Supervised Machine Learning Method for Anomaly Detection

Asma Salim Yahya*

Department of Software, College of Computer Science and Mathematics, University of Mosul, Nineveh, Iraq.

*Corresponding author: asma_alkhairi@uomosul.edu.iq

https://doi.org/10.29072/basjs.20240207

ARTICLE INFO

Keywords

Anomaly Detection, CNN algorithm, Cybersecurity, LSTM algorithm, Supervised Machine Learning.

ABSTRACT

Cybersecurity is an essential topic, as most of our daily activities are controlled by web applications. These applications become susceptible to various threats that lead to unauthorized access to personal data. Therefore, protecting application data has become essential. Supervised machine learning is widely utilized in various applications, such as spam detection; it functions as a powerful tool for automating decision-making and producing predictions based on historical data. This study employs supervised machine learning to classify anomalies in a network using the NSL-KDD dataset, which is utilized to assess intrusion detection techniques. This dataset contains no repeated items in the training subset, making the approach impartial to any particular items. This research utilizes approaches such as CNN, LSTM, hybrid CNN-LSTM, RBFN, MLP, and SVM. Evaluating multiple algorithms and analyzing their results to select the most efficient option is typically a wise strategy. The results of the implemented models were evaluated and compared based on detection rate, time efficiency, and accuracy. The findings demonstrate that the CNN-LSTM hybrid model exceeded the benchmark methods, with a detection rate of 99.61% and an accuracy of 99.8%.

Received 05 Apr 2024; Received in revised form 03 Jun 2024; Accepted 23 Aug 2024, Published 31 Aug 2024

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

1. Introduction

Intrusion Detection Systems (IDS) can be defined as systems that monitor the traffic of data within a system and distinguish abnormal transactions that may lead to unauthorized access to data [1]. The detection process depends on rules or previously defined events or suspicious activities [2]. The basic architecture of an IDS system is shown in Figure 1. The figure shows that the detection process needs information from many sources in order to detect anomalies [3,4]. The detection process is performed using a variety of approaches like statistical approaches, machine learning procedures, deep learning methods, etc. [5]. Other approaches in the literature combine the mentioned methods (hybrid) and suggest more efficient methods [6]. The performance of a method can be dignified using different metrics such as detection rate and time. Machine learning algorithms show a significant role in the ground of IDS. Many studies in the literature have done to improve the detection rate [7].

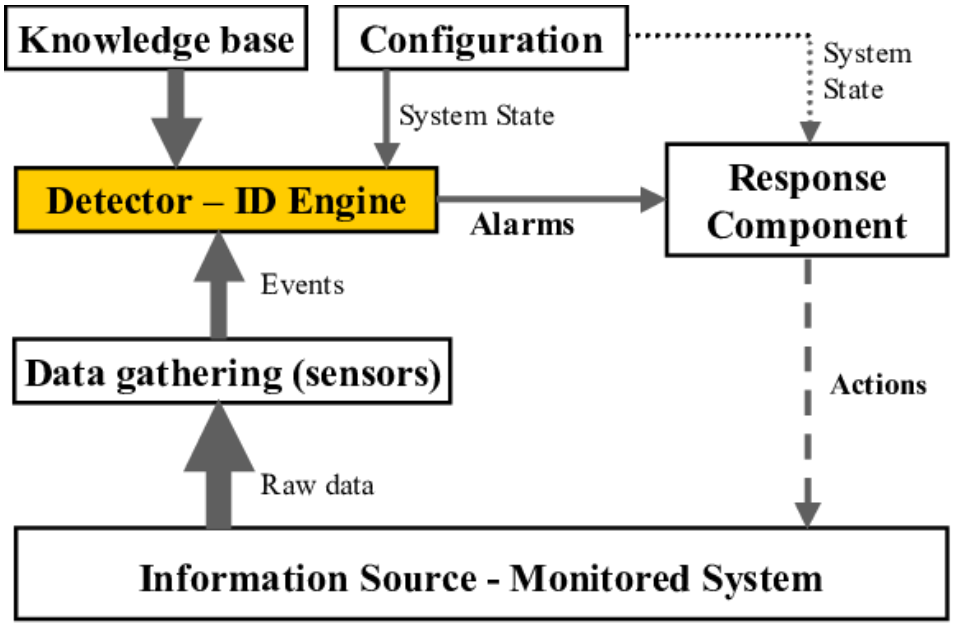

Figure 1: A typical IDS system [4].

