Cyber-Attacks in WSN & Security Optimization By a Novel Technique based Intensive Binary Pigeon Optimization (IBiPO) & Bi-LSTM-based IDS Framework

Faisal Nabi (faisal.nabi@jinnah.edu)
Muhammad Ali Jinnah University

Research Article

Keywords: Wireless Sensor Network (WSN), Intrusion Detection System (IDS), Data Preparation, Intensive Binary Pigeon Optimization (IBiPO), Bi-directional Long Short Term Memory (Bi-LSTM), Security

Posted Date: September 5th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3308713/v1

Abstract

With the global adoption of Internet services, service providers are having a difficult time securing their systems, especially against new attacks and intrusions. Various anomalous detection approaches have been developed for protecting WSN from cyber-attacks. However, those systems suffer from the major issues of a high number of false alarms, increased over-fitting, and complexity. Therefore, this paper motivates to develop a novel and intelligent IDS framework for protecting WSN from cyber-attacks. For this purpose, an Intensive Binary Pigeon Optimization (IBiPO) and Bi-directional Long Short Term Memory (Bi-LSTM) mechanisms are developed for accurate intrusion detection and classification.

INTRODUCTION

Wireless Sensor Networks (WSNs) [1, 2] is also termed as a heterogeneous system designed with the small controllers, sensors and generic processing components. Also, it is made up of thousands or hundreds of low-cost, self-organizing, wireless nodes that are used to monitor and regulate the environment. When creating a WSN, self-healing, dependability, adaptability, robustness, and security are the five primary factors that must be taken into account [3, 4]. It can also be used for a variety of military purposes, as well as for the monitoring of ocean, manufacturing equipment, earthquakes, and other natural disasters. In addition, it's likely that future applications may incorporate WSN concepts in their architectures, including those that monitor environment, transportation, site security, fires, and water quality. In this network [5], there may be one or more base stations, which are centralized control units. Then, a base station often serves as a gateway to another network, as well as a great data processing and storage facility and access point for human interaction. In order to retrieve data from the network and disseminate control information, it can also be utilized as a connector. It is imperative to guarantee a high level of security [6, 7] for the critical WSN applications in order to protect their data and infrastructure from breaches. In order to identify unusual activities and breaches, an Intrusion Detection System (IDS) [8] should be deployed. Moreover, it is a crucial component of security across any network type, since it provides the network with a high level of protection against potential dangers by stopping or identifying all intrusions and hosts. Its main objective is to make sure that every new attack may be detected. It is categorized into the types of misuse IDS and anomaly IDS [9], in which the anomaly IDS analyses statistical patterns and sophisticated ways to determine whether the behavior is healthy or not, whereas the misuse IDS uses signatures to find any new attacks.

In the existing works, various IDS frameworks [10, 11] have been developed to classify the intrusions or anomalies in the network. These techniques mostly involve intelligent classification methods and artificial intelligence algorithms. The pattern, detection rate, false alarm rate, and accuracy of each classifier can all be used to describe them. Moreover, various machine learning and deep learning based classification approaches are mainly used to accurately spot the intrusions in the network. For instance [12–14], the Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DR), Naïve Bayes (NB), and Artificial Neural Network (ANN) are the most popular and standard machine learning approaches used for developing an IDS framework. However, the traditional frameworks [15, 16] facing

some complications during the detection of network intrusions, which includes the followings: increased false alarm rate, complexity in system modeling, reduced detection accuracy, and high dimensionality of features. Therefore, this paper motivates to construct a novel and intelligent IDS framework for ensuring the security and confidentiality of WSN. The original contributions of this paper are as follows:

- To generate the balanced dataset for improving the process of intrusion detection and categorization, the data preparation is performed at first, which includes the operations of data cleaning, normalization, splitting, and clustering.
- To choose the optimal set of features for tuning the parameters of classifier, an Intensive Binary Pigeon Optimization (IBiPO) mechanism is deployed, which highly increases the accuracy and detection rate of classifier.
- To exactly predict the normal and attacking data flows by training and testing the optimized features, the Bi-directional Long Short Term Memory (Bi-LSTM) classification model is utilized.
- To assess the results of the proposed IBiPO + Bi-LSTM model, an extensive simulation analysis is carried out during performance evaluation.

The other sections of this paper are structured into the following units: Section 2 presents the complete literature review of the existing intrusion detection methodologies used for protecting WSNs. Also, it investigates the advantages and disadvantages of each mechanism based on its working model and operating principles. