JOURNAL OF NORTHEASTERN UNIVERSITY

Volume 25 Issue 04, 2022 ISSN: 1005-3026 https://dbdxxb.cn/ Original Research Paper

HYBRID ARTIFICIAL ECOSYSTEM-BASED OPTIMIZATION WITH LIGHT GRADIENT BOOSTING FOR INTRUSION DETECTION

Mohammed Eltahir Abdelhag[1], Saad Mamoun [2], Mohd Sarfaraz[1], Emad Addin Alsheikh[3]

[1]Department of Information Technology and Security, Jazan University, Jazan 45142, Saudi Arabia

[2]Imam Mohammad Ibn Saud Islamic University (IMSIU), College of Shari'a & Islamic Studies, AL-Ahsaa. Department Computer Science and Information. Saudi Arabia

[3] Department of Computer Science, Jazan University, Jazan 45142, Saudi Arabia

mohedtahir@gmail.com[1], smaahmed@imamu.edu.sa[2], msarfaraz@jazanu.edu.sa[1], ialsheikh@jazanu.edu.sa[3]

Abstract

The number of people and applications make internet usage dramatically increased by users over the past several years, resulting in more security problems. To provide a secure environment, businesses and institutions give more attention to providing more effective safeguards against modern attacks. Machine Learning (ML) algorithms show great positional to be used in Intrusion Detection (ID) systems which can monitor and distinguish whether a packet is a malicious or typical system behavior based on the data it contains. Therefore, this paper introduces an efficient model based on Artificial Ecosystem-based Optimization (AEO) and Light Gradient Boosting model (LGBM), named AEO-LGBM. The AEO is employed to select the most informative features from large datasets, while the LGBM model is used for classification. The AEO-LGBM is verified on two datasets for ID: KDD CUP99 and NSL-KDD and compared to Logistic Regression (LR), Multilayer Perceptron (MLP), and Support Vector Machine (SVM). The results validate the superiority of the introduced model over the examined techniques and recently reported models in the literature for ID.

Keywords: Feature selection, machine learning, Intrusion Detection, metaheuristic algorithms,

1. Introduction

With the increased use of internet services, cybersecurity issues have become one of the most serious challenges that have specific risks on individual and business operations [1]. A variety of security mechanisms such as firewalls, Intrusion Detection Prevention System (IDPS), encryption, and antivirus are used by organizations and enterprises to better deal with such cybersecurity attacks to their networks. [2, 3]. IDPSs prove themselves as a powerful technology for detecting and preventing several attack types. However, every packet passed on the network cannot be analysed in depth [4], so the desired protection level cannot be reached.

Intrusion Detection Systems (IDSs) can be grouped as signature-based or anomaly-based. Signature-based IDS compares the network data with the representation of different attacks in

Submitted: 21/10/2022 **Accepted**: 24/11/2022

the database to generate an alarm when a match is found [5]. The primary limitation of such IDS systems is that an attack remains undetected if not represented in the database. Anomaly-based IDS compares normal behaviors from the database to generate an alarm when detecting a deviation from the specified rules. The fundamental advantage of these IDS is that they can detect unknown attacks.

To achieve optimal network security requirements, researchers used Machine learning (ML) approaches to develop more intelligent ID systems that can look inside packet payloads and detect such attacks with high accuracy and a low false positive rate [6]. The use of ML techniques ginned special attention in ID in recent years because of their ability to classify hundreds of features into normal system behavior or attack attempts [7, 8]. The primary purpose of Feature Selection (FS) as a technique is to select an Optimal Feature Subset (OFS) in a given dataset, thus, optimizing the learning process by ML techniques.

In the present paper, AEO and LGBM are combined to develop a model named AEO-LGBM. The developed AEO- LGBM is based on AEO Meta-Heuristics (MH) optimization method and LGBM model. The AEO is used for searching OFS, while LGBM is for classification. This work aims to develop a novel AEO-LGBM model for ID; the AEO-based FS method is utilized to reduce the feature dimensionality of KDD-Cup 99. The LGBM uses NSL-KDD datasets and the reduced dataset for classification purposes. The efficiency of the AEO-LGBM is evaluated in terms of several evaluation measures and compared to LR, MLP, and SVM models. The results of the AEO- LGBM performed well when it is compared to the tested approaches and other existing models in the literature, making it more suitable and practical for ID

The rest of this paper is organized as follows: In section 2, some recent studies for ID using ML is provided. In section 3, a brief overview of AEO, LGBM, and the developed AEO-LGBM are given, in section 4, experimental results and discussion are presented. Finally, section 5 concludes this work.