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

2. Experimental

The researchers In [9] planned a machine learning-based IDS system for detecting network attacks. They involved six machine-learning algorithms that were mixed with the Principal Component Analysis (PCA). These algorithms are KNN, SVM, Cat Boost, XGBoost, and QDA. The results of the proposed system showed a superior performance with efficiency of accuracy of 99.9%. The dataset used in their work was UNSW-NB15 which includes network attacks with mixed activities. Moreover, the researchers in [10] combined Random Forest and Posterior algorithms to build an IDS system to improve the detection accuracy of the proposed method related to the traditional models in the literature. The machine learning and deep learning algorithms can compound to tackle issues in IDS systems. For instance, the work [11] proposed an IDS system that associations SMOTE approach and the XGBoost approach. The datasets used were CIC MalMem2022 KDD-CUP-99. Their results showed 99.99% of accuracy for the first and 100% for the second. In the literature, many works have been proposed to optimize the current machine-learning algorithms such as [12-16]. In a study performed by [17] ,various data stream methods were utilized on the CICIDS2017 datasets, encompassing multiple novel forms of attacks. The optimal algorithm that meets the requirements of high accuracy and short computation time was selected after the results. Moreover, the UNSW-NB15 dataset and CNN are applied by [18] to create a supervised network in order to save time and money, Recursive Feature Elimination (RFE) and Extreme Gradient Boosting (XGB). Also, bias toward the dataset's majority class is lessened via the Bayesian Gaussian Mixture Model (BGMM) and Synthetic Minority Oversampling Technique (SMOTE) with a 98.80% accuracy rate for binary classification and a 96.49% accuracy rate for classification into multiple categories, including the data demonstrate that this model outperforms existing techniques. One of the important problems in cybersecurity literature is the efficiency of the intrusion detection system that detects abnormal phenomena accurately and fast. Therefore, it is required to adopt a system that can detect anomalies in a network efficiently. This paper is about to propose a method that can do so. The remaining parts of this paper are: The research method proposed in this paper is illustrated in Section 2. The gotten results and deliberations are demonstrated in Section 3. Section 4 provides conclusions and some future works.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

3. Research Methodology

The main two steps of this work are extracting features from data and then performing the classification on the data. The dataset utilized in this work was NSL-KDD [19], which is used for purposes of testing methods of intrusion detection. This dataset does not have repeated items in the training subset, which makes the approach not biased to any of the items. The dataset has items related to attacks and normal activities that will be used to train and test the methods. It is worth mentioning that this dataset is widely used in the literature. The first step in this work was to extract the features, but before that, the data was pre-processed and partitioned to 80% for training and 20% for testing. The evaluation of the results was performed mainly using the detection rate (DR), which is calculated according to the following Eq. (1).

$$DR = \frac{TPR}{TPR(TPR+FNR)} \qquad (1)$$

Where the optimistic rate is denoted by TPR, and the untrue negative rate is signified by FNR. The other assessment metric used was accuracy, which can be calculated using the following Eq. (2).

$$Accuracy = \frac{TNR+TPR}{TNR+TPR+FNR+FPR} \qquad (2)$$

Where the total true negative rate is denoted by TNR and the total true confident rate is signified by TPR, the total false negative rate is denoted by FNR, and finally, the total false positive rate is denoted by FPR. The time consumption was also considered in the evaluation, which represents how long the detection takes. In deep learning, the Convolutional Neural Network (CNN) is considered an efficient algorithm for object recognition. It is widely used in recognizing patterns in data. It is also simple to implement as shown in Figure 2, adaptable, and consumes few parameters. In contrast, Long Short-Term Memory (LSTM) is another deep-learning algorithm resulting from Recurrent Neural Network (RNN). It is used to capture long-term dependencies in data objects. It is considered efficient and provides high accuracy. The architecture of LSTM is shown in Figure 3.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

