Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.10 : 2024
ISSN : **1906-9685**

Fake job detection Using Machine Learning

Mr. Y.Venkatesh, MCA Student, Department of Master of Computer Applications,
Vignan's Institute of Information Technology (A), Visakhapatnam-530049.

Abstract
The article provides an automated method that makes use of machine learning-based categorization approaches to stop fake jobs from appearing online. To find the optimum model for job fraud detection, fraudulent communications on the Internet are analyzed using several categories, and the results are compared. It is useful for spotting phony job posts among a big volume of postings. To identify fraudulent works, two primary classifier types are used: individual classifiers and group classifiers. Nonetheless, the outcomes of the experiment indicate that group classifications work best for detecting fraud in a single class. In this project, we do a comparative study using logistic regression, random forest, SVM, and Naviebayes classifiers.

Keywords
- Ensemble Classifiers
- Machine Learning
- Fraudulent Job
- Fraud Detection

Introduction

Today, job searchers have a ton of options for new, flexible careers because to the advancement of business and technology. Through these positions, job searchers investigate options according to their availability, skills, background, aptitude, etc. These days, the power of social media and the internet affects the hiring process. Social media has a significant impact on recruiting as it is a publicity-driven process that determines its success. Advertisements in electronic and social media media have opened up ever-new avenues for the dissemination of employment information. However, the number of job searchers who are harassed by scammers has surged due to the quick expansion of job sharing chances. Because they preserve the security and continuity of their personal, intellectual, and professional information, individuals do not express desire in taking on new employment. Therefore, the true motivation behind qualifying job postings sent via electronic and social media is a very challenging task: winning people over to your legitimacy and trust. The technology we use every day make life simpler and more advanced, but they do not produce a dangerous workplace. Hiring more people will advance significantly if occupations can be accurately vetted and predicted to be incorrect. False employment is a major time waster and makes it difficult for job seekers to get the position they desire. An automated technology that detects fraudulent employment creates a new avenue for HR issues.

A. Fake Job Posting:

Workplace Fraud Job scams are fraudulent online job advertising that often aim to get personal and professional information from job seekers rather than match them with a position. Scammers sometimes attempt to get money from job seekers unlawfully. More than 67% of individuals are at

high risk of searching for employment via internet ads, yet they are ignorant of phony jobs or job scams, according to a new survey conducted by Action Fraud in the UK. Nearly 700,000 job searchers in the UK reported losing over $500k as a result of falling for a scam. According to the survey, there has been an almost 300% rise in the UK over the last two years. Since they are often looking for a permanent job for which they are ready to spend additional money, students and new graduates are the ones that fall victim to scams the most. Techniques for preventing or protecting against cybercrime won't stop it since thieves often alter their methods of operation.

.

B. Common Types of Job Scams:

Fraudsters fabricate occupations in order to get personal information belonging to other individuals, including date of birth, ID, bank account, insurance, and income tax information. Advance payment scams happen when con artists demand money and provide excuses such expenses for security audits, administrative expenses, etc. Fraudsters may take on the role of employers and request documents such as driver's licenses, bank records, and passports before hiring someone. Verify. Scams involving illegal withdrawals happen when they deceive students into making payments into their accounts and then taking them out. Working with cash is tax-free thanks to this "cash" approach. To fool job searchers, fraudsters often fabricate official papers, bank websites, and corporate websites. Most employment scams attempt to deceive victims using email rather than in-person interactions. Typically, they aim to position themselves as headhunters or recruiting agency on social networking platforms such as LinkedIn. Usually, they make an effort to provide the job seeker the most accurate representation of the firm profile or websites. The job seeker is always aiming for a trap, obtaining information and benefits whether or not to earn money, regardless of the kind of employment scams they use..

Literature Survey

A collection of writings intended to analyze important facets of the state of the art and/or methodological stances on a certain subject is called a literature review. It is a kind of secondary source that deals with material that has been published on a certain topic and sometimes information that has been collected over time on a particular subject. Its ultimate goal is to keep the reader informed about the most recent research on the topic and to provide a foundation for future, perhaps justifiable study in the area that predates the research proposal. concise synopsis. from the original sources. It often follows an organized format and include both synthesis and summary. A synthesis is an informational rearrangement, while a summary is a summary of significant information from a source. It might provide a fresh perspective on antiquated information, fuse a modern and historical view, or chart significant debates and other intellectual advancements in the topic. A literature review may assess sources and advise the reader on the most pertinent or relevant sources, depending on the circumstances.

