IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 7, July 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107048 |

Fake Job Post Detection using Random Forest Algorithm

Rakesh G H [1], Manjunatha H T [2]

P.G. Student, Master of Computer Application, JNN College Engineering, Shivamogga, Karnataka, India[1]

Assistant Professor, Master of Computer Application, JNN College Engineering, Shivamogga, Karnataka, India[2]

ABSTRACT: The technology has been streamlined to one position over, and the idea of hiring the workers by the business, through online procedure is carried out. This makes the companies to get the workers of needed post more immediate and in a faster way. It'll be bring-effective as well. By exploring the internet, one can get the job fluently of their qualifications and the field they wish to work in it. The posted jobs may be fake or licit, which are ignorant by the people. To get relieve of these kind of problems we come up with a new software which is designed to prognosticate the job posts, as a result producing whether it's fake or legal bone . We are designing a system as Fake job Post vaticination using the conception of machine literacy, in that we're using Random Forest classifier that produces accurate results in an effective manner. The designed algorithm achieves the result of 98 as compared to the preliminarily used algorithms. The scholars or druggies who search for a job may find difficulties in relating the job posts that are fake and apply for the jobs, entering all the particular information without knowing about it. In some case they may get into the swindles like paying plutocrat in the form of operation freights in the need of job or the assurance of getting job after paying the plutocrat. The frame helps us to descry the posted jobs are fake or not.

KEYWORDS: Fake Job Post, Random Forest Classifier, Machine Learning, Legitimate Job, Decision Tree.

I. INTRODUCTION

Now a days, finding a job is challenging. Before going to any interview you have to apply for a job, get registered also further go for an interview. The first and foremost step is to apply for a job according to the conditions of a company and as per the field a user wants to get a job in it. When you explore on internet you may find several job bulletins, those job bulletins may be a phony jobs or legit jobs. user may not find it easy as it's hard to say, the posted job is a fake or legit . So, we bear a software to descry which is the fake job and which is not, helping a number of people not to expose their particular details to anyone by being alive of the fake job bulletins. The companies post about the job to make the hiring process more easy and immediate. We are using different data mining ways to break the issue of fake job announcement. On applying Random Forest Classifier, it gives the swish results, in relating the fake job bulletins which is better than the previously used.

This helps them to avoid financial losses like they may ask you to pay operation figure, for getting registered or they may ask capitalist in different forms, as a part of procedure in recovery or others. Every companies go for the online process of hiring workers, by posting the job details, if the information entered by the pupil or user matches the job details also they are hired by the company. The need for job by the people, exploring on internet may blindly have trust on anyone and expose their information to any fake job bulletins, which can be misused like bank information, etc.

One of the major problems in is employment fraud. Online recruitment frauds (ORF) have received significant attention [1]. Recently, numerous Companies want to list their job openings online so that job seekers may access them quickly and conveniently. However, given that the con artists offer employment to job seekers in exchange for money, this could be one of their scams. A reputable business may be targeted by fraudulent job postings if they compromise their credibility. These fraudulent job post detections attract a lot of interest in developing an automated method for recognizing bogus jobs and alerting people to them so they won't apply for them.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107048 |

II. RELATED WORK

Here are a few examples of literature reviews: Videos, et al. [1] made a substantial contribution to correctly identifying online process frauds. Online employment scams employ the Random Forest Classifier technique. Scams that use technology differ from those that use online hiring. SVM is used to choose features, while Random Forest Classifier is used to identify and categorize objects. The EMSCAD dataset, which has hundreds of data and is publicly accessible, was used by Alghamdi and Alharby, et al. [2]. Our final percentage is 97.41%. The two main areas of focus are a corporation's corporate logo and a few other important qualities. In their model, Tin Van Huynh, et al. [3] stated that it is important to take into account a candidate's knowledge and skills before hiring them. A person or student who is qualified for the role should be chosen by the commercial companies. With a variety of different neural networks, including Text CNN, BI-GRU-LSTM, and others, preconditioned data. This will result in efficient output with a f1 score of 72.71 percent. According to Jiawei Zhang et al. [4], the expansion of online social networking is accelerating daily in both political and economic dimensions. The users may be harmed by the false news reports. Knowing if news regarding a particular topic is fake or not is crucial. We employ machine learning (ML) algorithms to investigate the sources of the news and the topics they choose from online social networks in order to address the issue of fake news. Our goal is to generate news of the highest Caliber. et al., Thin Van Dang [5]. Virtual neurons are created using DNN and their weights are first set to random integers. After multiplying the weight by the input, the result falls within the range of 0 and 1. Weights are adjusted during training time so that output is divided into various groups. The over fitting issue is a result of the less than ideal patterns' extra layers. The model uses dense layers for data training. By reducing the layers for a few parameters that need training, a generic model can be produced. The rule is the activation function, while the Adam is the optimizer. As part of the training process, Adam evaluates each trainee's pace of learning depending on specific parameters. According to P. Wang et al. [6], the model's tenets are the building blocks of neural networks that behave similarly to how a human brain works. This enables a computer to compare two patterns and determine whether they are similar or dissimilar. A neuron is a structure that has some qualities and classifies things into groups. A neural network is made up of numerous layers of connections between various nodes. Jihadists are placed in layers and discuss Perceptron's in [7]are a part of each other. By adjusting the input layers' weight via hidden layers, the error rate can be reduced.

