Received: August 11, 2022. Revised: September 18, 2022. 110

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Generative Adversarial Network and Bayesian Optimization in Multi-class Support Vector Machine for Intrusion Detection System

Amit Kumar Pandey[1]* Prashant Singh[2] Dhyanendra Jain[3]
Anupam Kumar Sharma[4] Ashu Jain[5] Anjani Gupta[6]

[1]*Department of Computer Science and Engineering with Data Science, ABES Engineering College, Ghaziabad, India*
[2]*Department of Computer Science & Engineering, Sunder Deep Engineering College, Ghaziabad, India*
[3]*Department of Computer Science & Engineering – Artificial Intelligence & Machine Learning, ABES Engineering College, Ghaziabad, India*
[4]*School of Computing Science & Engineering, Galgotia University, Greater Noida, India*
[5]*Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management, New Delhi, India*
[6]*Department of Computer Science & Engineering, IMS Engineering College, Ghaziabad, India*
* Corresponding author's Email: amitpandey33@gmail.com

Abstract: Network Intrusion detection performances are highly affected by imbalance data problems due to the presence of less number of attack information in the dataset. Deep learning models are applied in existing methods to improve the efficiency that has limitations of overfitting problems. The Generative Adversarial Network (GAN) – Bayesian optimization Multi-class Support Vector Machine (BMSVM) is proposed to overcome imbalance and overfitting problems in intrusion detection systems. The Min-Max Normalization method is applied to normalize the input data to reduce the differences in features. GAN model is applied to generate minority class to balance the data instances to train the model. The proposed GAN-BMSVM model is compared with the classical sampling method, optimization, and classifier in the intrusion detection model in terms of Accuracy, Detection Rate (DR), and False Alarm Rate (FAR). The classical sampling methods are Near-miss, SMOTE and Autoencoder; traditional classifiers are KNN, RF, SVM, DNN and LSTM, and classical optimizations are PSO and WOA. The existing researches such as HCRNN, HLD, DONN, FL-NIDS and CNN-LSTM are used to evaluate the efficiency of GAN-BMSVM model. The GAN-BMSVM model has achieved 99.58% and 85.38 % accuracy for NSL-KDD and UNSW-NB15 dataset respectively, which is higher than the existing CNN-LSTM model.

Keywords: Bayesian optimization, Generative adversarial network, Imbalance, Multi-class support vector machine, Network intrusion detection.

1. Introduction

Network issues are growing concern with security challenges and operating system domain in the network. The security efforts are having a similar shift in experiencing it and the local centralized approaches are evolved with distributed network approaches. This has made an effort to cope with the interconnected platforms from the heterogeneous networks to obtain the solution [1-4]. Intrusion Detection provides security as it is evolving with the network environments. The attack scenarios are analyzed and the formal description finds the events which are needs to be monitored. The formal way of automatically determining the intrusion is by collecting the data for support intrusion analysis which instruct the components to look after the events involved in run-time attack detection. The IoT is prone to distinct security issues due to internet infrastructure as it was taken place during the

exchange of information in the heterogeneous networks [5-8]. Therefore, an intelligent management system is needed for attack detection due to malicious intrusions. The attacks that threaten sensitive data through the internet are successfully detected. The IoT devices face issues in providing security to the sensitive information towards the end devices which are not used to support the security mechanism to target the malicious attacks distinctly [9, 10].

The present research work uses an improved model for the classification of attacks to solves the imbalance data problem and overfitting problem in intrusion detection. The contributions of the proposed research are shown as follows:

1. The existing methods in intrusion detection systems have the limitation of imbalance data problem due to less number of data instances in attack class. GAN is applied in this proposed model to generate more minority class data instances to balance the dataset.
2. BMSVM is applied to perform hyper-parameter optimization in the MSVM model to effectively handle the input data and improve the performance efficiency. The balanced dataset is applied in the BMSVM model to effectively improve the attack classification.
3. GAN model provides higher performance in handling the imbalance dataset compared to existing SMOTE and near miss sampling methods. BMSVM model shows higher performance than existing classifiers.

The organization of the paper is given as follows: Section 2 describes various optimization models developed for Intrusion detection in the network. Section 3 describes the proposed method involved in IDS with the help of a flow chart. Section 4 shows the results and discussions for the proposed method and Section 5 describes the conclusion for the present research work and the importance of future work.

2. Literature review

Khan [11] developed a Hybrid Convolution Recurrent Neural Network-Based Network Intrusion Detection System. (HCRNN-ID). The developed model used the CSE-CIC IDS2018 dataset and the HCNRNN-ID system reduced the complexity in the system computationally using a deep learning model for providing effective security against malicious attacks. The limitation of the developed model was that the model tested only on single ID datasets which could be evaluated for other datasets for improving the results.