2. Related works

In this section, we briefly discuss some related work that is relevant to ML methods used for ID. Many ML approaches have been applied in the application of ID. [9] proposes a model for ID using Deep Extreme Learning Machine (DELM). The authors validated the results of the DELM using Knowledge Discovery Dataset (KDD)-Cup 99, and the final results showed an accuracy of 94.60%. In another work [10], Sparse Auto-Encoder (SAE) is employed for FS, and LR is then utilized for classification. The authors used Security Laboratory Knowledge Discovery Dataset (NSL-KDD) datasets to test SAE-LR. Results showed that their proposed model achieves an accuracy of 87.2%.

In the work of [11], an efficient detection model is proposed using FS and Recurrent Neural Network (RNN) model. In the FS stage, they used Oppositional Crow Search Algorithm (OCSA), which integrates the Crow Search Algorithm (CSA) and Opposition Based Learning

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

(OBL) methods to select OFS from KDD-Cup 99. The reported accuracy of 94.12% is better than other conventional methods. In [12], a Decision Tree (DT) based ID using the NSL-KDD dataset is presented. The authors utilized correlation-based FS to reduce the dates' dimensionality and improve the ID's classification accuracy. The outcomes showed that the proposed model performed better than others, with an accuracy of 90.3%.

In [13] work, an optimized Genetic based Enhanced Grey Wolf Optimization (GBEGWO) is presented for ID. The GBEGWO is used to select OFS from KDD-Cup 99, and the results showed that the improved FS method achieved better accuracy of 98.62% and performed better than the existing work. In [14], the authors designed an ID based on Extreme Machine Learning {EML} model in mobile edge computing and fog computing environment. They used KDD-Cup 99 dataset and the results showed that an accuracy of 99.07% is achieved. In [15]. a Deep Learning (DL) model is presented for ID. They used stacked symmetric Deep Auto-Encoder (DAE) and SVM to utilize both models power and reducing computational cost. The authors used KDD-Cup 99 to validate the robustness of their proposed model. The results showed that an accuracy of 99.65% is obtained by the proposed DL method,

In [16], a Spider Monkey Optimization (SMO) and Deep Neural Network (DNN) are proposed for dimensionality reduction, the authors used SMO, and the reduced dataset is fed into the DNN model. They evaluated the proposed SMO-DNN on model KDD-Cup 99 and NSL-KDD datasets, and the final results showed an accuracy of 99.4% and 92% attained for both datasets, respectively. Due to the wide availability of hacking tools that do not require many skills to launch an attack by users, accurate models are still needed to better distinguish such attacks from normal system behaviors. Therefore, this paper introduces an ML model named AEO-LGBM, which could provide a practical and accurate solution for ID.

3. Proposed method

3.1.Feature selection using AEO

Artificial Ecosystem-based Optimization (AEO) is a new MH method that is motivated by the energy flow in the natural ecosystem introduced by [17]. AEO uses three leading operators to achieve optimal solutions and they include.:

1. Production

In this operator, the producer (worst individual) in the population is updated with respect to the best individual within the given search space boundaries. It guides others to search different regions. The operator replaces the previous individual with a new one generated between the best and randomly produced (x_{rand}) individuals. This operator can be given as:

$$\mathrm{x}_1(t+1) = (1-\alpha)x_n(t) + \alpha x_{rand}(t) \quad (1)$$

$$\alpha = (1 - t/\mathrm{T})r_1 \quad (2)$$

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

$$x_{rand} = r(Ub - Lb) + \mathrm{L}b \quad 3$$

where n denotes the size of a population, $\mathrm{x}_1(t)$ leads others to explore the search space broadly; in the following iterations, $\mathrm{x}_1(t+1)$ leads the others to intensively exploit in a region around x_n, α is a linear weight coefficient to drift the individual's position linearly from a random towards the best individual through the pre-define iterations T, r is a random vector having a range of [0, 1], , r_1 is generated randomly within [0, 1], and Ub , Lb are upper and lower boundaries of the search space.

