UtilitasMathematica

ISSN 0315-3681 Volume 122 (2), 2025

ONLINE RECRUITMENT FRAUD DETECTION USING AI PATTERNS

Turlapati Sahithi [1], Vutukuri Padmaja [2], Devarasetti Prasad [3]

[1]M.Tech Scholor- CSE Department, DVR & Dr HS MIC College of Technology, Kanchikacherla, NTR District, AP,India.

[2]Assistant Professor, CSE Department, DVR & Dr HS MIC College of Technology, Kanchikacherla, NTR District, AP, India

[3] Professor & HOD of CSE Department, DVR & Dr HS MIC College of Technology, Kanchikacherla, NTR District, AP, India

ABSTRACT: The world population grows and demand for workers increases, leading to a rise in online job advertisements to connect employers with potential employees on a national scale. Machine learning is a powerful tool for finding complex financial security threats that constantly evolve and can be difficult to predict. The findings contribute to a deeper understanding of AI's capabilities and limitations providing insights that can guide the development and deployment of AI driven security systems. The fake jobs is precisely detected and classified from a pool of job posts of both fake and real jobs by using advanced deep learning as well as machine learning classification algorithms. The intervention of AI not only automates a particular task is improves efficiency by many folds. Cybercriminals is targets from globe time. In strategic decision making is important and logical decision model is one of the still unanswered cyber security method. Machine learning algorithms are constantly being improved to identify error data that might indicate a security threats. We used four classes of features: empirical rule set-based features, bag-of-word models, most recent state-of-the-art word embedding and transformer models for various machine learning classifier. The machine learning models were validated by evaluating them on a publicly available job description dataset

Index Terms: Artificial Intelligence, Fraud detection · Online recruitment fraud · word2vec · Transformers · Machine learning · Natural language processing, support vector machine, deep learning, and classification

1. INTRODUCTION

The global internet penetration is increased the advancement of web and telecommunication technology. Approximately 4901 million people as of 2023 is connected to the internet marking a significant shift is individuals interact and seek opportunities [1]. Advisory organizations such as the National Institute of Standards and Technologies (NIST) are also encouraging the use of more proactive and adaptive approaches by shifting towards real-time assessments, continuous monitoring and data-driven analysis to identify, protect against, detect, respond to, and catalogue cyber-attacks to prevent

future security incidents [2]. Fake job posts create inconsistency for the job seeker to find their preferable jobs causing a huge waste of their time. An automated system to predict false job post opens a new window to face difficulties in the field of Human Resource Management [3]. National Institute of Standards and Technologies (NIST) are also encouraging the use of more proactive and adaptive approaches by shifting towards real-time assessments, continuous monitoring, and data-driven analysis to identify, protect against, detect, respond to, and catalogue cyber-attacks to prevent future security incidents [4]. We develop machine learning system for identifying different fraudulent job advertisements. The fraudulent jobs were conceptualized after a thorough analysis of the literature: identity theft, corporate identity theft, and pyramid schemes or multi-level marketing [5]. Even when the most robust preventive measures are in place, hackers will attempt to circumvent them. It is doubtful that cyber dangers will ever be fully eradicated since hackers are clever and persistent, always looking for new methods to penetrate a company's defenses [6]. This phenomenon presents both positive economic growth indicators and potential risks for job seekers. While the rise in job postings fosters economic opportunities[7].

Figure 1 AI in Cyber Security

2. RELATED WORKS

The search results which we got were limited only to the papers published in the last four years, as the determination of this paper is to bring out the

newest trends of AI in cyber security. Last, the findings were categorized by the number of certifications [8]. A well-known cyber security framework proposed by NIST was used to understand the solution categories needed to protect, detect, react and defend against cyber-attacks [9]. The NIST cyber security framework's core describes the practices to improve the cyber security of any organization [10]. The International Labour Organization (ILO) defines forced labour as work demanded under threat or penalty, not voluntarily offered by an individual (ILO, 2023). Fraudulent job recruitment often involves luring jobseekers into providing personal information which is then resold to third parties, leading to spam emails and further fraud [11]. Cyber security is a broad term encompassing all measures taken in an effort to safeguard an entity from cyber threats, including securing data and mitigating damage from a cyber-security incident [12]. The cyber security landscape should include measures to protect the organization from crypto-jacking, data leaks, data phishing, and Internet of Things threats (IoT). One should use experienced IT professionals and ethical hackers to guarantee that your AI security solution is impenetrable [12]. One of the main game changers in the area of cyber security is the development of tools and methods that are supplemented as a sub-group by artificial intelligence (AI) [13]. The Majority Vote (MV) ensemble

learning algorithm takes the majority vote of the predictions of these base classifiers to make the prediction.