III. METHODOLOGY

A group of computer algorithms known as "machine learning" may learn from examples and get better over time without needing any explicit coding from a programmer. One of the common machine learning problems is making recommendations. A variety of jobs also make use of machine learning. The machine's brain is where all learning takes place. Machine learning is similar to how human learning works. People learn by experience. When we come across one, our chances of success are lower than they would be in a known circumstance. The identical instruction is given to machines. The system searches for a precedent to anticipate the outcome more precisely. When we give the computer a situation that is comparable, it can predict the outcome.

Learning and subsequent inference are the main goals of machine learning. The machine learns from the discoveries first. This conclusion was made possible by the facts. One of the most crucial talents of a data scientist is their ability to carefully choose the data they feed the computer. A feature vector is a group of characteristics that are utilized to address a problem. A feature vector can be viewed as a component of data that is used to address a challenge. Using some advanced techniques, the machine simplifies reality and turns this discovery into a model. As a result, during the learning phase, the data are summarized and expressed as a model.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107048 |

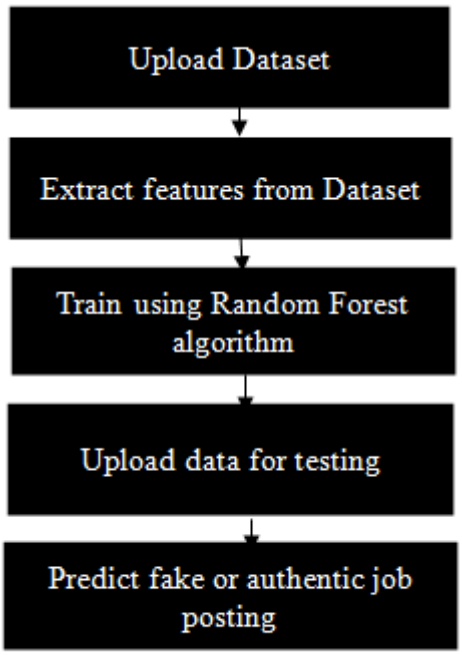

Figure 1: Methodology

There are two forms of machine learning. 1. Directed Education 2. Unsupervised Education
1. Supervised Learning: Using data that is sent into the computer, we train the system. The data stream is used as an input to generate output. It has a wide range of different classifiers and algorithms.
2. Unsupervised learning: In unsupervised learning, an algorithm examines input data without being given a specified output variable. It can be utilized when we require an algorithm to seek for trends and classify data for us because we are unsure of how to do so.

