assessed a number of prediction systems' accuracy with 97.16% , the random forest classifier offers the best results.

Manu Gupta et al (2023) [5] This paper explains for the purpose of trying to figure out the links between the many qualities and obtain insights into their multi-class classification, the research starts with exploratory data analysis, or EDA. Next, to get the datasets ready for training and testing, data pre-processing methods are used, such as natural language processing (NLP).Several kinds of machine learning techniques are used, including AdaBoost, Random Forest, Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Logistic Regression. To determine the efficacy of the classifier, performance evaluation metrics such as accuracy, precision, recall, F1-score, selectivity, and specificity are calculated. The outcomes show that the suggested model uses the random forest classifier to categorize employment information with an astounding accuracy of 99.2%.

Y.V Reddy et al (2023) [6] In this paper, various classifiers are employed to verify fraudulent posts on the internet, and the outcomes of those classifiers are compared to determine the most effective model for detecting employment scams. When it comes to identifying fake work posts, two fundamental classifier sorts are taken into thought: single classifiers and ensemble classifiers. Nevertheless, compared to single classifiers, ensemble classifiers are the most effective in detecting scams, according to experimental results.

Hina Afzal et al (2023) [7] This study uses principal component analysis (PCA) and chi-square to analyze selected characteristics in order to overcome the impact of feature select on limitation. Through the use of the synthetic minority oversampling technique (SMOTE), the impact of dataset imbalance is further studied. The suggested model's performance is contrasted with that of other state-of-the-art models and specific machine learning models. The results show that the suggested model produces the best results with 0.99 accuracy when SMOTE is combined with Chi-square-based selected features. K-fold cross-validation provides additional support for these findings.

Zahid Ullah et al (2023) [8] This work aims to develop an ensemble machine learning (ML) based smart secured framework for ORF identification and avoidance. To do this, an identification framework is constructed using four ensemble techniques. Random Forest (RF), Xtreme Gradient Boost (XGB), AdaBoost (AB), and Voting. To get better results, the raw data set had been pre-processed using a variety of cleansing and eliminating techniques. Accuracy, precision, sensitivity, F-measure, and ROC curves were used as metrics for performance evaluation for the employed approaches. These metrics showed that AB did the best, with a high accuracy of 98.374%.

Padma Jyothi et al (2023) [9] This study proposed to analyze different features of the URLs and use machine learning techniques to detect phishing URLs. To identify phishing websites, the study used ensemble learning approaches such as Gradient Boosting, XGBoost, Histogram Gradient Boosting, Light Gradient Boosting, and AdaBoost in addition to classification models such as Logistic Regression, Random Forest, Decision trees, KNN, Naive Bayes, and SVM. Three sets of datasets are employed in this methodology. With a score of 97.13%, XG Boost has the best accuracy in dataset 1.The least accurate algorithm is logistic regression, which has an accuracy rating of 91.73%.Boosting algorithms work incredibly well at enhancing the functionality of less powerful machine-learning algorithms. Further research can also make use of other methods like deep learning and neural networks.

Pablo Quihui et al (2023) [10] In an attempt to address this problem, the ability of various automated systems to discern between real and fraudulent job posts is compared using the Logistic Regression, Multi-layer Perceptron, Random Forest, and Decision Trees algorithms. The model was trained and tested using the Employment Scam Aegean Dataset (EMSCAD). Feature engineering was used on the data set to extract new features from the raw data in order to further enhance their findings. The outcomes showed that the Multi-layer Perceptron (MLP) and Logistic Regression (LR) are capable of correctly identifying phony job postings. Using accuracy, precision, recall, and f1-score as the measures they used to assess each model, these two produced the greatest results overall. Moreover, logistic regression performed admirably, achieving a fair balance between recall and precision. In comparison, the random forest showed a high degree of precision (a value of 0.99). Their results show how well MLP performs in correctly classifying fictitious job postings, with high scores in all the metrics.

Aravind Sasidharan Pillai (2023) [11] This study seeks to close the gap by using a Bidirectional Long Short-Term Memory (Bi-LSTM) model to detect phony job ads. This method successfully captures the underlying patterns and relationships within the data by taking into account both text and numeric aspects. With a 98.71% accuracy rate and a 0.91 ROC AUC score, this suggested method performs exceptionally well and shows promise for real-world use in the online employment market. The results of this study support the creation of reliable, automated solutions that may be used to stop the spread of phony job ads and enhance the credibility of the employment search process as a whole.

Syed Mahbub et al (2022) [12] The authors of this article proposed that the extraction of such contextual data be automated as well. The study comes to the conclusion that adding contextual information enhances the automated online recruiting fraud detection model's performance metrics. The study involves two practical implications. To begin, with minimal localization work, the contextual feature space-generating algorithm may be implemented to any dataset. Second, such learning models can be used at the backbone of online job recruitment platforms to identify and avoid online recruiting fraud. The study not only demonstrates the advantages of integrating contextual variables in identifying fraud using an actual-world dataset, but in addition illustrates how those contextual factors may be regularly obtained via the web using local commercial registries. This approach achieved an accuracy of 91.86%.