Ramaiah [12] used an optimized deep neural network for intrusion detection. A correlation tool was applied in the developed model and neural based attack of predominant independent variables which is detected by random forest in KDDCUP-99 dataset. However, the intrusion detection framework has required more memory and complex neural activities.

Mulyanto [13] applied focal loss on cost-sensitive neural networks for the intrusion detection system. The developed model used Bot-IoT, UNSW-NB15, and NSL-KDD datasets for the evaluation of results. The algorithm FLNIDS was used to overcome the imbalanced data problem. The detection accuracy of the focal loss model in imbalanced dataset was improved in the developed model. Nevertheless, the DNN model was caused because the focal loss was heavily dependent on the datasets.

Hsu [14] developed a Robust Network Intrusion Detection Scheme Using Long-Short Term Memory (LSTM) for the network attacks classification. The NSL-KDD dataset was applied for the evaluation of results LSTM technique was increased significantly while applying CNN. The difficulty level attack was more complicated because of the ML scheme which cannot provide optimum performance.

Su [15] applied the Attention method in the Bi-directional LSTM model to perform forward and backward LSTM to extract features for the intrusion detection system. However, the BAT-MC algorithm in the developed model failed to utilize the tools effectively resulting in poorer results. Gao [16] utilized an Adaptive Ensemble Model for Intrusion Detection. The developed model used NSL-KDD for the evaluation of the results and the method used an ensemble learning algorithm to detect the capacity of such high-level malicious attacks. However, the decision tree was not good as that of DNN. However, the developed model required a small number of types of attacks such as U2R, separate optimization methods for improving the detection capability of such high-level threat attacks.

3. Method

Input data is normalized using Min-Max Normalization method to reduce the differences in the input data.

GAN model is applied in normalized data to generate more data instances related to minority class to balance the data. MSVM with Bayesian optimization is applied for the intrusion detection system to improve its performance. The flow of the GAN-BMSVM model in intrusion detection is shown in Fig. 1.

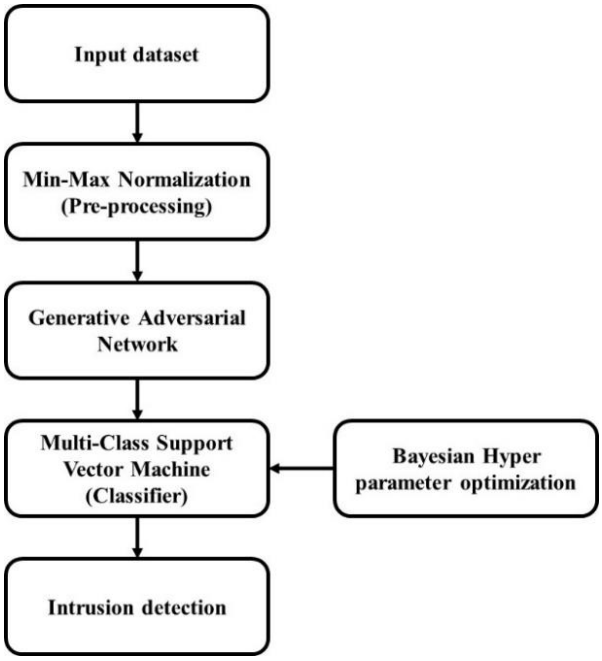

Figure. 1 The flow of GAN-BMSVM model in intrusion detection system

3.1 Min-max normalization

Min max normalization method uses the minimum and maximum value in the given features to measure the standard value in the given features, which are shown in Eqs. (1) and (2)

$$x_{std} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

$$x_{scaled} = x_{std} \times (max - min) + min \quad (2)$$

Where feature range is denoted as min, and max. The x_{scaled} is applied as X in the GAN model.

3.2 Generative adversarial network

The GAN model is applied to generate more data instances related to minority classes. Consider training dataset $X \subseteq R^{M \times T}$ in each stream of measurements, M and streams T and the testing dataset is $X^{test} \subseteq R^{N \times T}$ in each stream with N measurements and T streams. The GAN model is applied to generate more data similar to the minority class instances.