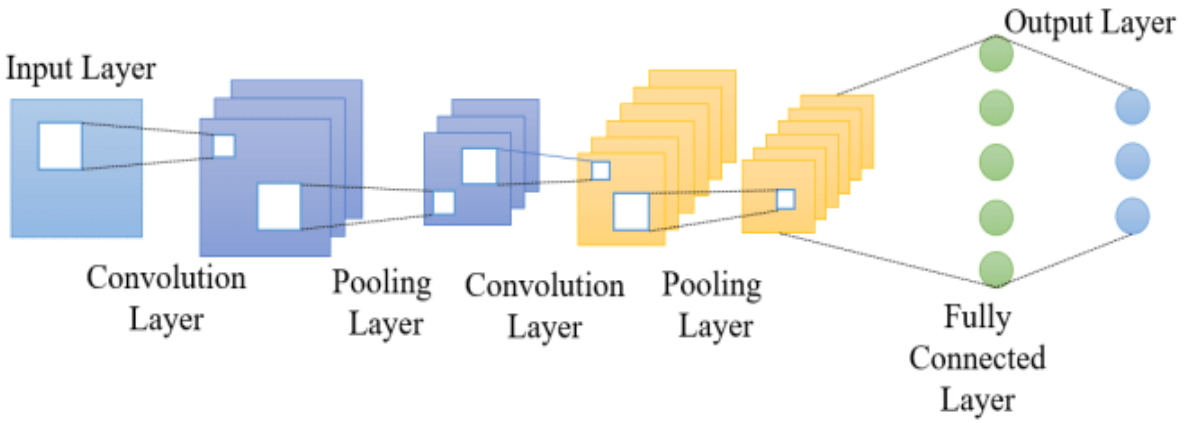

Figure 2: CNN typical architecture [20].

This work used a combination of CNN and LSTM models to generate one efficient model. The CNN-LSTM with multilevel feature extraction is the main model that is used to achieve the purpose of this research. The construction of the network is exposed in Figure 4. As known, CNN is not too sensitive to time step, but the case is different when dealing with LSTM because it is sensitive to each time step order. Also, to make the long sequence be recognized and shorter, the CNN can be involved as a preprocessing step by distinguishing the high-level features. Therefore, the CNN is used as a preprocessing layer to the LSTM layer. One of the most important features of this integration is that it provides a better amount of training time, which improves the efficiency and the whole performance of the method. Also, as proved in the literature, the accuracy of CNN-LSTM outperforms the accuracy of most of the available algorithms [20-26]. The general workflow of the suggested method is exposed in Figure 5.

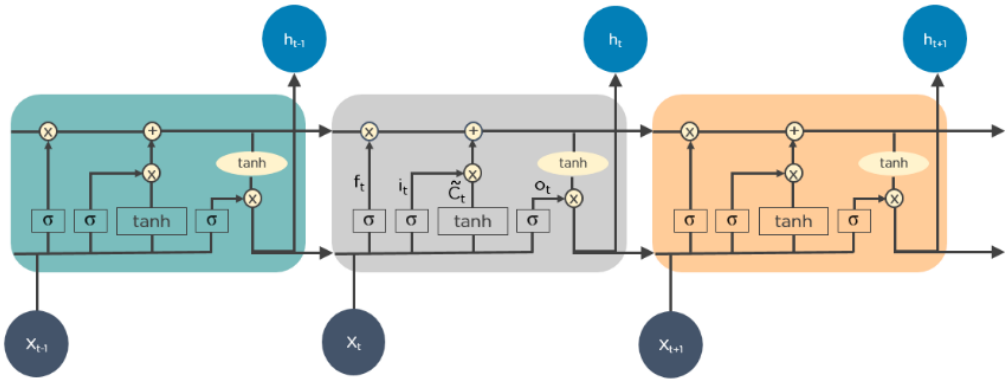

Figure 3: LSTM typical architecture [24].