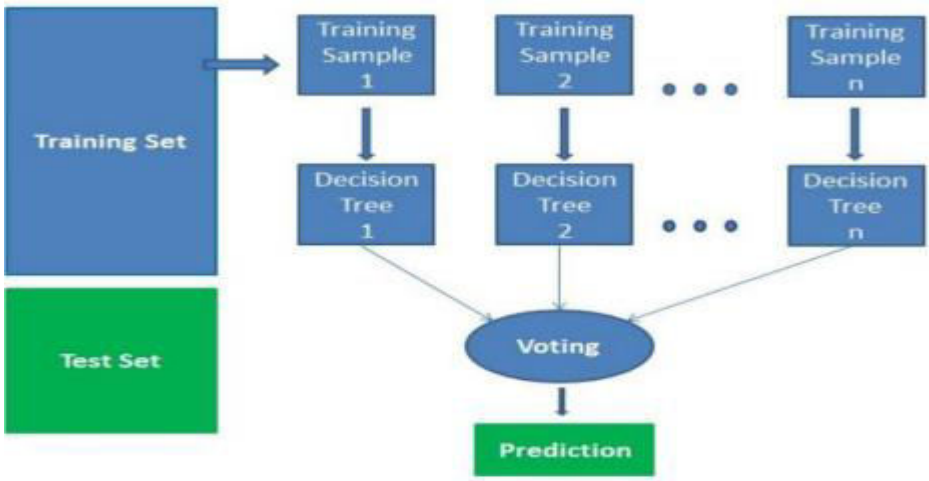

Figure 2: Random Forest Classifier

Random Forest Classifier: The term "random forest classifier" refers to a set of decision tree classifiers. We receive the results based on a voting process that is based on the majority. Here are the steps:
1. Choose a random sample from the provided dataset.
2. A decision tree is built for each sample that is there, producing a prediction result for each sample.
3. Votes have been cast on each forecast outcome.
4. Select the outcome that received the most votes and was expected.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107048 |

IV. EXPERIMENTAL RESULTS

It is the outcome of a false job advertisement, which indicates that the information about a position is fake. The information you provided regarding a job such as experience, credentials, employment, corporate logo, etc. is incorrect. It is the outcome of a false job advertisement, which indicates that the information about a position is fake. The information you provided regarding a job such as experience, credentials, employment, corporate logo, etc. is incorrect.