Window size of s_w sliding window is applied to effectively learn from X. Multivariate sub-sequences $X = \{x_i, i = 1,2, \dots, m\} \subseteq \mathcal{R}^{s_w \times T}$ is derived from multivariate time series using a step size s_s, the number of sub-sequences is denoted as $m = \frac{M - s_w}{s_s}$. Random space of multivariant sub-sequences set is denoted as $Z = \{z_i, i = 1,2, \dots, m\}$. GAN model is applied with Z and X to train discriminator and generator using minimax game of two-player, as shown in Eq. (3).

$$\min_G \max_D V(D,G) = \varepsilon_{x \sim p_{data}(X)}[logD(x)] + \varepsilon_{z \sim p_z(Z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (3)$$

Discriminator (D) and generator (G) of GAN are applied with the LSTM-RNN model. The model has been trained with particular iterations and applied the model to generate the data instances related to minority classes. Discrimination and Reconstruction Anomaly Score (DR-Score) is combined with test data to generate more data related to minority classes.

Multivariate sub-sequences $X^{tes} = x_j^{tes}, \quad j = 1,2, \dots, n$ with a sliding window for detection in testing dataset $x^{test} \subseteq \mathcal{R}^{N \times T}$, where $n = \frac{N - s_w}{s_s}$. The DR-Score (DRS) of testing dataset labels of each sub-sequences are given as follows in Eq. (4).

$$A_t^{tes} = \begin{cases} 1, & if H(DRS_t, 1) > \tau \\ 0, & else \end{cases} \quad (4)$$

Testing dataset with label vector is denoted as $A_t^{tes} \subseteq \mathcal{R}^{N \times 1}$ and minority class of non-zero value is measured using cross entropy error $H(.,..)$ in minority class score and this is higher than predefined value τ. The GAN structure is shown in Fig. 2.

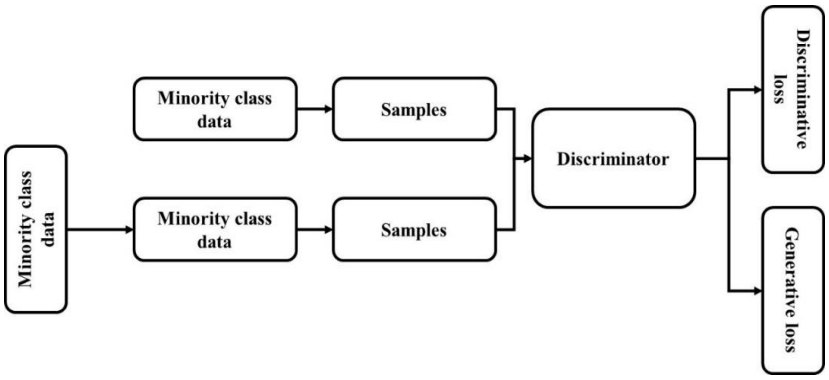

Figure. 2 Structure of GAN

3.3 Bayesian - multi-class support vector machine

Simplex coding in dimensional space is used to solve the problem of classification boundary. Simplex space is dimensional space and original mapping to simple space is optimized based classification error minimization. Simplex space decision boundary and object distance is measured to calculate classification error. The simplex coding guarantees the class predictor for each data instances and SVM solution has no ambiguity area.

The labels are denoted as $y_i \in \{1, \dots, K\}$ denotes samples $i \in \{1, \dots, n\}$ class, and p features with input data is $x_i \in R^p$. The translation vector is denoted as $t \in R^{K-1}$ in bias term and weight matrix is denoted as $W \in R^{p\times(K-1)}$. Linear function is applied for dimensional space $z_i' = x_i'W + t'$ of $(K-1)$ with sample i. SVM original space of kernel changes requires pre-processing on kernel matrix. A positive definite nucleus is denoted as $k: R^p \times R^p \to R^+$ that satisfies Mercer's theorem and reproducing core of Hilbert space which is denoted as H_k. The k action of definition map is $\psi(x) = k(x, .)$, and $k(x_i, x_j) =< \psi(x_i), \psi(x_j) >_{H_k}$. The ψ is defined as $n \times l$ matrix with row $\psi(x_i)$ and kernel matrix K is defined as $n \times n$ matrix with $k(x_i, x_i)$. Simplex space is mapped in Eq. (5). Number of iteration, population size, α, and cost value C were parameters used for optimization.

$$Z = \psi W + 1t' \tag{5}$$

Sample error i is measured using distance of each classification boundary. The sample i to class k and j distance is measured in Eq. (6).

$$q_i^{k_j} = (x_i'W + t')(g_k - g_j) \tag{6}$$

Huber hinge loss is given in Eq. (7).

$$h(q) = \begin{cases} 1 - q - \frac{k+1}{2} & if\ q \le -k \\ \frac{1}{2(k+1)}(1-q)^2 & if\ q \in (-k, 1] \\ 0 & if\ q > 1 \end{cases} \tag{7}$$

The total error is measured for each sample summed by l_p norm, as given in Eq. (8).