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

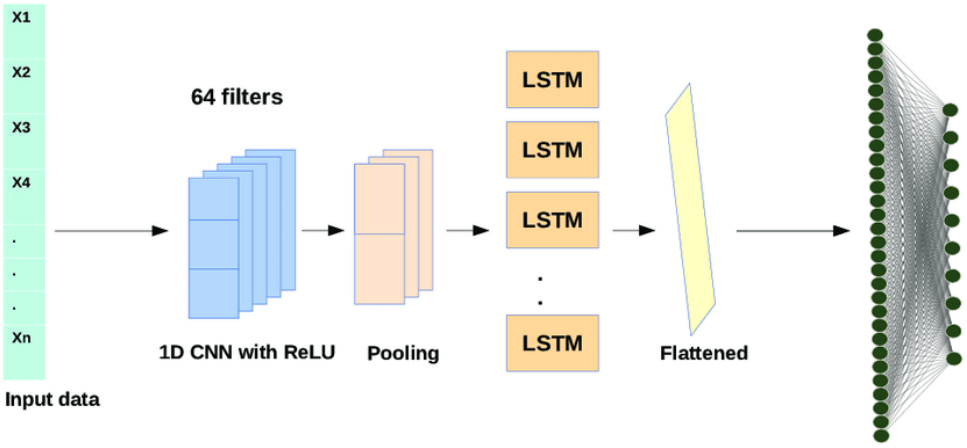

Figure 4: CNN-LSTM hybrid architecture [25].

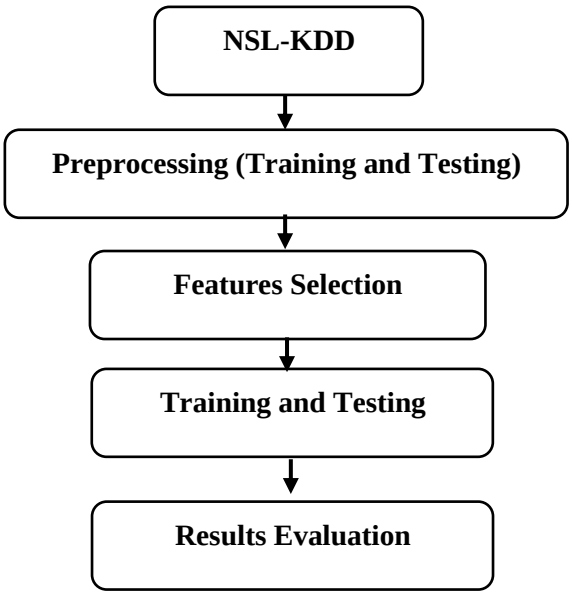

Figure 5: Workflow diagram of the proposed method.

4. Results and discussions

After the operation of the CNN-LSTM model, the results were benchmarked with other algorithms in the literature. The first benchmarking method is Radial Basis Function Networks (RBFN) [27], which includes the input layer, hidden layer, and output layer. The other benchmarking algorithm

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

is Multilayer Perceptron (MLP) as shown in figure (6) [28], which is the basic form of neural network with multilayer. The last benchmarking method is Support Vector Machine (SVM) [29].