2. Consumption

This operator starts after the production operator is completed. Each consumer may eat either a randomly chosen consumer with a low level of energy or a producer, or both to obtain food energy. A Levy flight-like random walk, called a Consumption Factor (CF), is employed to enhance exploration capability, and it is defined as follows:

$$\mathrm{CF} = \frac{1}{2}\frac{v_1}{|v_2|}, \quad v_1 \text{and } v_2 \in Norm\,(0,1) \quad 4$$

where, $Norm\,(0,1)$ presents a normal distribution with zero mean and unity standard deviation.

Different types of consumers adopt different consumption behaviors to update their positions and these strategies include:

- Herbivore behavior: A consumer would eat only the producer if it is chosen as an herbivore, and this behavior can be framed as:

$$\mathrm{x}_i(t+1) = \mathrm{x}_i(t) + \mathrm{CF}.\left(\mathrm{x}_i(t) - \mathrm{x}_1(t)\right), \quad i \in [2, \cdots n] \quad 5$$

- Carnivore behavior: A carnivore consumer would eat only higher energy level consumers, and it can be modeled as:

$$\mathrm{x}_i(t+1) = \mathrm{x}_i(t) + \mathrm{CF}.\left(\mathrm{x}_i(t) - \mathrm{x}_j(t)\right), \quad i \in [3, \cdots n] \quad 6$$

$$\mathrm{j} = \mathrm{ran}di([2i-1])$$

- Omnivore behavior: A consumer chosen as an omnivore can randomly eat a higher energy level consumer or a producer and this behavior can be formulated as:

$$\mathrm{x}_i(t+1) = \mathrm{x}_i(t) + \mathrm{CF}.(r_2\,.(\mathrm{x}_i(t) - \mathrm{x}_1(t))) + (1 - r_2)(\mathrm{x}_i(t) - \mathrm{x}_j(t)) \quad i\,[3, \cdots n] \quad 7$$

$$\mathrm{j} = \mathrm{ran}di([2i-1])$$

where r_2 is a rnndom number in the range of [0, 1]

3. De-composition

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

This is the final phase of the ecosystem as everyone in the agent dissolves, and the decomposer provides the necessary nutrients for the producer's growth by breaking down the remains of dead individuals in the population. The de-composition operator can be expressed as:

$$\mathrm{x}_i(t+1) = \mathrm{x}_n(t) + \mathrm{D}e.\,(e\,.\,\mathrm{x}_n(t) - \mathrm{h}\,.\,\mathrm{x}_j(t)), \qquad i = 1, \cdots n \tag{8}$$

$$De = 3u. \quad u \in \mathrm{N}(0,1)$$

$$e = r_3\ .\mathrm{rand}i([1\ 2]) - 1$$

$$h = 2\ .\ r_3 - 1$$

where e, h, and De, are weight coefficients designed to model de-composition behavior

3.2. Light Gradient Boosting model (LGBM)

LGBM ensembles the decisions from a set of weak learners to build a strong model [18]. LGBM conserves the high accuracy of Gradient Boosted Decision Trees (GBDT) while reducing computational time and memory consumption [19].

A histogram-based algorithm in LGBM boosts the model's capability to deal with high dimensional data and prevent model overfitting. Boosting technique transforms continuous eigenvalues into *l* integers to generate a *k*-width histogram with restricted depth. Furthermore, Local Voting Decision (LVD) selects top-*k* samples from the distributed initial samples of multiple trees. For the *k* iteration process, the top-2*k* attributes are computed by a global voting decision by collecting the most important *k* LVD attributes. LGBM selected optimized leaves using the Leaf-wise method. The objective function of LGBM is given by [19]:

$$\text{Objective Function (t)} = \mathrm{L(t)} + \Omega(\mathrm{t}) + \mathrm{c} \tag{9}$$

where, $\mathrm{L(t)}$ and $\Omega(\mathrm{t})$ represent loss function and regular, c extra parameter to prevent the model's overfitting, and t is the sampling time.