$$l_p = \left(\sum_{j=1, j\neq y_i}^{K} h^p\left(q_i^{y,j}\right)\right)^{\frac{1}{p}} \tag{8}$$

Optional sample weights are denoted as $\omega_i = \frac{n}{n_k K}, i \in G_k$ related to different group sizes or apply extra weight value to errors classification. The sample set is denoted as $G_k = \{i: y_i = k\}$ that belongs to each class k and the number of G_k samples are n_k. The total loss function of MSVM is denoted in Eq. (9).

$$L_{MSVM}(W, t) = \frac{1}{n}\sum_{k=1}^{K}\sum_{i\in G}\omega_i\left(\sum_{j\neq k} h^p\left(q_i^{kj}\right)\right)^{\frac{1}{p}} + \lambda trW'W \tag{9}$$

Where the regularization term is denoted as λ and penalty term is denoted as $\lambda trWW'$ to avoid overfitting. Penalty term effect is similar to Ridge Regression, applies l_2 norm row vector in W near to zero. The penalty term is $\lambda WW'$ for $K = 2$ with loss function in Eq. (10) which is improved by two-class SVM with a Huber hinge loss basis.

The simplex space is mapped with optimal $z_m = x_m'W + t'$ for unknown sample x_m to predict the class label of x_m, as in Eq. (10).

$$\hat{y}_m = argmin\left||z_m' - g_k'|\right|^2, for\ k = 1, \dots, K \tag{10}$$

Once the classification is performed, the performance is evaluated and compared with existing methods.

4. Simulation setup

The GAN-MSVM implementation details in intrusion detection were given in this section.

Dataset: The NSL-KDD [18], KDDCUP99 [19], UNSW-NB15 [20], and CICIDS2017 [21] datasets were used to evaluate GAN-BMSVM model in intrusion detection. The NSL-KDD is an improved dataset version of KDDCUP99 dataset with reduced redundancy and test data is not consist of duplicate records. KDDCUP99 has 38 numeric features and three categorical features in the dataset.

Metrics: Accuracy, Detection Rate (DR), and False Alarm Rate (FAR) were measured from the performance of the GAN-BMSVM model. The formula for Accuracy, DR, FAR are given in Eqs. (11) to (13).

$$Accuracy(\%) = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{11}$$

$$Detection\ Rate(\%) = \frac{TP}{TP+FN} \times 100 \tag{12}$$

$$FAR(\%) = \frac{FP}{FP+TP} \times 100 \tag{13}$$

Parameter settings: GAN has 4 hidden layer, 0.01 learning rate, and Adam optimizer is used.

System Requirement: Intel i9 processor, 128 GB RAM, 22 GB graphics, and windows 10 OS system were used to implement the GAN-BMSVM model.

5. Result

The Bayesian optimization with SVM model is proposed with GAN model to solve imbalance data problem in intrusion detection. Quantitative and comparative analysis of GAN-BMSVM model is given in this section.

The GAN-BMSVM model validation loss value for various epochs was evaluated and shown in Fig. 3. The GAN-BMSVM model has lower validation loss in the 8th epochs and loss value is increased in 9th epochs due to overfitting.

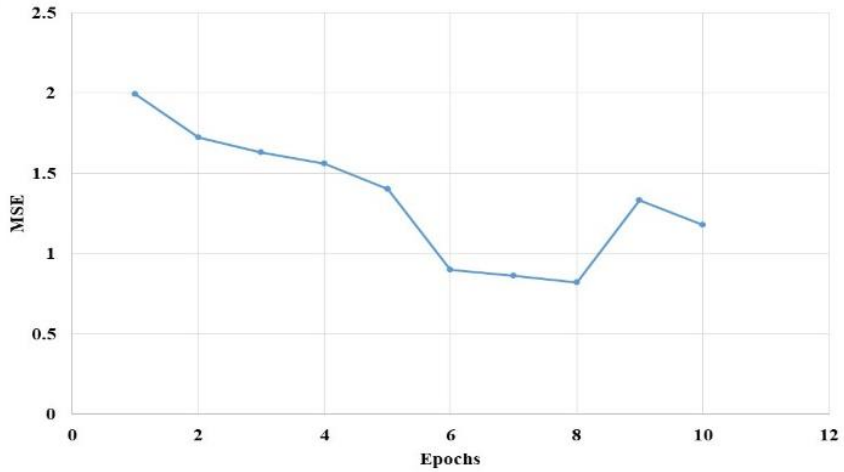

Figure. 3 MSE lose value analysis on various epochs

Table 1. Quantitative analysis of GAN-BMSVM

Methods	Accuracy (%)	DR (%)	FAR (%)
SVM	64.2	66.5	15.1
GAN-SVM	94.2	95.1	5.3
GAN-BMSVM	98.5	98.3	1.92