Banu Priya Prathaban et al (2022) [13] The Prediction of Employment Scam Model (POESM), which aims to categorize fraudulent and legitimate digital job posting advertising, is implemented in this study. They employed eight techniques in this method, including Logistic Regression, Naive Bayes, Multiple Layer Perceptron, K-Nearest Neighbor, Decision Tree, Random Forests, Adaboost, Gradient Boosting Classifiers. Supervised machine learning methods are used for gathering key information from the dataset with the aim to analyze it. According to empirical findings based on MSE, F1-Score, Cohen-Kappa Score and Accuracy, the Random Forest classifier is the most effective at predicting online fake requirement about 97.98% accuracy than the other classifiers. However, Hyperparameter tuning can also be used to increase the accuracy of this model. Also this model can be connected to a graphical user interface (GUI) application and be used as a desktop application by linking this to a cloud platform.

Bishwajeet Pandey et al (2022) [14] In this research, they explore the performance of various machine learning algorithms in detecting a job scams on job recruitment sites depending on linguistic behavior defined by the employer.This study employed machine learning and deep learning models. Job portals will benefit from the usage of employer-defined language features, which will allow them to recognize and classify such job postings as spam or remove them automatically. There are three machine learning algorithms are used such as Random Forest, Support Vector Machine and Bi-Directional Long Short-Term Memory. The Bi-Directional LSTMs model had the maximum accuracy of 98.679%, while Support Vector Machines had the lowest accuracy of 95.773%.

Marcel Naude et al (2022) [15] This research investigates the extent to which various classification of fraudulent jobs can be classified. Furthermore, this article aims to determine which characteristics are most important in categorizing the type of fraudulent job. In this paper, researchers create and verify a machine learning technique for identifying multi-level marketing, corporate identity theft, and identity theft among phony job advertisements. For different machine learning classifiers, they utilized four sorts of highlights: observational run the show set-based highlights, bag-of-word models, the foremost later state-of-the-art word embeddings, and transformer models. The machine learning models are verified by comparing them to a readily accessible dataset of job postings. According to our findings, word embeddings and transformer-based features consistently beat the customized rule-set based features class. Ultimately, a Gradient Boosting classifier with a mix of parts-of-speech tags, bag-of-words vectors, and empirical rule-set based features produced an F1-score of 0.88. As a result, the indicated job kinds can be separated based on contextual and/or language characteristics. The length of the job description and company profile was an especially useful element. In this method, Multi-level advertising jobs, on the other hand, used more sequential punctuation to pose questions. Future work in this field may investigate other approaches to learning contextual and semantic information from job adverts, as well as new natural language processing techniques.

Aashir Amaar et al (2022) [16] The researchers suggested a strategy for detecting fraudulent job adverts from online recruitment sites that employs natural language processing and supervised machine learning techniques. They extracted features from data using two feature extraction techniques: Term Frequency-Inverse Document Frequency (TF-IDF) and Bag-of-Words (BoW). Six machine learning models were utilized in this study such as Random Forest (RF), Linear Regression (LR), Support Vector Machine(SVM), Extra Tree Classifier (ETC), K-nearest neighbour (KNN) and Multilayer perceptron (MLP) to determine if these job ads were fraudulent or authentic. Then, they compared all models with both BoW and TF-IDF characteristics to assess the overall performance of the classifier. One of the difficulties in this study is the dataset that used. The model over-fitted on majority class data due to an unbalanced ratio of real and false job posting samples. To circumvent this constraint, authors applied the adaptive synthetic sampling strategy (ADASYN), which helps to balance the ratio of target classes by artificially increasing the amount of samples for the minority class. Here they carried out two trials, one with balanced data and the other with data that was unbalanced. ETC obtained 99.9% accuracy through experimental analysis by employing ADASYN as oversampling and TF-IDF as feature extraction. Furthermore, this work conducts an in-depth comparison of their suggested approach with cutting-edge deep learning models and alternative re-sampling strategies.

Priya Khandagale et al (2022) [17] The study suggests an automated method to stop fraudulent job postings on the internet that uses machine learning-based classification techniques. This can be a fraud run by fraudsters who offer to hire people in return for money. Several categorization methods are employed in a machine learning technique to identify fraudulent posts. The system would use historical data from both phony and real job advertisements to train the model to identify jobs as true or fraudulent. Supervised learning algorithms are a promising first step towards addressing the issue of recognizing scammers on job advertisements as classification techniques. Here the researchers are used four machine learning algorithms such as Logistic Regression, Naive Bayes, SVM and Random Forest, choosing the one who best predicts if a job advertisement headline is real or fake, based on accuracy score. The experiments' outcomes demonstrate the effectiveness of Random Forest. In classification, the classifier outperforms its peers in classification. The accuracy percentage of the suggested approach was 97%.

Iffatun Nessa et al (2022) [18] Scammers using inter-net access are constantly coming up with fresh concepts in this study. Recruiting fraud is frequently committed by sending phony emails, SMS messages, phony web pages, bogus social media accounts, phony job postings, online recruiting services like LinkedIn posing as official company correspondence. They proposed a one-layer gated recurrent unit (GRU) model that can distinguish between frauds and legitimate hiring. It also assesses efficacy and obtained 93.51% of AUC score on the Employment Scam Aegean Dataset (EMSCAD) dataset, making it useful for identifying phony or bogus advertisements for jobs.

C. Prashanth et al (2022) [19] The paper discusses how the development of the internet and the easy access to social media platforms like Twitter paved the way for the greatest dissemination of wisdom in the history of humanity. Advertisements of fake job openings are used to lure applicants, giving scammers access to personal data such home address, emails, phone number, date of birth, bank account information, and full identity theft. The authors of this research created Reveal,a machine learning-based web tool to help applicants avoid applying for jobs that are not legitimate and trustworthy by identifying phony job ads online.