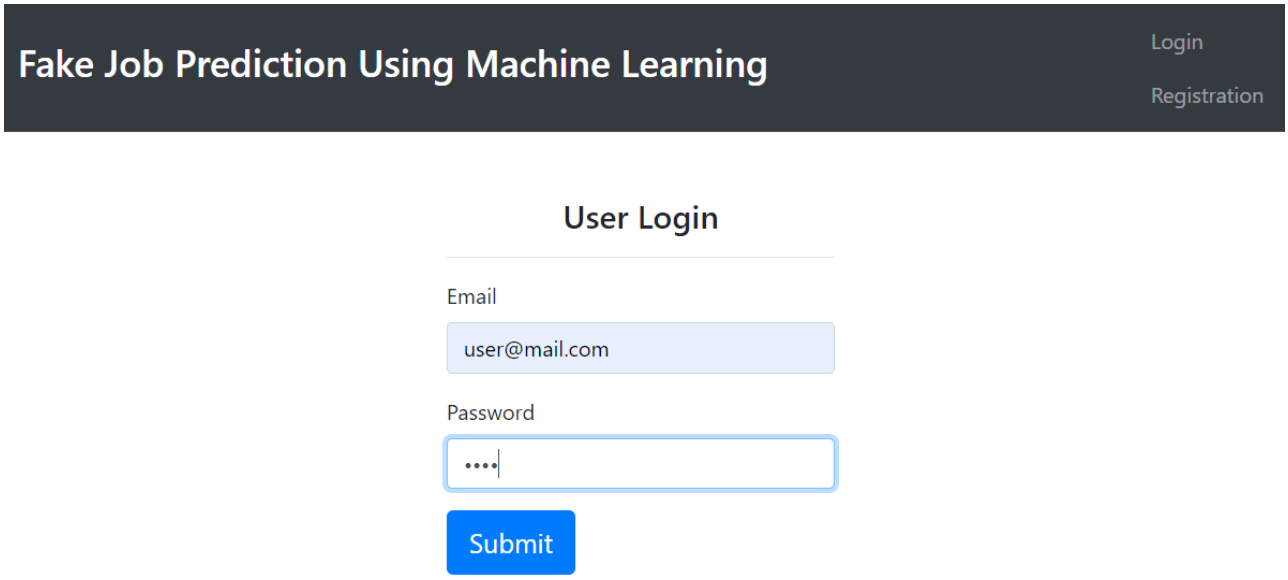

Figure 3: Login Page

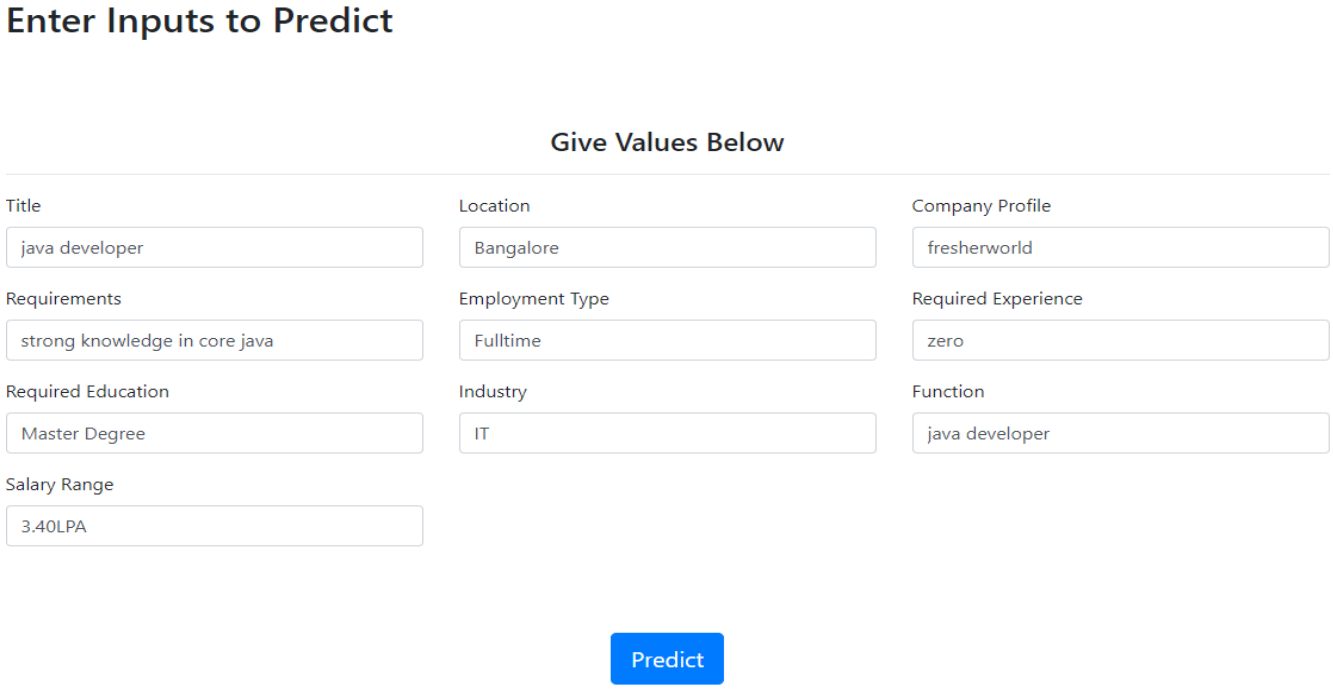

Figure 4: Fake Job Post

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107048 |

When you click submit after inputting all the information including the company name, employment type, experience, qualifications, etc. it makes a prediction and displays the outcomes. The information you provided is therefore about a valid job posting