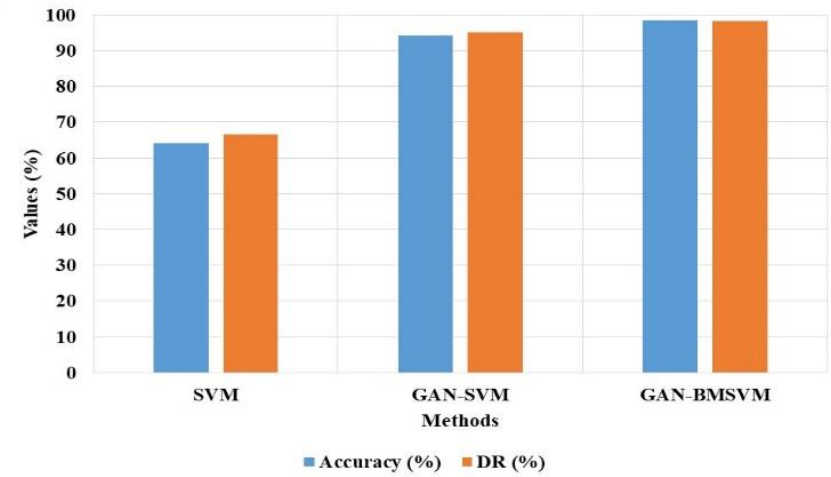

Figure. 4 Quantitative analysis on intrusion detection

Table 2. Sampling method comparison

Methods	Accuracy (%)	DR (%)	FAR (%)
Nearmiss	86.2	85.1	4.7
SMOTE	88.3	87.4	3.8
Autoencoder	93.4	94.2	2.4
GAN-BMSVM	98.5	98.3	1.92

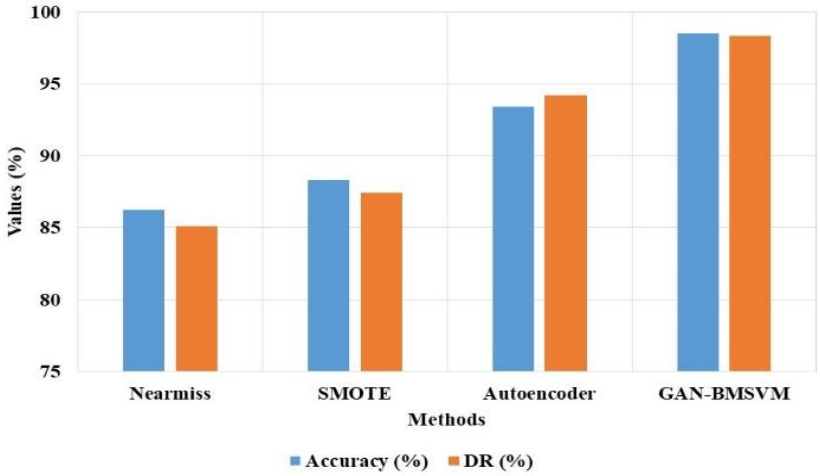

Figure. 5 Sampling method comparison for GAN-BMSVM

The quantitative analysis of GAN-BMSVM model in intrusion detection is shown in Table 1 and Fig. 4. SVM model has imbalance data problem and GAN-SVM training process is affected by overfitting. GAN model overcomes the limitation of imbalance data problem by generating the minority classes in the dataset. Bayesian optimization applied in SVM helps to select the optimal parameters to protect the model from overfitting problems in classification. The accuracy of GAN-BMSVM model is 98.5 %, GAN-SVM model is 94.2 % and SVM model has 64.2 % accuracy.

Sampling method is commonly applied to overcome the imbalance data problem in the intrusion detection system. A commonly applied method is SMOTE and Nearmiss is under-sampling method to reduce the minority class. GAN model is applied in the proposed model to generate the minority class and effectively train the SVM model in classification. Fig. 5 and Table 2 give the GAN-BMSVM model comparison with sampling methods of SMOTE, and Nearmiss method in an intrusion detection system. GAN-BMSVM model has higher performance due to GAN model generating the minority class instance to balance the data, Bayesian optimization is applied to overcome overfitting in SVM and MSVM model effectively handle high-dimensional data. SVM-Nearmiss method has lower efficiency due to less number data is available for SVM in intrusion detection. The GAN-BMSVM model has 98.5 % accuracy, SMOTE has 88.3 % accuracy, and nearmiss has 86.2 % accuracy.