Baraneetharan. E (2022) [20] Through the use of machine learning algorithms, researcher hope to reduce the quantity of these phony and fraudulent attempts. This paper uses classifiers like Extreme Gradient Boosting, K-Nearest Neighbor, and Support Vector Machine to predict fake ads. Metrics like accuracy, precision, recall, and F1 measures are used to assess how well machine learning algorithms work.

Riktesh Srivastava (2022) [21] In this paper, the author provided a method to easily identify fake job platforms using a variety of prediction models, such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest, Naïve Bayes, or Logistics Regression.In this method, dataset was used from kaggle. Thirty percent of these job posts assess the model's effectiveness, and the remaining seventy percent train it. Four assessment metrics—Classification Accuracy (CA), Precision, Recall, and F-1 score—are used to forecast each model's results. The study discovered that it is appropriate from two angles: job searchers are protected from fraudulent job posts and the websites are able to detect phony jobs before being published.

Ibrahim M. Nasser et al (2021) [22] proposed an article on "Online Recruitment Fraud Detection Using ANN" The authors focused on this paper, an Artificial Neural Network-based technique to identify fake job advertisements is developed. The public Employment Scam Aegean Dataset (EMSCAD) is combined with suitable content methods of preparation to train and evaluate the proposed model. This model's accuracy, recall, and f-measure are 91.84%, 96.02%, and 93.88%, respectively. According to the findings, the suggested ANN-based model improves equivalent current approaches to detect fraudulent hiring.

Sultana Umme Habiba et al (2021) [23] In this article, the authors highlighted how, in recent years, developments in contemporary technology and social communication have made advertising new job openings a very common issue in today's society. Similar to many other classification tasks, fake job posting prediction has several challenges. In order to determine whether a job ad is real or fraudulent, this study recommends utilizing a variety of data mining techniques and classification algorithms, including KNN, decision trees, support vector machines, nave bayes classifiers, random forest classifiers, multilayer perceptrons, and deep neural networks. Using the 18001 sample Employment Scam Aegean Dataset (EMSCAD), they conducted a study. Deep neural networks fare quite well in this categorization task as classifiers with accuracy 98%.

Hridita Tabassum et al (2021) [24] The researchers offered a method for detecting ORF in this study. The authors used a publicly available dataset as their basis and built their own set of data on the employment sector in Bangladesh. Further, the methods that have been employed are Gradient Boosting, AdaBoost, Decision Tree Classifier, Random Forest Classifier, Voting Classifier, LightGBM, and Logistic Regression.They determined the efficacy of multiple predictive algorithms, with Gradient Boosting (95.17%) and LightGBM (95.17%) giving the best efficacy. They attempted to develop an accurate approach for identifying phony job postings with this study.

Asmitha Shree Rameshprabu et al (2021) [25] The authors of this article arrange fake or genuine posts using AI calculations and accessible informative sources. The purpose of this audit is to combine content, information, and meta-information related to the tasks at hand. By using the information gathered, classification models that can adapt to the dishonest job arrangements can be created. assembling information models through request calculations such as RandomForest, K Nearest Neighbors, and Logistic Regression. With 99.8 percent accuracy, the random forest provides good performance

Ahamed Shibly et al (2021) [26] In this paper, the researchers explained many machine learning methods are currently in use to identify these types of bogus posts. However, the effectiveness of these algorithms needs to be evaluated and contrasted in order to select the best algorithm to use for spotting fakes. This study compares the performance of two-class boosted decision trees and two-class decision forests algorithms through the usage of a suggested model using Microsoft Azure Machine Learning Studio. Researchers compared those two algorithms using F1 Score, recall, accuracy, and precision. The outcomes demonstrated that the two-class forest decision algorithm is not as effective as the two-class boosted decision tree in identifying phony job postings. Thus, it is possible to locate and recognize bogus or gossipy posts on Twitter, messages and social media posts using a two-class decision forest algorithm.

Charan Lokku et al (2021) [27] This work employed TF-IDF vectorizer for feature extraction and natural language processing (NLP) to examine the job posting's sentiments and patterns. Synthetic Minority Oversampling Technique (SMOTE) was used in this model to balance the information and for categorization. Random Forest was also used to predict output with high accuracy; even for large datasets, it runs efficiently, improves model accuracy, and avoids the risk of over fitting.

Ammar Odeh et al (2021) [28] The most advanced methods for detecting phishing websites using machine learning techniques are presented in this study. Based on ML approaches, this research finds solutions to the phishing issue with the website. Throughout the literature, influential machine learning approaches like Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes (NB), and AdaBoosting are analyzed. In addition, this survey research finds that deep learning-based methods outperform traditional machine learning methods in identifying phishing websites. The paper highlights several challenges to machine learning techniques, such as overfitting, low precision, and ineffectiveness when sufficient training data is unavailable.