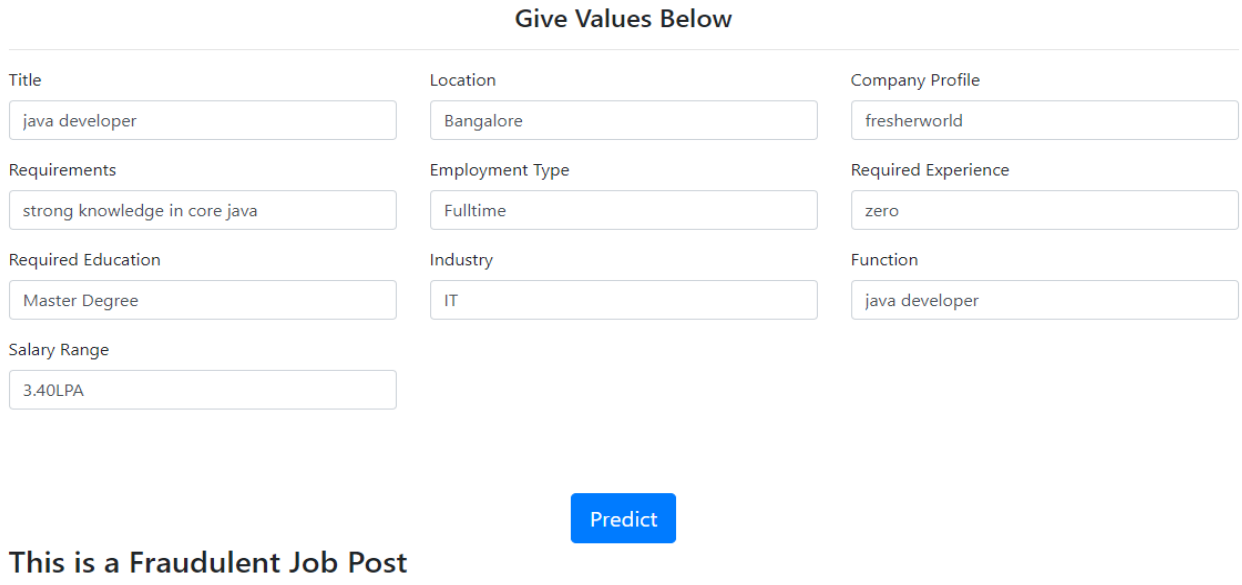

Figure 5: Fraudulent Job Post

If you get the job post in the form of message or job details described in the message format, then paste the text in the textbox and click on predict to get the results. In comparison to previously employed algorithms like SVM, Decision Tree Classifier, etc., which provide accuracy of 90%, Random Forest Classifier provides accuracy of 98%. By minimizing job fraud and scams, we are making the online hiring process safer. As a result, you might choose to apply for employment online.

V. CONCLUSION

The detection of work scams has recently grown to be a significant issue on a global scale. In this project, we looked at the effects of employment scams because they might be a very profitable area of research and make it challenging to spot false job postings. We used the EMSCAD dataset, which contains current job advertisements. In comparison to previously employed algorithms like SVM, Decision Tree Classifier, etc., which provide accuracy of 90%, Random Forest Classifier provides accuracy of 98%. By minimizing job fraud and scams, we are making the online hiring process safer. As a result, you might choose to apply for employment online. Consequently, preventing a person's financial losses and protecting that person's personal information.

REFERENCES

[1] S. Vidros, C. Kolias , G. Kambourakis ,and L. Akoglu, "Automatic Detection of Online Recruitment Frauds: Characteristics, Methods, and a Public Dataset", Future Internet 2017, 9, 6; doi:10.3390/fi9010006A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting.", IEEE Transactions on Image Processing, vol. 13, no.9, pp. 1200–1212, 2004.

[2] B. Alghamdi, F. Alharby, "An Intelligent Model for Online Recruitment Fraud Detection Journal of Information Security, 2019, Vol 10, pp. 155176,.

[3] Tin Van Huynh1, Kiet Van Nguyen, Ngan Luu-Thuy Nguyen1, and Anh Gia-Tuan Nguyen, "Job Prediction: From Deep Neural Network Models to Applications", RIVF International Conference on Computing and Communication Technologies (RIVF), 2020T.Eftychios A. Pnevmatikakis, Petros Maragos "An Inpainting System For Automatic Image Structure-Texture Restoration With Text Removal", IEEE trans. 978-1-4244-1764, 2008

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107048 |

[4] Jiawei Zhang, Bowen Dong, Philip S. Yu, "FAKEDETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network", IEEE 36th International Conference on Data Engineering (ICDE), 2020.Aria Pezeshk and Richard L. Tutwiler, "Automatic Feature Extraction and Text Recognition from Scanned Topographic Maps", IEEE Transactions on geosciences and remote sensing, VOL. 49, NO. 12, 2011

[5] T. Van Huynh, V. D. Nguyen, K. Van Nguyen, N. L.-T. Nguyen, and A.G.-T. Nguyen, "Hate Speech Detection on Vietnamese Social Media Text using the Bi-GRU-LSTM-CNN Model," arXiv Prepr. arXiv1911.03644, 2019

[6] Manjunatha HT and AjitDanti, "Segmentation of Traffic Sign Board in a cluttered background using Using Digital Image Processing", National Conference on Network Security, Image Processing and Information Technology, March 2017.

[7] Manjunatha HT and AjitDanti," Indian traffic sign board recognition using Normalized Correlation Method", International Journal of Computer Engineering and Applications (IJCEA), Volume XII, Issue III, March 18, ISSN 2321-3469

[8] Manjunatha HT and AjitDanti. "Detection and Classification of Potholes in Indian Roads using Wavelet Based Energy Modules "IEEE- 978-1-5386-9319-3/19© 2019, SCOUPS Nature

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details