Review of Literature Survey

Title : Predicting of Job Failure in Compute Cloud Based on Online Extreme

Learning Machine: A Comparative Study

Author: CHUNHONG LIU1, 2. JINGJING HAN2 , YANLEI SHANG1 ,

CHUANCHANG LIU1 , BO CHENG1 ,AND JUNLIANG CHEN1

Year: 2017

Encouraging operational breakdowns ahead of time and taking targeted efforts to eliminate them may greatly increase resource use efficiency. The offline operational model that is often utilized by current machine learning-based forecasting techniques is unsuitable for live forecasting in real-world scenarios where data comes sequentially. This study suggests a novel approach to forecast the end state of a network task, based on Network Sequential Extreme Learning Machine (OS-ELM) technology, in order to address this issue. Using this approach, real-time data is gathered based on the arrival sequence of the works, allowing for the prediction of the work's status and the updating of

the operational model. The incremental learning technique combined with the onlinear method allows for quick learning and strong generalization. When the model is updated in 0.01 seconds, the suggested strategy has a 93% prediction accuracy, according to a comparison study conducted using Google tracking data. The method developed in this paper has many advantages over some state-of-the-art methods, including better false negative performance, less time-consuming model generation and updating, and higher prediction accuracy and precision compared to Extreme Learning Machine (ELM), Network Sequence Support Vector Machine (OS-SVM), and Support Vector Machine (SVM).

Title: Machine Learning and Job Posting Classification: A Comparative Study

Author: Ibrahim M. Nasser1 and Amjad H. Alzaanin2

Year: 2020

For the text classification issue, we examine a number of machine learning classifiers in this study, including Multinomial Naive Bayes, Support Vector Machine, Decision Tree, K Nearest Neighbors, and Random Forest. Right and bad jobs are included in the information we utilize. After cleaning and preprocessing our data, we extracted features using TF-IDF. We trained and assessed the classifiers once they were deployed. The precision, recall, f-measure, and precision are evaluation metrics. Every classifier's output was compiled and contrasted with the others.

Conclusion

These days, detecting labor fraud is a big problem all over the globe. In this piece, we examined the effects of work fraud, which is a fruitful field of study even though it may be difficult to spot fake employment. We examined the EMSCAD dataset, which included actual fraudulent job postings. In this paper, we evaluated a deep learning model (Deep Neural Network) as well as many machine learning methods (SVM, KNN, Naïve Bayes, Random Forest, and MLP). In this paper, classifiers based on deep learning and conventional machine learning are compared and evaluated. We discovered that the Random Forest Classifier has the best classification performance among traditional machine learning techniques, with 99% accuracy (time 9) for DNNs and 97.7% average for deep neural networks.

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/fi9010006.

[2] B. Alghamdi, F. Alharby, “An Intelligent Model for Online Recruitment Fraud Detection”, Journal of Information Security, 2019, Vol 10, pp. 155176, https://doi.org/10.4236/iis.2019.103009 .

[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), 2020.

[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.

[5] Scanlon, J.R. and Gerber, M.S., “Automatic Detection of Cyber Recruitment by Violent Extremists”, Security Informatics, 3, 5, 2014, https://doi.org/10.1186/s13388-014-0005-5

[6] Y. Kim, “Convolutional neural networks for sentence classification,” arXiv Prepr. arXiv1408.5882, 2014.

[7] 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.

[8] P. Wang, B. Xu, J. Xu, G. Tian, C.-L. Liu, and H. Hao, “Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification,” Neurocomputing, vol. 174, pp. 806814, 2016.

[9] C. Li, G. Zhan, and Z. Li, “News Text Classification Based on Improved BiLSTM-CNN,” in 2018 9th International Conference on Information Technology in Medicine and Education (ITME), 2018, pp. 890-893.

[10] K. R. Remya and J. S. Ramya, “Using weighted majority voting classifier combination for relation classification in biomedical texts,” International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014, pp. 1205-1209.

[11] Yasin, A. and Abuhasan, A. (2016) An Intelligent Classification Model for Phishing Email Detection. International Journal of Network Security& Its Applications, 8, 55-72. https://doi.org/10.5121/imsa.2016.8405

[12] Vong Anh Ho, Duong Huynh-Cong Nguyen, Danh Hoang Nguyen, Linh Thi-Van Pham, Duc-Vu Nguyen, Kiet Van Nguyen, and Ngan Luu-Thuy Nguyen."Emotion Recognition for Vietnamese Social Media Text", arXiv Prepr. arXiv:1911.09339, 2019.

[13] Thin Van Dang, Vu Duc Nguyen, Kiet Van Nguyen and Ngan Luu-Thuy Nguyen, “Deep learning for aspect detection on vietnamese reviews” in In Proceeding of the 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), 2018, pp. 104-109.

[14] Li, H.; Chen, Z.; Liu, B.; Wei, X.; Shao, J. Spotting fake reviews via collective positive-unlabeled learning. In Proceedings of the 2014 IEEE International Conference on Data Mining (ICDM), Shenzhen, China, 14-17 December 2014; pp. 899-904.

[15] Ott, M.; Cardie, C.; Hancock, J. Estimating the prevalence of deception in online review communities. InProceedings of the 21st international conference on World Wide Web, Lyon, France, 16-20 April 2012; ACM: New York, NY, USA, 2012; pp. 201-210.

[16] Nizamani, S., Memon, N., Glasdam, M. and Nguyen, D.D. (2014) Detection of Fraudulent Emails by Employing Advanced Feature Abundance. Egyptian Informatics Journal, Vol.15, pp.169-174. https://doi.org/10.1016/j.eij.2014.07.002