Shawni Dutta et al (2020) [29] The proposed method introduces classification techniques in machine learning algorithms for checking fraudulent post on the web. Two different classifier types were employed in this method: ensemble classifiers like AdaBoost and Gradient Boosting Algorithm and single classifiers like Naive Bayes , Multi-Layer Perceptron, K-nearest Neighbor and Decision Tree. To determine the most effective model for detecting job scams, the outcomes of those classifiers are compared. A Kaggle dataset that offers details on a job that might or might not be suspicious is used in this situation. In this collection, 17,880 job postings are present. According to experimental findings, the Random Forest classifier works better than its peer classification technology. The proposed approach achieved an accuracy of 98.27%.

Bandar Alghamdi et al (2019) [30] This study makes a significant addition by presenting a reliable detection model that uses an ensemble technique based on a Random forest classifier to identify Online Recruitment Fraud (ORF). To accomplish the study's objectives, the researcher presented the detection model. The support vector machine method is used for feature selection,to conduct this phase, the researcher used Weka tools while the ensemble classifier with Random Forest is used for classification and detection. Moreover, the data was analyzed using a feature selection method, which removed irrelevant and redundant information from the data. The model is applied using the Employment Scam Aegean Dataset (EMSCAD). This Model achieved the best performance with a recall of 0.974.

Yan Ding et al (2019) [31] This research develops a combination detection approach, Search & Heuristic Rule & Logistic Regression (SHLR), to discover hiding strategies on phishing websites and improve the filtering efficiency of legitimate webpages.The approach consists of three steps. First, the webpage's title tag text is entered as search keywords into the Baidu search engine, and the webpage is declared legitimate if the domain name of any of the top-10 results appears; otherwise, additional examination is undertaken. Secondly, if the webpage cannot be classified as legitimate, it is further evaluated to determine whether it is a fake website using the character features' heuristic guidelines. The first two steps can swiftly filter webpages to suit real-time detection requirements. Finally, a logistic regression classifier is utilized to evaluate the remaining pages in order to improve the detection method's adaptability and accuracy. Based on uniform/universal resource locator (URL) lexical information, the SHLR can filter 61.9% of authentic webpages and identify 22.9% of fraudulent webpages. The SHLR's accuracy is 98.9%, indicating that it has a good phishing detection performance.

IV. PERFORMANCE ANALYSIS OF REVIEW FAKE ONLINE RECRUITMENT

S. NO.	PAPER TITLE	AUTHOR	DATASET USED	TECHNIQUES USED	ADVANTAGES	LIMITATIONS	EVALUATION METRICS [OUTCOMES]
1.	An investigation of crowd sourcing methods in enhancing the machine learning approach for detecting online recruitment fraud.	Krishnadas Nanath et al., [2023]	EMSCAD (Employment Scam Aegean Dataset)	Machine Learning Techniques[NB,GLM,LR,DT,RF] and Crowd Sourcing Techniques	It provides stable result.	Large crowd sourcing would require more time and the feature selection lacked sufficient depth.	Accuracy**:** DT**:**80% NB**:**<80% (Less accuracy)
2.	A Novel Fake Job Posting Detection: An Empirical Study and Performance Evaluation Using ML and Ensemble Techniques	Cheekati Srikanth et al.,[2023]	EMSCAD	Ensemble approaches [AdaBoost, Gradient Boost, Stacking classifier, XGBoost, Bagging, RF] and featurization methods	Combination of ensemble model and featurization methods such as Laplace smoothed Response coding and the BoW technique (Bag-of-Words) outperforms most advanced research on bogus job posting identification.	-	Bagging(Imbalan-ced dataset): Accuracy: 98.85% F1-score: 0.88 XGBoost(Balanc-ed dataset): Accuracy: 97.89% F1-score: 0.98
3.	Fake News Detection Using Machine Learning	P.Santhiya et al., [2023]	EMSCAD	Machine Learning Techniques [SVM, RF, NB] and Feature Extraction methods:BoW&TF-IDF	Three models provide respectable outcomes.	Less effectiveness and accuracy	Accuracy: RF:Td-idf-98.18%
4.	Predicting Online Job Recruitment Fraudulent Using Machine Learning	Ishrat Jahan Mouri et al., [2023]	EMSCAD	Classification models[DT,RF,NB,LR]	It calculated and assessed a number of prediction systems' accuracy.	-	Accuracy: RF: 97.16%
5.	Real and Fake Job Classification Using NLP and Machine Learning Techniques	Manu Gupta et al., [2023]	Publicly available Dataset	NLP and Machine Learning Techniques[AdaBoost,RF,NB,SVM,KNN and LR]	Robustness and efficiency of the model in precise.	-	RF: Accuracy: 99.2**%** selectivity value: 96%
6.	Online Fake Job Advert Detection Application Using Machine Learning.	Y.V Reddy et al., [2023]	Publicly available Dataset	Single Classifiers and Ensemble Classifiers	It compares both classifiers and provides accurate result.	More time consuming due to long procedures to do the process.	Ensemble classifiers are more effective than Single classifiers
7.	Identifying fake job posting using selective features and resampling techniques	Hina Afzal et al.,[2023]	Publicly available Dataset	Feature Selection based on PCA and Chi-Square and the Synthetic Minority Oversampling Technique (SMOTE)	K-fold cross-validation provides additional support for the findings.	-	Accuracy: SMOTE- Chi-square:0.99
8.	A smart secured framework for detecting and averting online recruitment fraud using ensemble machine learning techniques	Zahid Ullah et al.,[2023]	Publicly available Dataset (kaggle)	Ensemble Techniques[RF,XGB,AB and Voting]	It shows more reliablility and higher accuracy rates.	-	Accuracy: AdaBoost: 98.374%.
9.	A Machine Learning Approach to Identifying PhishingWebsites: A Comparative Study of Classification Models and Ensemble Learning	Padma Jyothi et al., [2023]	Kaggle and UCI machine learning repository	Ensemble learning approaches[GB, XGBoost, HGB, LGB, and AdaBoost] and Classification Models[LR,RF,DT,KNN,NB and SVM]	High accuracy and Boosting algorithms work incredibly well at enhancing the functionality of less powerful machine-learning algorithms	Complex computation because of three large datasets are used.	High Accuracy: XG Boost-97.13% Low Accuracy: LR-91.73%