The model fitness represented by the loss function can be given as [20] :

$$\mathrm{L(t)} = \sum_{n=1}^{n} \{a_i(t) - p_i(t)\}^2 \tag{10}$$

where a_i is the actual values p_i predicted values for n samples

3.3.Proposed AEO –LGBM

LGBM is a robust classifier with high performance reported for several applications. Although it selects features based on their importance for improving classification performance at each boosting stage, the curse of dimensionality affects its performance. A large number of input-features limits the LGBM's ability to select salient features at each level and requires

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

more estimators, increasing model complexity for performance generalization. AEO method has been widely used for selecting salient features in several applications due to its strong exploration ability. This section gives details of the proposed AEO-LGBM that uses AEO for determining OFS, which is used to train LGBM.

Firstly, k-fold cross-validation divides the dataset into training and test subsets. The training data is given as input to AEO for determining OFS. The OFS of training data is input to LGBM for training an ID classification model. A feature reduction is applied to the testing data based on the OFS training data. The trained LGBM model classifies the OFS of testing data. The predicted class labels and original test labels are input to the evaluation module to calculate measures such as accuracy, precision, recall, F1-score, and receiver operating curve (ROC). The process is repeated until each k-fold validation set is used as testing data.

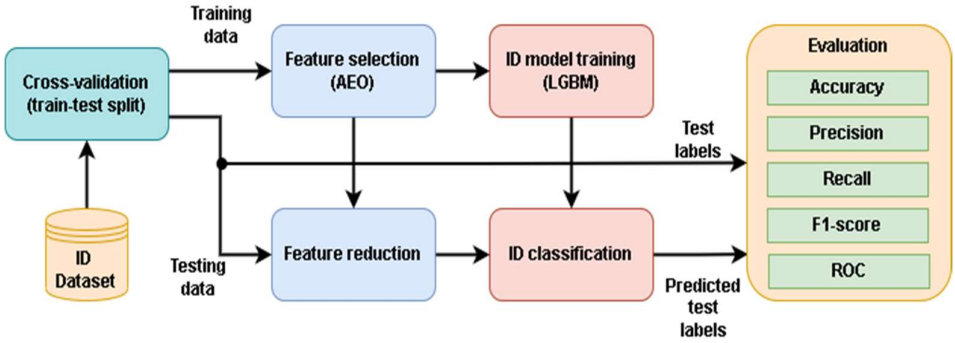


Figure 1. AEO-LGBM workflow

Table 1. Prameters of the proposed AEO-LGBM model

Algorithm	Parameters
AEO	r_1, r_2 and r_3 are random numbers in range [0, 1]
LGBM	Boosting type = traditional GBDT, number of estimators = 50, maximum depth = 6, learning rate = 0.1

4. Simulation results

4.1.Experimental setup

The AEO- LGBM is validated by conducting experiments on two benchmarked ID datasets: KDD- CUP99 and NSL-KDD. The KDD- CUP99 dataset includes DoS, Remote to Local (R2L), User to Root (U2R), and probing attack properties. It is one of the most widely used datasets for assessing ID models comprising about 5 million lines for seven weeks of network traffic [20]. The NSL-KDD is an upgraded version of KDD- CUP99, which contains a 42-

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

dimensional feature in each record. It avoids unnecessary and repetitive records from the KDD-CUP99 dataset with the same properties [21].

10 -fold- Cross Validation (CV) method is employed to avoid possible bias in selecting the training and testing datasets. All the experiments are implemented using Python and executed on a 3.13 GHz PC with 32 GB RAM and Win 11 operating system. The main characteristics of those datasets are presented in Table 2.

Table 2. KDD- CUP99 and NSL-KDD datasets characteristics

Dataset	Year	No. of features	No. of samples
KDD-CUP99	1998	43	494020
NSL-KDD	2009	43	125973

4.2. Evaluation measures

A set of evaluation measures can be used to validate the proposed AEO-LGBM model and its efficiency. In this work, accuracy, precision, recall, and F-measure and these measures are calculated as follows:

$$AC = \frac{TP + TN}{TP + TN + FN + FP} \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{F} - \text{measure} = \frac{2\,P\,R}{P + R} \quad (14)$$

where True Positive (TP) and True Negative (TN) denote the correctly detected churner or not cases. False Negative (FN) represents misclassified positive and False Positive (FP) is misclassified negative.