GAN model is applied to balance the data and tested with various classifier models to analyze the efficiency, as in Fig. 6 and Table 3. GAN-BMSVM model has higher performance due to Bayesian optimizer providing optimal parameter to overcome overfitting problem and MSVM model effectively handle high dimensional data. Long Short Term Memory (LSTM) model has a second higher performance due to its efficiency in remembering historic data for classification and this model also have a limitation of vanishing gradient problem. Deep Neural Network (DNN), Random Forest (RF), and K-Nearest Neighbour (KNN) have lower efficiency in handling high dimensional data. KNN model has outlier sensitivity and DNN has an overfitting problem in classification. GAN-BMSVM model has 98.5 % accuracy, LSTM has 91.2 %, DNN has 88.3 %, SVM has 64.2 %, and KNN has 75.3 % accuracy in intrusion detection.

Various optimization methods such as Particle Swarm Optimization (PSO), and Whale Optimization Algorithm (WOA) are compared with Bayesian optimization in SVM, as shown in Fig. 7 and Table 4. PSO method is easily trap into local optima and WOA method has lower convergence in parameter

Table 3. Classifier comparison in intrusion detection

Methods	Accuracy (%)	DR (%)	FAR (%)
KNN	75.3	76.5	9.3
RF	78.1	75.3	8.7
SVM	64.2	66.5	15.1
DNN	88.3	89.4	4.3
LSTM	91.2	92.4	3.4
GAN-BMSVM	98.5	98.3	1.92

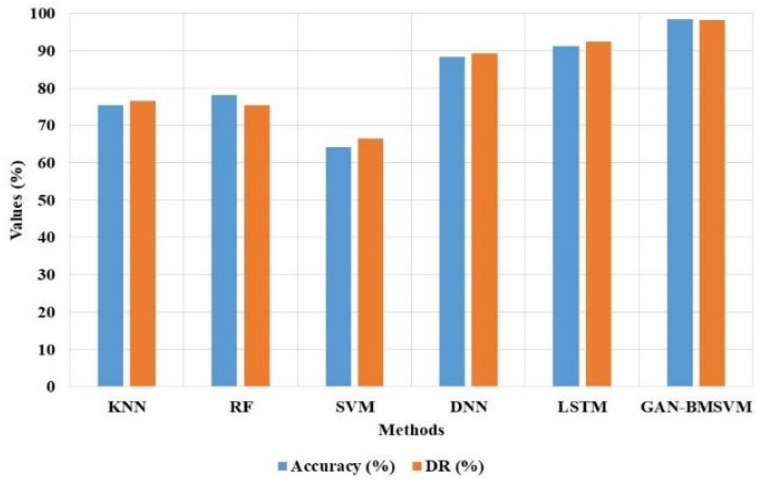

Figure. 6 Classifier comparison in intrusion detection

Table 4. Hyper parameter optimization comparison

Methods	Accuracy (%)	DR (%)	FAR (%)
SVM	64.2	66.5	15.1
SVM-PSO	82.4	83.5	12.1
SVM-WOA	85.3	86.1	11.7
GAN-BMSVM	98.5	98.3	1.92

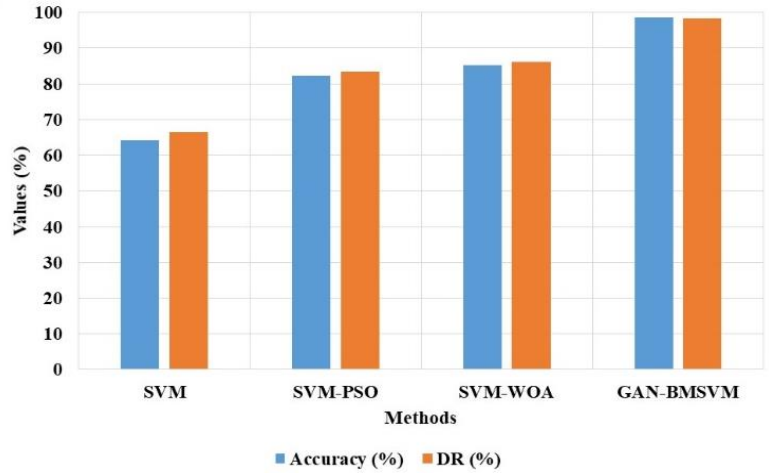

Figure. 7 Hyper parameter optimization comparison

selection. Penalty term in Bayesian optimization selects the unique features to avoid overfitting in the parameter selection and improves the classification efficiency. GAN-BMSVM model has 98.5 % accuracy, SVM-WOA has 85.3 %, SVM-PSO has 82.4 %, and SVM has 64.2 % accuracy.