S. NO.	PAPER TITLE	AUTHOR	DATASET USED	TECHNIQUES USED	ADVANTAGES	LIMITATIONS	EVALUATION METRICS [OUTCOMES]
	Techniques						
10.	Fake Job Detection with MachineLearning: A Comparison	Pablo Quihui et al., [2023]	EMSCAD	Machine Learning models [LR, Multi-layer Perceptron, RF, DT]	Performance improved	Less reliable	MLP: Recall:0.91 F1-Score:0.94 Accuracy:0.99
11.	Detecting Fake Job Postings Using Bidirectional LSTM	Aravind Sasidharan Pillai [2023]	GitHub Dataset	Bi-LSTM with ensemble learning methods [RF, LightGBM, GBM]	More Reliable	-	Accuracy:98.71% ROC AUC score: 0.91
12.	Online Recruitment Fraud Detection: A Study on Contextual Features in Australian Job Industries	Syed Mahbub et al., [2022]	EMSCAD	NB,RF with contextual features and content-based features	The inclusion of contextual features improves the performance measure.	More contextual features need to produce better result.	Accuracy: RF: 91.86%
13.	Verification of Job Authenticity using Prediction of Online Employment Scam Model (POESM)	Banu Priya Prathaban et al.,[2022]	Kaggle Dataset	LR,NB,MLP,KNN,DT,RF, AdaBoost,GB	Hyperparameter tuning can be used to increase the accuracy of this model.	-	Accuracy: RF: 97.98%
14.	Effective Identification of Spam Jobs Postings Using Employer Defined Linguistic Feature	Bishwajeet Pandey et al.,[2022]	Private Dataset	RF,SVM,Bi-LSTM	Linguistic Features is more advantageous to this model.	-	High Accuracy: Bi-LSTM : 98.68% Low Accuracy: SVM :95.773%
15.	A machine learning approach to detecting fraudulent job types	Marcel Naude et al.,[2022]	EMSCAD	LR,SGD,KNN,DT,SVM,RF,AB, GB with features[BoW, empirical ruleset, word embeddings, transformers]	More number of features provides better accurate result	The limited scope and platform-specific nature of the dataset may raise concerns about the generalizability of the results.	GB: F1-Score:0.88
16.	Detection of Fake Job Postings by Utilizing Machine Learning and Natural Language Processing Approaches	Aashir Amaar et al.,[2022]	Publicly available unbalanced Dataset	RF,LR,SVM,ETC,KNN,MLP with Feature Extraction Techniques[BoW and TF-IDF] and ADASYN strategy	High accuracy and in-depth comparison among data	-	Accuracy: ETC: 99.9%
17.	Fake Job Detection Using Machine Learning	Priya K et al., [2022]	Kaggle Dataset	LR,NB,SVM,RF	Less time consuming	-	Accuracy: RF:97%
18.	Recruitment Scam Detection Using Gated Recurrent Unit	Iffatun Nessa et al.,[2022]	EMSCAD	Gated Recurrent Unit(GRU)	-	-	AUC score: 93.51%
19.	Reveal: Online Fake Job Advert Detection Application using Machine Learning	C. Prashanth et al., [2022]	Publicly available Dataset	NLP,Web Scraping,ML techniques	-	-	-
20.	Detection of Fake Job Advertisements using Machine Learning algorithms	Baraneetharan, E [2022]	Publicly available Dataset	EGB, KNN, SVM	-	-	-