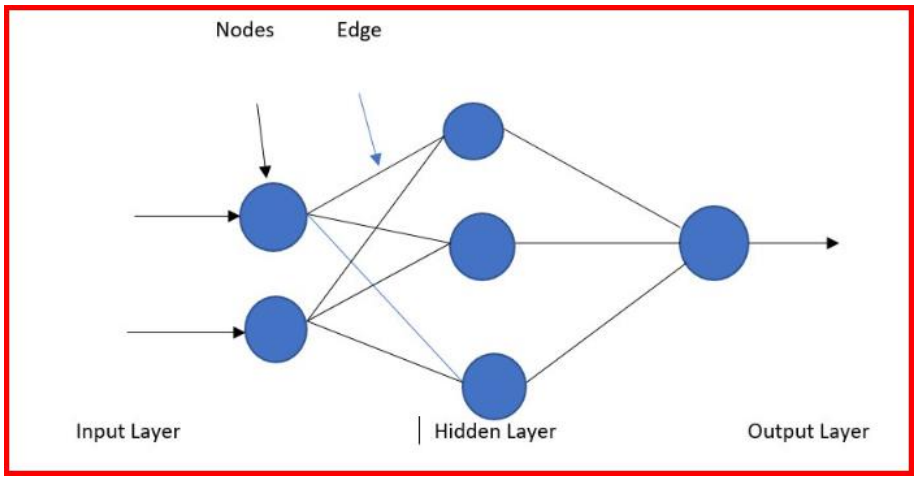

Figure 6: Multilayer perceptron (MLP) neural network.

The CNN and LSTM were implemented separately. In addition, the hybrid CNN-LSTM and the benchmarking methods were also implemented. Table 1 shows the findings of all the used methods. According to the table, the performance of the CNN-LSTM model overtakes the other benchmarking models in relations of detection rate and accuracy. However, the time consumed for CNN-LSTM was longer than the other models. This is because the combination of CNN and LSTM needs more computations and procedures. Also, the architecture of the CNN-LSTM is more complex than the other. On the other hand, the SVM model provides better performance in terms of time and the detection rate along with the accuracy are considered acceptable. The CNN and LSTM models provides higher detection rate and accuracy compared to the benchmarking, which means they are still good options for anomaly detection. The visualization of the presentation of each model in terms of DR, Correctness, and time are demonstrated in the Figures 7, 8, and 9 respectively.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

Table 1 - Performance of the models considered in this work.

Method	DR	Accuracy	Time
CNN	96.32%	94.8%	1.6 min
LSTM	97.45%	96%	1.8 min
CNN-LSTM	**99.61%**	**99.8%**	1.9 min
RBFN	68.92%	78.1%	1.7 min
MLP	95.73%	92%	1.2 min
SVM	93.27%	90.6%	**1.1 min**

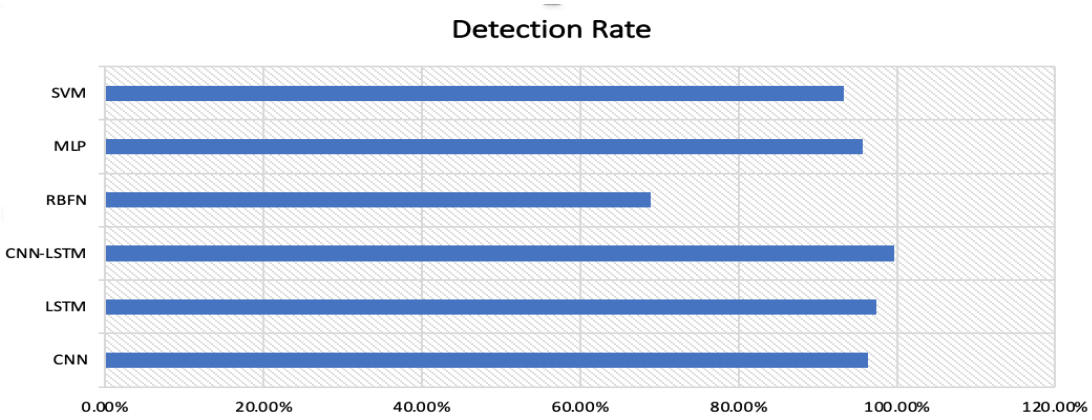

Figure 7: Visualization of the detection rate (DR) of all the models.