5.1 Comparative analysis

Recent methods such as HCRNN [11], FL-NIDS [14] in intrusion detection system were compared with GAN-BMSVM model on CSE-CIC-IDS2018 and NSL-KDD dataset are shown in Table 5 and Table 6.

From the Table 5 and 6, it clearly shows that proposed GAN-BMSVM achieved better results in terms of accuracy, F1-Score and Recall. The proposed GAN-BMSVM achieved the maximum accuracy of 99.20% and 99.58% on CSE-CIC-IDS2018 and NSL-KDD datasets respectively. Table 7 shows the comparative analysis of accuracy on NSL-KDD and UNSW-NB15 datasets.

Table 5. Comparative analysis on CSE-CIC-IDS2018 dataset

Methods	Accuracy (%)	F1-Score (%)	Recall (%)
HCRNN [11]	97.75	97.60	97.12
GAN-BMSVM	99.20	99.17	98.76

Table 6. Comparative analysis on NSL-KDD dataset

Methods	Accuracy (%)	F1-Score (%)	Recall (%)
FL-NIDS [14]	76.00	51.96	52.41
GAN-BMSVM	99.58	99.29	98.87

Table 7. Comparative analysis of accuracy

Dataset	Methods	Accuracy (%)
NSL-KDD	CNN-LSTM [16]	99.47
	GAN-BMSVM	99.58
UNSW-NB15	CNN-LSTM [16]	73.00
	GAN-BMSVM	85.38

From the Table 7, it clearly displays that the GAN-BMSVM model has higher performance than existing methods in intrusion detection due to its capacity to handle imbalance data problems. Existing CNN-LSTM [16] model have limitations of overfitting problems due to the deep learning method generating more data for feature learning. The existing methods in the intrusion detection model were affected by imbalance data problem in classification. The GAN model generates the data instances of minority class to balance the data and the BMSVM model selects optimal parameters to overcome the overfitting problem. In NSL-KDD dataset, the proposed GAN-BMSVM achieved 99.58% which is better than existing CNN-LSTM which attained 99.47%. While, the GAN-BMSVM model has 85.38% accuracy in the UNSW-NB15 dataset and the CNN-LSTM model has 73% accuracy.

6. Conclusion

The GAN-BMSVM model is proposed to overcome the imbalance data problem and optimal parameter selection in the MSVM model for an effective intrusion detection system. The GAN-BMSVM model has three advantages: GAN model generates more minority data to balance the dataset, MSVM model effectively handles high dimensional features, and Bayesian optimization selects optimal parameter settings for MSVM in intrusion detection. Due to the above-mentioned reasons, GAN-BMSVM attaining better results. The proposed GAN-BMSVM model is evaluated using UNSW-NB15 and NSL-KDD datasets. From the result analysis, it clearly shows that proposed GAN-BMSVM model achieves high accuracy of 99.58% and 85.38% on NSL-KDD and UNSW-NB15 datasets respectively. Deep learning-based models such as CNN, and LSTM have limitations of overfitting and vanishing gradient problems. MSVM model has the limitation of imbalance data problem that is overcome using GAN model. KNN model is sensitive to outliers and the DNN model has an overfitting problem in classification. Sampling methods such as SMOTE have the limitation of providing similar features and Nearmiss method highly reduces the data instance that degrades the training performance. Future work of this method is applied in IoT based networks to test the efficiency of the proposed model.

Notation List

Notation	Description
s_w	Window size
$X \subseteq R^{M \times T}$	Training data
$X^{test} \subseteq R^{N \times T}$	Testing data
min, and max	Feature range
H_k	Hilbert space
M	Measurements
T	Streams
Z	Multivariate sub-sequence set
s_s	Step size
m	The number of sub-sequences
D	Discriminator
G	Generator
$A_t^{tes} \subseteq \mathcal{R}^{N \times 1}$	Testing dataset with label vector
$H(.,..)$	Cross Entropy Error
y_i	Labels
$x_i \in R^p$	Input data
$t \in R^{K-1}$	Translation vector
z_i'	Linear Function
$k: R^p \times R^p \to R^+$	Positive Definite Nucleus
K	Kernel matrix
C	Cost
i	Sample error
$h(q)$	Huber hinge loss
L_{MSVM}	Loss function of MSVM
ω_i	Optional sample weights
G_k	Sample set
k	Class
λ	Regularization term
$\lambda trWW'$	Penalty term
x_m	Unknown sample
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, have been done by 1st and 3rd author. Formal analysis, investigation, resources, data curation, writing—original draft preparation have been done by 5th author. Writing—review and editing, visualization, have been done by 6th author. The supervision and project administration, have been done by 2nd and 4th author.