3. SYSTEM MODELS

These tasks include removing irrelevant and duplicate data, handling missing values, and cleaning nominal and text data. For nominal data cleaning, label encoding is employed to handle categorical variables [13]. The functions provide a comprehensive view of the lifecycle for managing cyber security overtime [14]. The proposed botnet detection model based on machine learning using DNS query data. The model is built on the analysis that Threats of CS Threats routinely send lookup queries to the DNS system to find IP addresses of servers using automatically generated domain names [15]. In practical terms, AI refers to a number of different technologies are used in a variety of ways. The data extraction and pre-processing phase, data was

extracted from missing values were addressed by removing columns with a missing value percentage of 70% or more [16].

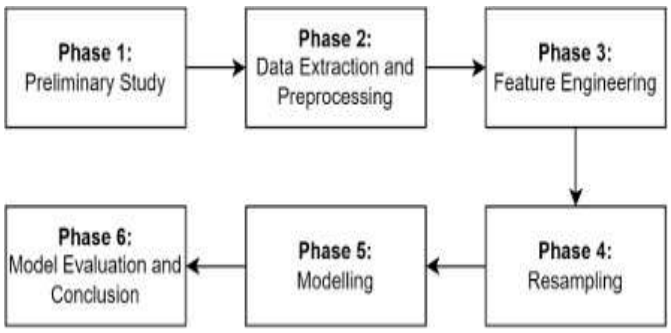


Fig 2. System Architecture..

4. PROPOSED SYSTEM

In our proposed model uses to different machine learning techniques and classification algorithm like KNN, decision tree, support vector machine, naive bayes classifier, random forest classifier, multilayer perceptron and deep neural network to predict a job post if it is real or fraudulent. We select and view the imported dataset for future purpose and we get missing values and fill the default values to the dataset. We encode the label in the dataset. And we split the dataset to the Train and Test data for predict the fraud or non-fraud [17]. There are Random

forest algorithm, KNN classifiers and Ada-Boost Algorithm. Now, we fit the training data from the dataset.

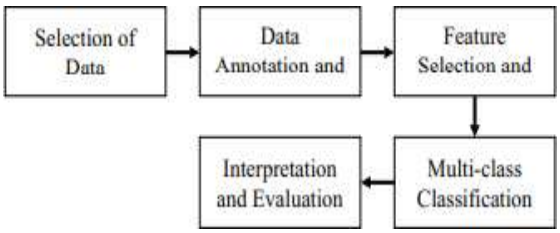


Figure 3 Research design detection.

5. Algorithms:

For the prediction, multiple supervised learning algorithms are trained using

the training set, after which using the testing set performance evaluation

A. **Random**

Forest STEP 1:

START

STEP 2: SPLIT dataset into 67 percent training set, 33 percent testing set

STEP 3: FOR train dataset CALL RF Classifier TRAINRF Classifier

STEP 4: FOR test dataset CALL RF Classifier PREDICT the label COMPUTE Accuracy Score SAVE Accuracy Score DISPLAY Confusion Matrix

STEP 5: STOP

B. **KNN Classifier**

Train() Input: train set,

test set **Output:** Trained

model

Step 1: Read Train set and test

set **Step 2**: Build KNN classifier

Step 3: Train the model using

fit()

Step 4: Performance Graph Returned Trained Model

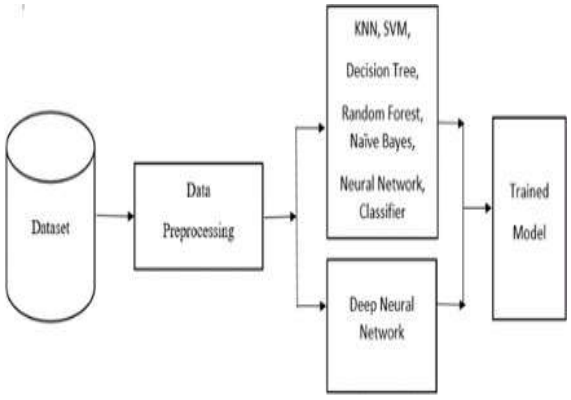


Fig. 4. Multiple supervised learning algorithms.

C. **Naïve Bayes**

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.

STEP 1: START

STEP 2: SPLIT dataset into 67 percent training set, 33 percent testing set

STEP 3: FOR train dataset CALL Multinomial NB TRAIN Multinomial NB

STEP 4: FOR test dataset CALL Multinomial NB PREDICT the label COMPUTE Accuracy Score SAVE Accuracy Score DISPLAY Confusion Matrix

STEP 5: STOP

D. **Support Vector Machine (SVM)**

The SVM method is based on the Vapnik Chervonenkis (VC) dimension theory of statistical learning theory and the principle of structural risk minimization.