4.3.Experimental results and discussion

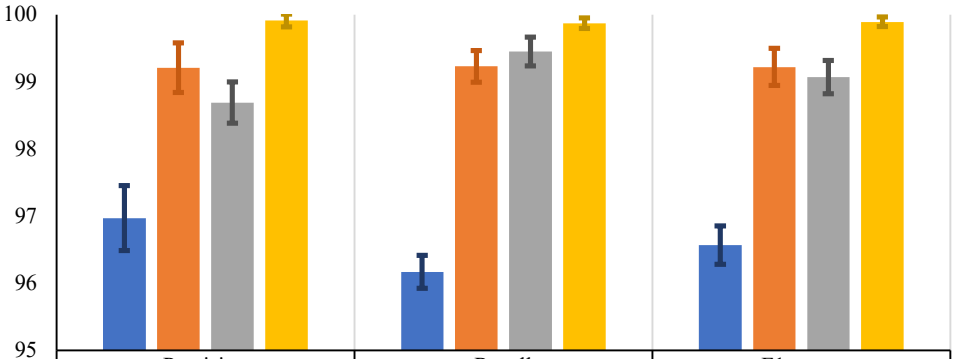
To examine the effectiveness of the AEO-LGBM model, KDD-CUP99 and NSL-KDD datasets are used. Table 3 gives the mean and Standard Deviation (SD) of the achieved accuracies by LR, MLP, SVM, and proposed AEO-LGBM models. It can be seen that the proposed AEO-LGBM gained higher mean accuracy than the other comparative models, indicating that FS using AEO has boosted the overall accuracy of the LGBM model by reducing the redundant features during the classification phase. The SD of the proposed AEO-LGBM model is the smallest among the three, demonstrating its stability more than the other three models.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

Table 3. Performance comparison of the accuracy of the AEO-LGBM, LR, MLP, and SVM models.

Model		KDD-CUP99	NSL-KDD
LR	Mean	96.83%	98.25%
	SD	0.6824	0.5263
MLP	Mean	99.23%	98.81%
	SD	0.4723	0.5623
SVM	Mean	98.89%	99.85%
	SD	0.4690	0.4642
Proposed	Mean	99.92%	99.88%
AEO-LGBM	SD	0.4376	0.3513

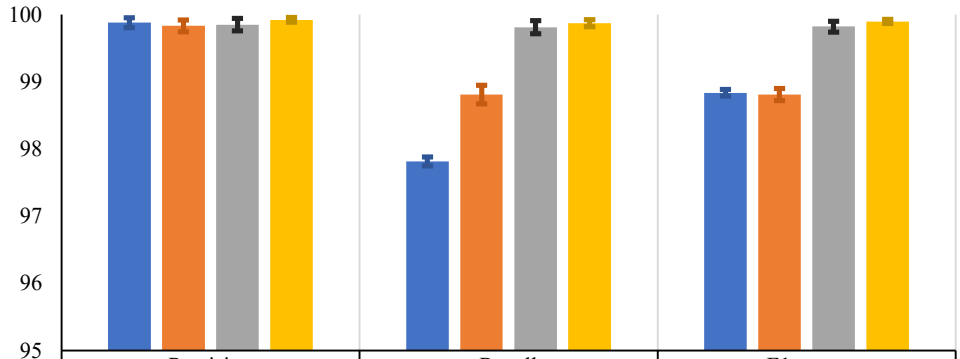
Figures 2 and 3 provide a comparative analysis of the proposed AEO-LGBM, LR, MLP, and SVM using both datasets' precision, recall, and F-measure evaluation measures. The evaluation measure type is plotted on the horizontal axis with values on the vertical axis, and different colors mark different IDs. Figure 2 shows that the proposed AEO-LGBM has ranked first in both datasets, while LR ranked last on the KDD-CUP99 dataset. Almost all the models attained similar precision values on the NSL-KDD dataset with a slight advantage over the proposed AEO-LGBM.



	Precision	Recall	F1-score
LR	96.97	96.17	96.57
MLP	99.21	99.23	99.22
SVM	98.69	99.45	99.07
AEO-LGBM	99.91	99.87	99.89

Figure 2. Quantitative comparison of the proposed AEO-LGBM, LR, MLP, and SVM models using KDD-CUP99 dataset