References

[1] R. V. Mendonça, A. A. Teodoro, R. L. Rosa, M. Saadi, D. C. Melgarejo, P. H. Nardelli, and D. Z. Rodríguez, "Intrusion detection system based on fast hierarchical deep convolutional neural network", *IEEE Access*, Vol. 9, pp. 61024-61034, 2021.

[2] W. Liang, L. Xiao, K. Zhang, M. Tang, D. He, and K. C. Li, "Data fusion approach for collaborative anomaly intrusion detection in blockchain-based systems", *IEEE Internet of Things Journal*, Vol. 9, No. 16, pp. 14741-14751, 2021.

[3] S. N. Mighan and M. Kahani, "A novel scalable intrusion detection system based on deep learning", *International Journal of Information Security*, Vol. 20, No. 3, pp. 387-403, 2021.

[4] A. Yazdinejadna, R. M. Parizi, A. Dehghantanha, and M. S. Khan, "A kangaroo-based intrusion detection system on software-defined networks", *Computer Networks*, Vol. 184, p. 107688, 2021.

[5] N. Oliveira, I. Praça, E. Maia, and O. Sousa, "Intelligent cyberattack detection and classification for network-based intrusion detection systems", *Applied Sciences*, Vol. 11, No. 4, p. 1674, 2021.

[6] J. Liu, D. Yang, M. Lian, and M. Li, "Research on intrusion detection based on particle swarm optimization in IoT", *IEEE Access*, Vol. 9, pp. 38254-38268, 2021.

[7] X. Kan, Y. Fan, Z. Fang, L. Cao, N. N. Xiong, D. Yang, and X. Li, "A novel IoT network intrusion detection approach based on adaptive particle swarm optimization convolutional neural network", *Information Sciences*, Vol. 568, pp. 147-162, 2021.

[8] Z. Wang, Y. Zeng, Y. Liu, and D. Li, "Deep belief network integrating improved kernel-based extreme learning machine for network intrusion detection", *IEEE Access*, Vol. 9, pp. 16062-16091, 2021.

[9] M. Mulyanto, M. Faisal, S. W. Prakosa, and J. S. Leu, "Effectiveness of focal loss for minority classification in network intrusion detection systems", *Symmetry*, Vol. 13, No. 1, p. 4, 2021.

[10] X. Li, P. Yi, W. Wei, Y. Jiang, and L. Tian, "LNNLS-KH: a feature selection method for network intrusion detection", *Security and Communication Networks*, Vol. 2021, 2021.

[11] M. A. Khan, "HCRNNIDS: hybrid convolutional recurrent neural network-based network intrusion detection system", *Processes*, Vol. 9, No. 5, p. 834, 2021.

[12] T. Kim and W. Pak, "Hybrid classification for high-speed and high-accuracy network intrusion detection system", *IEEE Access*, Vol. 9, pp. 83806-83817, 2021.

[13] M. Ramaiah, V. Chandrasekaran, V. Ravi, and N. Kumar, "An intrusion detection system using optimized deep neural network architecture", *Transactions on Emerging Telecommunications Technologies*, Vol. 32, No. 4, p. e4221, 2021.

[14] M. Mulyanto, M. Faisal, S. W. Prakosa, and J. S. Leu, "Effectiveness of focal loss for minority classification in network intrusion detection systems", *Symmetry*, Vol. 13, No. 1, p. 4, 2021.

[15] M. Ajdani and H. Ghaffary, "Design network intrusion detection system using support vector machine", *International Journal of Communication Systems*, Vol. 34, No. 3, pp. e4689, 2021.

[16] C. M. Hsu, H. Y. Hsieh, S. W. Prakosa, M. Z. Azhari, and J. S. Leu, "Using long-short-term memory based convolutional neural networks for network intrusion detection", In: *Proc. of the International Wireless Internet Conference*, pp. 86-94, 2018.

[17] T. Su, H. Sun, J. Zhu, S. Wang, and Y. Li, "BAT: Deep learning methods on network intrusion detection using NSL-KDD dataset", *IEEE Access*, Vol. 8, pp. 29575-29585, 2020.

[18] S. S. Kaushik and P. R. Deshmukh, "Detection of attacks in an intrusion detection system", *International Journal of Computer Science and Information Technologies*, Vol. 2, No. 3, pp. 982-986, 2011.

[19] http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[20] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)", In: *Proc. of the Military Communications and Information Systems Conference (MilCIS)*, pp. 1-6, 2015.

[21] S. Iman, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", In: *Proc. of 4th International Conference on Information Systems Security and Privacy (ICISSP)*, pp. 108-16, 2018.