S. NO.	PAPER TITLE	AUTHOR	DATASET USED	TECHNIQUES USED	ADVANTAGES	LIMITATIONS	EVALUATION METRICS [OUTCOMES]
21.	Identification of Online Recruitment Fraud (ORF)through Predictive Models	Riktesh Srivastava [2022]	Kaggle Dataset	SVM, ANN, RF, NB,LR	-	Deepening analysis on variables not done	ANN: CA-0.95 precision-0.907, recall-0.950 F-1 score-0.928 RF: Acc:95.2%
22.	Online Recruitment Fraud Detection using ANN	Ibrahim M. Nasser et al., [2021]	EMSCAD	ANN	It improves equivalent current approaches for fraud detection	-	ANN: Accuracy: 91.84% Recall: 96.02% F-measure: 93.88%
23.	A Comparative Study on Fake Job Post Prediction Using Different Data mining Techniques	Sultana Umme Habiba et al., [2021]	EMSCAD (Employment Scam Aegean Dataset	KNN,DT,SVM,NB,RF,MLP, DNN with variety of data mining techniques	conversion to categorical form	-	DNN: Accuracy:98%
24.	Detecting Online Recruitment Fraud Using Machine Learning	Hridita Tabassum et al., [2021]	Publicly available Dataset	LR, AB, DT, RF Voting, LightGBM, GBoosting	More accurate approach	-	Accuracy: GB: 95.17% LightGBM: 95.17%
25.	Ensemble Modeling on Job Scam Detection	Asmitha Shree Rameshprabu et al., [2021]	EMSCAD	RF, KNN, LR	Good performance	-	Accuracy: RF: 99.8%
26.	Performance Comparison of Two Class Boosted Decision Tree and Two Class Decision Forest Algorithms in Predicting Fake Job Postings	Ahamed Shibly et al.,[2021]	EMSCAD	Two-class boosted decision trees and two-class decision forests algorithms	Model is properly trained and tested using MS Azure machine learning tool.	Just two methods were examined and contrasted	Two-class boosted decision tree: Accuracy:0.938 Precison:0.720 Recall:0.750 F1 Score:0.735
27.	Classification of Genuinity in Job Posting Using Machine Learning	Charan Lokku et al.,[2021]	Kaggle Dataset	NLP, SMOTE, RF with Feature Extraction [TF-IDF]	It runs efficiently for large datasets, improves model accuracy, avoids the risk of overfitting	-	RF: Accuracy:99%
28.	Machine Learning Techniques for Detection of Website Phishing: A Review for Promises and Challenges	Ammar Odeh et al.,[2021]	Publicly available Dataset	RF, SVM, NB, AdaBoost	It identifies deep learning-based techniques with good performance	Overfitting, low precision, and ineffectiveness when sufficient training data is unavailable	Ensemble learning techniques achieve better results. Measurement of parameters not mentioned
29.	Fake Job Recruitment Detection Using Machine Learning Approach	Shawni Dutta et al., [2020]	Kaggle Dataset	NB, KNN, DT, RF, AB, GBoosting	Data cleaning, missing values removed.	-	RF: Accuracy: 98.27%.
30.	An Intelligent Model for Online Recruitment Fraud Detection	Bandar Alghamdi et al.,[2019]	Kaggle Dataset	RF	Filled missing values in MS Excel	-	RF: Accuracy: 97.41%
31.	A keyword based combination approach for detecting phishing webpages.	Yan Ding et al.,[2019]	Publicly available Dataset	SHLR	High performance	-	SHLR : Accuracy : 98.9%

Tab. 1 **Performance Analysis of fake online recruitment-overview**

V. RESULT ANALYSIS

S.No.	ML MODELS	ACCURACY
1.	XGBoost	0.978
2.	Random Forest	0.992
3.	AdaBoost	0.983
4.	Multi Layer Perceptron	0.989
5.	Artificial Neural Network	0.918
6.	Decision Tree	0.930

Tab. 2 Analysis of ML models based on Accuracy

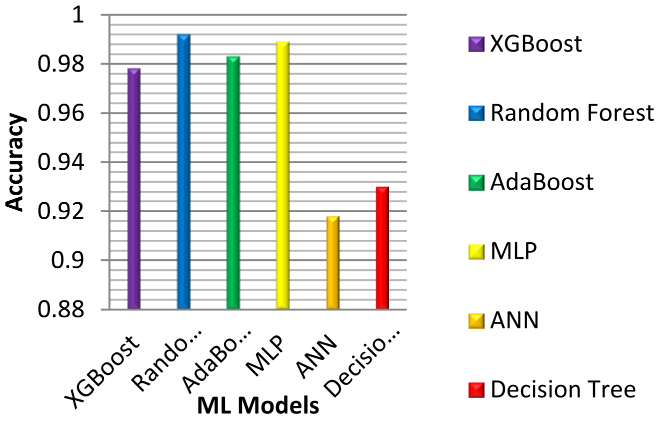

Fig. 2 Comparison of model Time complexities

Consider this study and research indicate that the machine learning approach is the most effective means of guiding job searchers away from scammers' traps. While numerous supervised algorithms are available for classification, studies indicate that Random Forest—which performs better on large datasets—is the most effective method for distinguishing between genuine and fraudulent job posts. Another significant reason is that the oversampling strategy increased the accuracy overall by balancing the dataset and obtaining the best results across all evaluation metrics.

CHALLENGES

Several challenges arise when utilizing machine learning for fake internet recruitment:

Data Quality: The caliber of the training data that machine learning models use greatly influences the models' efficacy. Obtaining high-quality labeled data (i.e., samples of actual and false job advertisements) in the context of bogus internet recruitment might be difficult. Furthermore, inadequate or biased data can result in erroneous predictions from the model.

Feature Selection: It can be challenging to pinpoint the important characteristics that set real job advertisements apart from fraudulent Careful thought ones and domain expertise are needed to extract significant features from unstructured text data (like job descriptions) and other metadata.

Model Generalization: The successful detection of fraudulent job posts depends on how well the trained machine learning model generalizes to new data. Underfitting may result in missed opportunities to identify fraudulent postings, while overfitting to the training data may cause poor performance on fresh data.

Adversarial Attacks: Dishonest people may try to hide from detection by creating job posts that look similar to real ones but are nonetheless false. Developing robust and resilient systems is difficult because adversarial attacks might use flaws in the machine learning model to avoid detection.

Dynamic Character of Fake listings: Scammers modify their tactics in order to evade detection, causing fake job listings to change over time. For machine learning models to be effective, they need to be able to adjust to shifting patterns of fraudulent conduct.