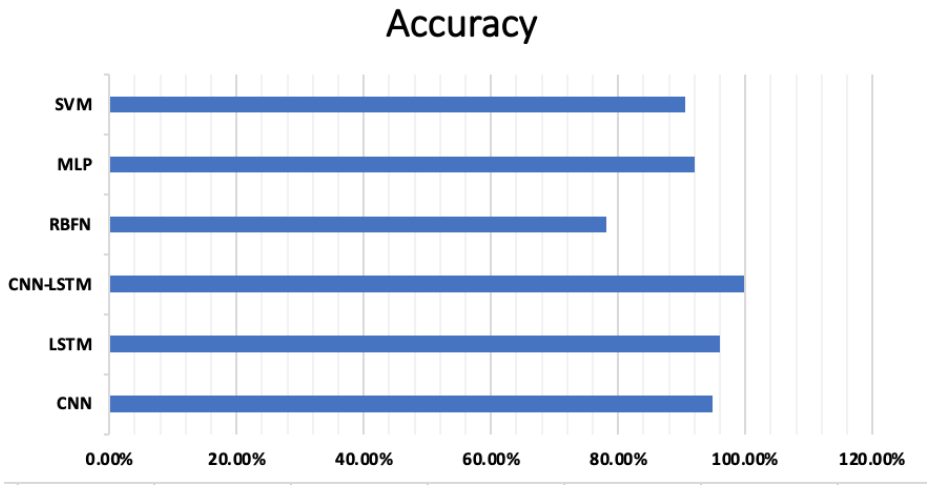

Figure 8: Visualization of the accuracy of all the models.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

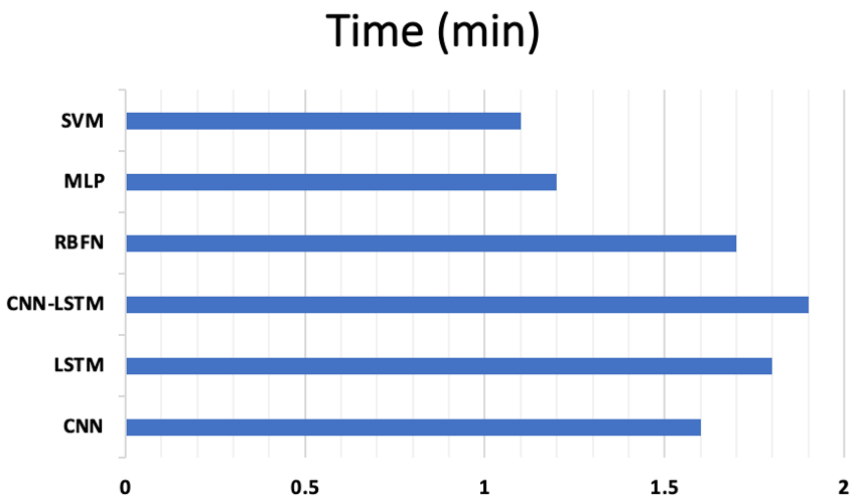

Figure 9: Visualization of the time consumption of all the models.

5. Conclusions

This study aims to employ machine learning methodologies for anomaly detection. To this end, numerous supervised machine learning models were employed to categorize anomalies inside a network utilizing the NSL-KDD dataset. The employed models included CNN, LSTM, hybrid CNN-LSTM, RBFN, MLP, and SVM. The efficacy of each model was assessed using three metrics: detection rate, accuracy, and time consumption. The outcomes of the executed models were assessed and compared against one another. The results indicate that the CNN-LSTM hybrid model surpassed the benchmark approaches for accuracy and detection rate. The most efficient model regarding time utilization was SVM. Future research should employ many models across various datasets to identify the most effective approach for intrusion detection tasks.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

References

[1] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, Survey of intrusion detection systems: techniques, datasets and challenges, Cybersecur, 1(2019)1-22, https://doi.org/10.1186/s42400-019-0038-7.

[2] Z. Ahmad, A. Sh. Khan, C. W. Shiang, J. Abdullah F. Ahmad, Network intrusion detection system: A systematic study of machine learning and deep learning approaches, T Emerg Telecommun T, 1(2021)4150, https://doi.org/10.1002/ett.4150.

[3] C. Zhang, D. Jia, L. Wang, W. Wang, F. Liu, A. Yang, Comparative research on network intrusion detection methods based on machine learning, Comput. Secur., 121(2022)102861, https://doi.org/10.1016/j.cose.2022.102861.