STEP 1: START

STEP 2: SPLIT dataset into 67 percent training set, 33 percent testing set

STEP 3: FOR train dataset CALL SVM Classifier TRAIN SVM Classifier

STEP 4: FOR test dataset CALL SVM Classifier PREDICT the label COMPUTE Accuracy Score DISPLAY Confusion Matrix

STEP 5: STOP

6. INTERVENTION OF AI

The use of AI in cyber-attacks is a new and emerging trend. It is not yet clear how this will affect the future of cybercrime. There are several different AI and machine learning techniques used in cyber security. The most common ones include strategies that use AI to identify and monitor malicious activities, detect cyber threats, and protect an organization's networks [21]. Cyber security professionals who can adopt successful cognitive technologies and guide their human element on a holistic approach will be more successful in defending against cyber-attacks [22].

AI based mitigation of Cyber threats

Malware detection and identification: Artificial intelligence being used for malware detection and identification is still in its infancy, but it has the potential to revolutionize the way we deal with cybercrime. AI can help identify malicious files before they reach the end-user and, by doing so, can provide significant security benefits. Many different AI/ML approaches have been used to detect malware, some more successful than others [23].

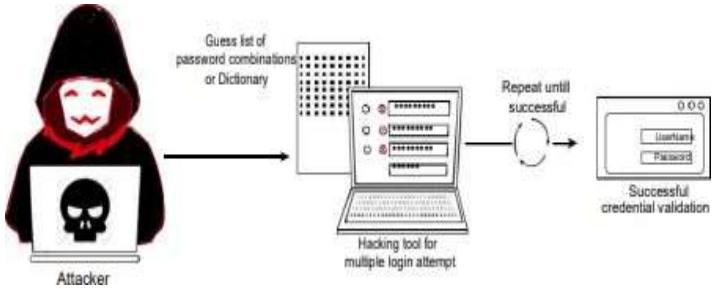
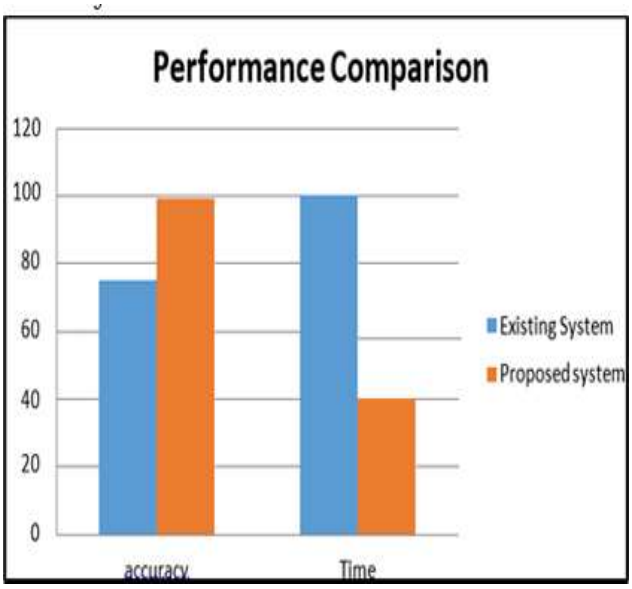


Fig. 5. Depiction of generalized principle of a Password Attack

7. PERFOMANCE METRICS

The accuracy is determined by accuracy score of the ML models, which is measured in percentage. The average time taken is determined by comparing the evaluation time taken for training and prediction by a model. The performance of



this proposed approach is evaluated using some measures like, Accuracy: Accuracy of classifier refers to the ability of classifier. It predicts the class label correctly and the accuracy of the predictor refers to how well a given predictor can guess the value of predicted attribute for a new data. Even though the organization's information is essential to any organization, the cost of implementing AI technology is much higher, limiting the number of individuals who will use the technology for the safety of their data and information.

Fig. 6. Distribution of use of AI domain during

8. CONCLUSIONS AND FUTURE OPPORTUNITIES

The main contribution of this research lies in the development and evaluation of an effective prediction model for fraudulent job advertisement classification. In a scenario where malicious intelligence and cyber threats are rising exponentially, sophisticated cyber security strategies cannot be ignored. Also security against large-scale threats, with very minimal resources, has been demonstrated smart approaches are used. The evolution of AI in cyber security was studied with respect to different functions, solution categories, specific use cases, and the type of AI technique used. The way they get these probabilities is by using KNN, which describes the probability of a feature which has misclassification and less prediction. In this proposed model, initially for training 80% data is being used, and for testing 20% of data are pre-processed.

The Classifications classifier gives high accuracy results that are comparable or superior to other fraud detection techniques in spite of working with reduced data and also compared with graph. Research shows that artificial intelligence has seemingly positively affected cyber security and risks. The continuation of AI and machine learning will take the cyber security field to a new level of intelligence. In future discovery of additional information based on cause-event Fraud detection well as prediction of detection based on cause events. The working of the proposed approach in a web application

9. REFERENCES

[1] Bhardwaj, M.D. Alshehri, K. Kaushik, H.J. Alyamani, M. Kumar, Secure framework against cyber-attacks on cyber-physical robotic systems, J. Electron. Imaging 31 (6) (2022), 061802-061802.