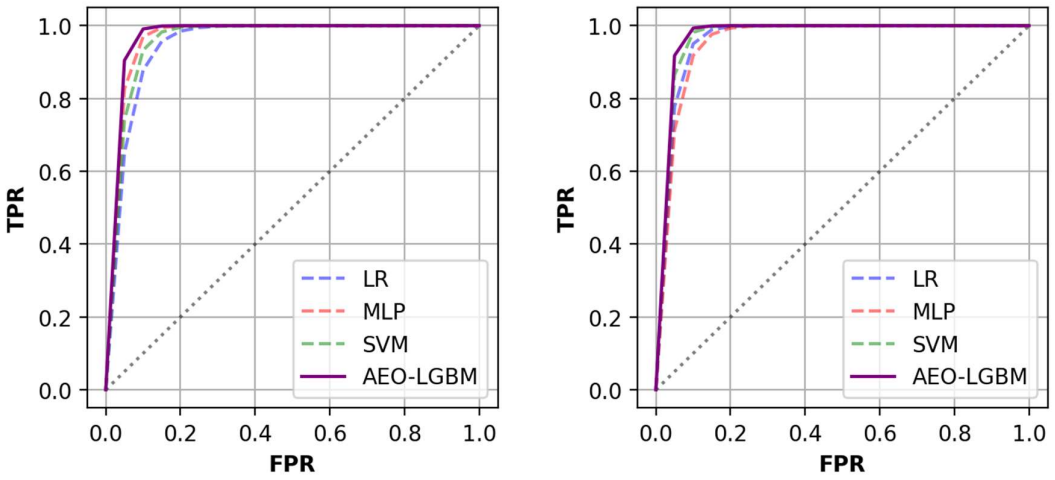
Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/



	Precision	Recall	F1-score
LR	99.88	97.81	98.83
MLP	99.83	98.81	98.81
SVM	99.85	99.81	99.82
AEO-LGBM	99.92	99.87	99.90

Figure 3. Quantitative comparison of the proposed AEO-LGBM, LR, MLP, and SVM models using the NSL-KDD dataset

The Receiver Operating Characteristic (ROC) curve represents the variation of True-Positive Rates (TPR) and False-Positive Rates (FPR) for different possible thresholds. A well-trained model should have the least FPR while the highest TPR; hence the curve must be biased toward the top left corner. The results of the ROC curves are shown in Figure 4 for both datasets.



(a) (b)

Figure 4. ROC curve for different models using a) KDD-CUP99 and b) NSL-KDD datasets

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

From Figure 4, the proposed AEO-LGBM curve dominates the other classifiers at all thresholds for both KDD-CUP99 and NSL-KDD datasets, indicating that the proposed AEO-LGBM offers the best result.

4.4.*Comparison* with other reported models

Recently, several works have been introduced for ID. Table 4 depicts the accuracy of the achievements of the existing models. The studies in Table 4 used KDD CUP99 and NSL-KDD datasets to validate their proposed model's efficiency. As per the results in table 4, the proposed AEO-LGBM model provides higher accuracy in both KDD CUP99 and NSL-KDD datasets than the existing models.

Table 4. Accuracy comparison between the earlier reported models and the proposed AEO-LGBM.for ID

Model	KDD CUP99	NSL-KDD
DELM [9]	94.60%	-
SAE-LR [10]	-	87.20%
OCSA-RNN [11]	94.12%	-
DT [12]	-	90.30%
GBEGWO [13]	98.62%	-
EML [14]	99.07 %	-
DAE-SVM [15]	99.65%	-
SMO-DNN [16]	99.40%	92.00%
Proposed AEO-LGBM	99.92%	99.88%

5. Conclusion and future work

Advent in intelligent devices has enabled the internet as integral to every aspect of our daily life. This brings new forms of attacks which should be identified to contain their risks. In this regard, ID systems have been developed to monitor and identify such types of attacks in network traffic. In this work, the AEO-LGBM model is developed and presented for ID. The AEO-LGBM uses ERO for FS and LGBM as a learning model. Two datasets are used to test the efficacy of the AEO-LGBM model, including KDD CUP99 and NSL-KDD. Results show that the introduced AEO-LGBM gained better results than LR, MLP and SVM models in several evaluation measures.Moreover, the AEO-LGBM produced superior performance compared to other reported approaches for ID in the literature. In the future, the AEO-LGBM model will be used in applications such as signal processing, website phishing attacks, and malware detection. Another possible lane is to work on MH methods for FS in the application of ID due to the great potential of these methods in other domains.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

References

[1] Choo, K. K. R., Gai, K., Chiaraviglio, L., & Yang, Q. (2021). A multidisciplinary approach to Internet of Things (IoT) cybersecurity and risk management. Computers & Security, 102, 102136.