[4] I. Naqvi, A. Chaudhary, A. Kumar, A Systematic Review of the Intrusion Detection Techniques in VANETS, TEM J. 2(2022)900-907, https://doi.org/10.18421/TEM112-51.

[5] Z. Yang, X. Liu, T. Li, D. Wu, J. Wang, Y. Zhao and H. Han, A systematic literature review of methods and datasets for anomaly-based network intrusion detection, Comput. Secur., 16(2022) 102675, https://doi.org/10.1016/j.cose.2022.102675.

[6] A.K. Balyan, S. Ahuja, U.K. Lilhore, S. K. Sharma, P. Manoharan, A.D. Algarni, K. A. Raahemifar, A hybrid intrusion detection model using ega-pso and improved random forest method, Sens., 16(2022)5986, https://doi.org/10.3390/s22165986.

[7] M.A. Umer, K.N. Junejo, M.T. Jilani, A.P. Mathur, Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations, Int. J Crit Infrastruct Prot, 38(2022)100516, https://doi.org/10.48550/arXiv.2202.11917.

[8] Y.K. Saheed, A.I. Abiodun, S. Misra, M.K. Holone, R.A. Colomo-Palacios, A machine learning-based intrusion detection for detecting internet of things network attacks, Alexandria Eng. J., 12(2022)9395-940, https://doi.org/10.1016/j.aej.2022.02.063.

[9] M.A. Talukder, K.F. Hasan, M.M. Islam, M.A. Uddin, A. Akhter, M.A. Yousuf, F. Alharbi, M. A. Moni, A dependable hybrid machine learning model for network intrusion detection, J. Inf. Secur. Appl., 72(2023)103405, https://doi.org/10.48550/arXiv.2212.04546.

[10] H. Bangui, M. Ge and B. Buhnova, A hybrid machine learning model for intrusion detection in VANET, Computing, 3(2022)503-531, https://doi.org/10.1007/s00607-021-01001-0.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

[11] O.A. Alzubi, J.A. Alzubi, M. Alazab, A. Alrabea, A. Awajan, I. Qiqieh, Optimized machine learning-based intrusion detection system for fog and edge computing environment, J. Power Electron., 19(2022)3007, https://doi.org/10.3390/electronics11193007.

[12] A.S. Alqahtani, FSO-LSTM IDS: hybrid optimized and ensembled deep-learning network-based intrusion detection system for smart networks, J Supercomput, 7(2022)9438-9455, https://doi.org/10.1007/s11227-021-04285-3.

[13] N. Omer, A.H. Samak, A.I. Taloba, R.M. Abd El-Aziz, A novel optimized probabilistic neural network approach for intrusion detection and categorization, Alexandria Eng. J., 72(2023)351-361, https://doi.org/10.1016/j.aej.2023.03.093.

[14] A. El-Ghamry, A. Darwish, A.E. Hassanien, An optimized CNN-based intrusion detection system for reducing risks in smart farming, ACM Trans. Internet Things, *22*(2023)100709, https://doi.org/10.1016/J.IOT.2023.100709.

[15] P.R. Kanna, P. Santhi, Hybrid intrusion detection using MapReduce based black widow optimized convolutional long short-term memory neural networks, Expert Syst. Appl., 194(2023)116545, https://doi.org/10.1016/j.eswa.2022.116545.

[16] A.A. Abdulrahman, M.K Ibrahem, Intrusion Detection System Using Data Stream Classification, Iraqi J Sci, 1(2021)319-328, https://doi.org/10.24996/ijs.2021.62.1.30.

[17] W.F. Kamil, I.J. Mohammed, Adapted CNN-SMOTE-BGMM Deep Learning Framework for Network Intrusion Detection using Unbalanced Dataset, Iraqi J Sci, 9(2023)4846-4864, https://doi.org/10.24996/ijs.2023.64.9.43.