Interpretability and Explainability: Gaining confidence in machine learning models' judgments requires an understanding of how they produce predictions, particularly in crucial applications like identifying fraudulent job postings. It can be difficult to ensure that these models are interpretable and explainable, especially when dealing with complicated models like deep learning systems.

Privacy Concerns: Protecting individuals' privacy while managing sensitive personal information found in job advertising is a major concern. To reduce privacy threats, ensuring adherence to data protection laws and putting in place suitable security measures is crucial.

Strong data collecting and pre-processing methods, cautious model selection and assessment, continual monitoring and adaption to new threats, and open communication regarding the capabilities and limitations of the implemented system are all necessary to meet these challenges. Furthermore, to effectively use machine learning to detect bogus online recruitment, cybersecurity specialists, data scientists, and domain experts must collaborate.

VI. CONCLUSION

This article provides a summary of all earlier research endeavours aimed at identifying phony job listings that are nearly universally accessible online. The survey ends with a summary of the issues and potential paths forward in the field of online fake recruitment prediction, highlighting the necessity of multidisciplinary cooperation, moral considerations, and ongoing model adaption to ever-evolving deceptive tactics. In order to address the growing issue of online fake recruiting and promote a safer digital environment for job seekers, this survey intends to be a useful resource for researchers, practitioners, and policymakers by integrating current study findings, techniques, and problems. In this research explains Random Forest gives more accurate and relevant result but deep learning models have been used in certain articles to accurately conduct specific operations on large datasets and produce correct results.

VII. FUTURE ENHANCEMENT

The field of predicting and stopping online recruitment frauds with ensemble approaches is dynamic, and it requires ongoing improvements to keep up with new threats and boost model efficiency. Future applications of some of the concepts, such Hyperparameter Tuning, Model Diversity, Advanced Text Processing, and Cross-Platform Integration, will take into account the interconnectedness of online activities and allow the model to be extended to evaluate data from numerous online platforms and also investigating several methods to extract contextual and semantic information from the job postings, and may take into consideration extra methods for natural language processing, include GloVe and fastText for word embeddings and Latent Semantic Analysis.

VIII. REFERENCE

[1] Krishnadas Nanath, Liting Olney, “An investigation of crowdsourcing methods in enhancing the machine learning approach for detecting online recruitment fraud”, International Journal of Information Management Data Insights, Volume 3, Issue 1, 2023, 100167, ISSN 2667-0968, https://doi.org/10.1016/j.jjimei.2023.100167.

[2] Srikanth, C., Rashmi, M., Ramu, S., Guddeti, R.M.R. (2023) “A Novel Fake Job Posting Detection: An Empirical Study and Performance Evaluation Using ML and Ensemble Techniques” In: Rao, U.P., Alazab, M., Gohil, B.N., Chelliah, P.R. (eds) Security, Privacy and Data Analytics. ISPDA 2022. Lecture Notes in Electrical Engineering, vol. 1049. Springer, Singapore. https://doi.org/10.1007/978-981-99-3569-7_16.

[3] P. Santhiya, S. Kavitha, T. Aravindh, S. Archana and A. V. Praveen, "Fake News Detection Using Machine Learning," *2023 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2023, pp. 1-8, doi: 10.1109/ICCCI56745.2023.10128339.

[4] Mouri, I.J., Barua, B., Mesbahuddin Sarker, M., Barros, A., Whaiduzzaman, M. (2023). Predicting Online Job Recruitment Fraudulent Using Machine Learning. In: Bindhu, V., Tavares, J.M.R.S., Vuppalapati, C. (eds) Proceedings of Fourth International Conference on Communication, Computing and Electronics Systems . Lecture Notes in Electrical Engineering,,vol 977. Springer, Singapore. https://doi.org/10.1007/978-981-19-7753-4_55

[5] Gupta, M., Sridevi Piratla, N., Chakrapani, S.K., Pasem, Y. (2024). Real and Fake Job Classification Using NLP and Machine Learning Techniques. In: Tiwari, S., Trivedi, M.C., Kolhe, M.L., Singh, B.K. (eds) Advances in Data and Information Sciences. ICDIS 2023. Lecture Notes in Networks and Systems, vol 796. Springer, Singapore. https://doi.org/10.1007/978-981-99-6906-7_14

[6] Reddy, Y. V., Neeraj, B. S., Reddy, K. P., & Reddy, P. B. (2023). Online Fake Job Advert Detection Application Using Machine Learning. *Journal of Engineering Sciences*, *14*(03).

[7] Afzal, H., Rustam, F., Aljedaani, W. *et al.* Identifying fake job posting using selective features and resampling techniques. *Multimed Tools Appl* (2023). https://doi.org/10.1007/s11042-023-15173-8

[8] Ullah Z, Jamjoom M. 2023. A smart secured framework for detecting and averting online recruitment fraud using ensemble machine learning techniques. *PeerJ Computer Science* 9:e1234 https://doi.org/10.7717/peerj-cs.1234

[9] Uppalapati PJ, Gontla BK, Gundu P, Hussain SM, Narasimharo K. A Machine Learning Approach to Identifying Phishing Websites: A Comparative Study of Classification Models and Ensemble Learning Techniques. EAI Endorsed Scal Inf Syst [Internet]. 2023 J un. 23 [cited 2024 Jan. 23];10(5). Available from: https://publications.eai.eu/index.php/sis/article/view/3300

[10] Quihui, Pablo & Pérez Espinosa, Guillermo & Vázquez, Alberto. (2023). Fake Job Detection with Machine Learning: A Comparison.