[2] Jaw, E., & Wang, X. (2021). Feature selection and ensemble-based intrusion detection system: an efficient and comprehensive approach. Symmetry, 13(10), 1764.

[3] Krishnaveni, S., Sivamohan, S., Sridhar, S. S., & Prabakaran, S. (2021). Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing. Cluster Computing, 24(3), 1761-1779.

[4] Balasaraswathi, V. R., Mary Shamala, L., Hamid, Y., Pachhaiammal Alias Priya, M., Shobana, M., & Sugumaran, M. (2022). An Efficient Feature Selection for Intrusion Detection System Using B-HKNN and C2 Search Based Learning Model. Neural Processing Letters, 1-25.

[5] Ford, V., & Siraj, A. (2014, October). Applications of machine learning in cyber security. In Proceedings of the 27th international conference on computer applications in industry and engineering (Vol. 118). Kota Kinabalu, Malaysia: IEEE Xplore.

[6] Gupta, A., Gupta, R., & Kukreja, G. (2021). Cyber security using machine learning: techniques and business applications. In Applications of Artificial Intelligence in Business, Education and Healthcare (pp. 385-406). Springer, Cham.

[7] Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, 32(1), e4150.

[8] Kaushik, D., Garg, M., Gupta, A., & Pramanik, S. (2022). Application of machine learning and deep learning in cybersecurity: An innovative approach. In An Interdisciplinary Approach to Modern Network Security (pp. 89-109). CRC Press.

[9] Khan, M. A., Abbas, S., Rehman, A., Saeed, Y., Zeb, A., Uddin, M. I., ... & Ali, A. (2020). A machine learning approach for blockchain-based smart home networks security. IEEE Network, 35(3), 223-229.

[10] Gurung, S., Ghose, M. K., & Subedi, A. (2019). Deep learning approach on network intrusion detection system using NSL-KDD dataset. International Journal of Computer Network and Information Security, 11(3), 8-14.

[11] SaiSindhuTheja, R., & Shyam, G. K. (2021). An efficient metaheuristic algorithm based feature selection and recurrent neural network for DoS attack detection in cloud computing environment. Applied Soft Computing, 100, 106997.

[12] Ingre, B., Yadav, A., & Soni, A. K. (2017, March). Decision tree based intrusion detection system for NSL-KDD dataset. In International conference on information and communication technology for intelligent systems (pp. 207-218). Springer, Cham.

[13] Yerriswamy, T., & Murtugudde, G. (2021). An efficient algorithm for anomaly intrusion detection in a network. Global Transitions Proceedings, 2(2), 255-260.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/

[14] An, X., Zhou, X., Lü, X., Lin, F., & Yang, L. (2018). Sample selected extreme learning machine based intrusion detection in fog computing and MEC. Wireless Communications and Mobile Computing, 2018.

[15] Imran, M., Haider, N., Shoaib, M., & Razzak, I. (2022). An intelligent and efficient network intrusion detection system using deep learning. Computers and Electrical Engineering, 99, 107764.

[16] Khare, N., Devan, P., Chowdhary, C. L., Bhattacharya, S., Singh, G., Singh, S., & Yoon, B. (2020). Smo-dnn: Spider monkey optimization and deep neural network hybrid classifier model for intrusion detection. Electronics, 9(4), 692.

[17] Zhao, W., Wang, L., & Zhang, Z. (2020). Artificial ecosystem-based optimization: a novel nature-inspired meta-heuristic algorithm. Neural Computing and Applications, 32(13), 9383-9425.

[18] Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., & Rehman, M. U. (2019). A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting. Ieee Access, 7, 28309-28318.

[19] Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X., & Zeng, W. (2019). Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. Agricultural water management, 225, 105758.

[20] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009, July). A detailed analysis of the KDD CUP 99 data set. In 2009 IEEE symposium on computational intelligence for security and defense applications (pp. 1-6). Ieee.

[21] Sapre, S., Ahmadi, P., & Islam, K. (2019). A robust comparison of the KDDCup99 and NSL-KDD IoT network intrusion detection datasets through various machine learning algorithms. arXiv preprint arXiv:1912.13204.

Copyright © 2022. Journal of Northeastern University. Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at https://dbdxxb.cn/