[18] M Tavallaee, E. Bagheri, W. Lu and A. A. Ghorbani, A detailed analysis of the KDD CUP 99 data set, In *2009 IEEE* symposium on computational intelligence for security and defence applications, 1(2019)1-6, IEEE, http://dx.doi.org/10.1109/CISDA.2009.5356528.

[19] P. Li, J. Zhang, and P. Krebs, "Prediction of flow based on a CNN-LSTM combined deep learning approach, WATER-SUI, 6(2022)993, https://doi.org/10.3390/w14060993.

[20] A. E. Abulwafa, A Survey of Deep Learning Algorithms and its Applications, Nile J Commun. Comput. Sci., 1(2022)28-49, https://doi.org/10.21608/njccs.2022.139054.1000.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

[21] R. Vankdothu, M.A. Hameed, H. Fatima, A Brain Tumor Identification and Classification Using Deep Learning Based on CNN-LSTM Method, Comput. Electr. Eng., 101(2022)107960, https://doi.org/10.1016/j.compeleceng.2022.107960.

[23] A. Heidari, N.J. Navimipour, M. Unal, A Secure Intrusion Detection Platform Using Blockchain and Radial Basis Function Neural Networks for Internet of Drones, IEEE Internet Things, 10(2023)99, https://doi.org/10.1109/IOT.2023.3237661.

[24] Y. Yin, J. Jang-Jaccard, W. Xu, A. Singh, J. Zhu, F. Sabrina, J. Kwak, IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset, Int. J. Big Data Intel., 1(2023)1-26, https://doi.org/10.1186/s40537-023-00694-8.

[25] M. A. Almaiah, O. Almomani, A. Alsaaidah, S. Al-Otaibi, N. Bani-Hani, K. A. Hwaitat, A. Al-Zahrani, A. Lutfi, A. Bani-Awad, T. H. Aldhyani, Performance investigation of principal component analysis for intrusion detection system using different support vector machine kernels, Electronics, 21(2022)3571, http://dx.doi.org/10.3390/electronics11213571.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

طريقة التعلم الآلي الخاضعة للإشراف للكشف عن الحالات الشاذة

أسماء سالم يحيى

قسم البرمجيات، كلية علوم الحاسوب والرياضيات، جامعة الموصل، الموصل، العراق

المستخلص

أصبح الأمن السيبراني أحد أهم المواضيع في الوقت الحاضر. تخضع غالبية أنشطة حياتنا اليومية باستخدام التطبيقات عبر الإنترنت. حيث من الممكن أن تتعرض هذه التطبيقات لكمية مختلفة من التهديدات والتي تتسبب في الوصول غير المصرح به إلى البيانات الشخصية. ولذلك، أصبحت حماية بيانات التطبيقات أمرًا بالغ الأهمية. يُستخدم التعلم الآلي الخاضع للإشراف على نطاق واسع في العديد من التطبيقات، مثل اكتشاف البريد العشوائي. يعتبر التعلم الآلي أداة قوية لأتمتة عمليات صنع القرار وإجراء التنبؤات بناءً على البيانات المسجلة. يتضمن هذا العمل تعلمًا آليًا خاضعًا للإشراف لتصنيف الحالات الشاذة في الشبكة باستخدام مجموعة بيانات NSL-KDD والتي تُستخدم لأغراض اختبار طرق كشف التسلل. لا تحتوي مجموعة البيانات هذه على عناصر متكررة في مجموعة التدريب الفرعية، مما يجعل العمل غير متحيز لأي من العناصر. الطرق المستخدمة في هذا البحث هي CNN، LSTM، الهجين CNN-LSTM، RBFN، MLP، وSVM. غالبًا ما يكون من الأمور الجيدة تجربة خوارزميات متعددة ومقارنة أدائها لاختيار الأفضل. تم تقييم نتائج النماذج المنفذة وقياسها باستخدام معدل اكتشاف المقاييس واستهلاك الوقت والدقة. أظهرت النتائج أن النموذج الهجين لـ CNN-LSTM تفوق في الأداء على طرق القياس من حيث معدل الكشف (99.61%)، والدقة (99.8%).

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).