[2] P. Chithaluru, A.T. Fadi, M. Kumar, T. Stephan, Computational intelligence inspired adaptive opportunistic clustering approach for industrial IoT networks, IEEE Internet Things J (2023), https://doi.org/10.1109/JIOT.2022.3231605.

[3] M. Barrett, Technical Report, National Institute of Standards and Technology, Gaithersburg, MD, USA, 2018.

[4] I. Wiafe, F.N. Koranteng, E.N. Obeng, N. Assyne, A. Wiafe, S.R. Gulliver, Artificial intelligence for cybersecurity: a systematic mapping of literature, IEEE Access 8 (2020) 146598–146612.

[5] Z. Zhang, H. Ning, F. Shi, F. Farha, Y. Xu, J. Xu, F. Zhang, K.K.R. Choo, Artificial intelligence in cyber security: research advances, challenges, and opportunities, Artif. Intell. Rev. 55 (2022) 1029–1053.

[6] J. Martínez Torres, C. Iglesias Comesana, ˜ P.J. García-Nieto, Machine learning techniques applied to cybersecurity, Int. J. Mach. Learn. Cybern. 10
(10) (2019) 2823–2836.

[7] T.C. Truong, I. Zelinka, J. Plucar, M. Candík, ˇ V. Sulc, ˇ Artificial intelligence and cybersecurity: past, presence, and future, in: Artificial

intelligence and evolutionary computations in engineering systems, 2020, pp. 351–363.

[8] S. Samoili, M.L. Cobo, E. Gomez, G. De Prato, F. Martinez-Plumed, B. Delipetrev, A.I. Watch, Technical report, Joint Research Center (Seville site), 2020.

[9] High-Level Expert Group on Artificial Intelligence. (HLEG AI), A definition of AI: main capabilities and disciplines, (2019). Retrieved from Brussels https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341.

[10] D. Zhao, A. Strotmann, Analysis and visualization of citation networks, Synthesis lectures on information concepts, retrieval, and services, 7 1 (2015) 1–207.

[11] M. Zajko, "Canada's cyber security and the changing threat landscape", Critical Studies on Security, vol. 3, no. 2, pp. 147-161, 2015.

[12] R. Winkels, Eleventh International Conference on Artificial Intelligence and Law: proceedings: June 4-8, 2007, Stanford Law School, Stanford, California. Place of publication not identified: ACM, 2007.

[13] H. Bidgoli, Handbook of information security. Hoboken, NJ: John Wiley, 2006.

[14] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology", Artificial Intelligence, vol. 175, no. 5-6, pp. 988-1019, 2011.

[15] C. Blackwell and H. Zhu, Cyberpatterns: Unifying Design Patterns with Security and Attack Patterns, 2nd ed. Cham : Springer International Publishing, 2014.

[16] A. R. Dengel, K. Berns, T. M. Breuel, F. Bomarius, and T. R. RothBerghofer, KI 2008: advances in artificial intelligence 31st Annual

[17] D. Feng, D. Lin, and M. Yung, Information Security and Cryptology First SKLOIS Conference, CISC 2005, Beijing, China, December 15- 17, 2005, Proceedings. Berlin: Springer, 2005.

[18] J. G. Siegel, The artificial intelligence handbook: business applications in accounting, banking, finance, management, marketing. Mason, OH: Thomson/South-Western, 2003.

[19] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology," Artificial Intelligence, vol. 175, no. 5-6, pp. 988–1019, 2011.

[20] J. R. Vacca, Computer and information security handbook. Waltham, MA, USA: Morgan Kaufmann Publishers, 2013

[21] Rajani, P., Adike, S., & Abhishek, S. G. K. (2020). ARTIFICIAL INTELLIGENCE : THE NEW AGE. 8(2), 1398–1403.

[22] Rosenblatt, F. (1957). The Perceptron - A Perceiving and Recognizing Automaton. In Report 85, Cornell Aeronautical Laboratory (pp. 460–461).

[23] Sadiku, M. N. O., Fagbohungbe, O. I., & Musa, S. M. (2020). Artificial Intelligence in Cyber Security. International Journal of Engineering Research and Advanced Technology,

[24] Shankarapani, M. K., Ramamoorthy, S., Movva, R. S., & Mukkamala, S. (2011). Malware detection using assembly and API call sequences. Journal in Computer Virology, 7(2), 107– 119.

[25] Tyugu, E. (2011). Artificial intelligence in cyber defense. 2011 3rd International Conference on Cyber Conflict, ICCC 2011 - Proceedings, 95–105.