[11] Pillai, Aravind Sasidharan. (2023). DETECTING FAKE JOB POSTINGS USING BIDIRECTIONAL LSTM. international research journal of science and technology. 2582-5208. 10.56726/IRJMETS35202.

[12] S. Mahbub, E. Pardede and A. S. M. Kayes, "Online Recruitment Fraud Detection: A Study on Contextual Features in Australian Job Industries," in *IEEE Access*, vol. 10, pp. 82776-82787, 2022, doi:10.1109/ACCESS.2022.3197225.

[13] B. P. Prathaban, S. Rajendran, G. Lakshmi and D. Menaka, "Verification of Job Authenticity using Prediction of Online Employment Scam Model (POESM)," *2022 1st International Conference on Computational Science and Technology (ICCST)*, CHENNAI, India, 2022, pp. 1-6, doi: 10.1109/ICCST55948.2022.10040305.

[14] B. Pandey, T. Kala, N. Bhoj, H. Gohel, A. Kumar and P. Sivaram, "Effective Identification of Spam Jobs Postings Using Employer Defined Linguistic Feature," *2022 1st International Conference on AI in Cybersecurity (ICAIC)*, Victoria, TX, USA, 2022, pp. 1-6, doi: 10.1109/ICAIC53980.2022.9897059.

[15] Naudé, M., Adebayo, K.J. & Nanda, R. A machine learning approach to detecting fraudulent job types. *AI & Soc* **38**, 1013–1024 (2023). https://doi.org/10.1007/s00146-022-01469-0

[16] Amaar, A., Aljedaani, W., Rustam, F. *et al.* Detection of Fake Job Postings by Utilizing Machine Learning and Natural Language Processing Approaches. *Neural Process Lett* 54, 2219–2247 (2022). https://doi.org/10.1007/s11063-021-10727-z

[17] Khandagale, Priya, et al. "Fake Job Detection using Machine Learning." *International Journal for Research in Applied Science & Engineering Technology (IJRASET), 10 (5)* 1826 (1822).

[18] Nessa, Iffatun, et al. "Recruitment scam detection using gated recurrent unit." *2022 IEEE 10th Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, 2022.

[19] Prashanth, C., et al. "Reveal: Online Fake Job Advert Detection Application using Machine Learning." *2022 IEEE Delhi Section Conference (DELCON)*. IEEE, 2022.

[20] Baraneetharan, E. "Detection of Fake Job Advertisements using Machine Learning algorithms." *Journal of Artificial Intelligence* 4.3 (2022): 200-210

[21] Srivastava, R. (2022). Identification of online recruitment fraud (orf) through predictive models. Emirati Journal of Business, Economics and Social Studies, 1(1). https://doi.org/10.54878/ejbess.170

[22] I. M. Nasser, A. H. Alzaanin and A. Y. Maghari, "Online Recruitment Fraud Detection using ANN," *2021 Palestinian International Conference on Information and Communication Technology (PICICT)*, Gaza, Palestine, State of, 2021, pp. 13-17, doi: 10.1109/PICICT53635.2021.00015

[23] Habiba, Sultana Umme, Md Khairul Islam, and Farzana Tasnim. "A comparative study on fake job post prediction using different data mining techniques." *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*. IEEE, 2021.

[24] Tabassum, Hridita, et al. "Detecting online recruitment fraud using machine learning." *2021 9th International Conference on Information and Communication Technology (ICoICT)*. IEEE, 2021.

[25] Shree, R. Asmitha, et al. "Ensemble modeling on job scam detection." *Journal of Physics: Conference Series*. Vol. 1916. No. 1. IOP Publishing, 2021, **DOI** 10.1088/1742-6596/1916/1/012167

[26] Shibly, Ahamed & Sharma, Uzzal & Naleer, Hmm. (2021). Performance Comparison of Two Class Boosted Decision Tree and Two Class Decision Forest Algorithms in Predicting Fake Job Postings Keywords:Two class decision forest, Fake job postings, machine learning, MS Azure and Two class boosted decision tree. 25. 2462-2472

[27] Kolli, Kesava & Charan, Lokku & Puganuru, Santhosh. (2021). Classification of Genuinity in Job Posting Using Machine Learning. International Journal for Research in Applied Science and Engineering Technology. 9. 10.22214/ijraset.2021.39580

[28] A. Odeh, I. Keshta and E. Abdelfattah, "Machine LearningTechniquesfor Detection of Website Phishing: A Review for Promises and Challenges," *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, NV, USA, 2021, pp. 0813-0818, doi: 10.1109/CCWC51732.2021.9375997

[29] Dutta S, Bandyopadhyay S (2020) Fake job recruitment detection using machine learning approach. Int J Eng Trends Technol 68(4):48–53. https:// doi. org/ 10. 14445/ 22315 381/ ijett- v68i4 p209s

[30] B. Alghamdi and F. Alharby, ―*An Intelligent Model for Online Recruitment Fraud Detection,"* J. Inf. Secur., vol. 10,no. 03, pp.155– 176, 2019, doi: 10.4236/jis.2019.103009.

[31] Ding, Yan, et al. "A keyword-based combination approach for detecting phishing webpages." *computers & security* 84 (2019): 256- 275.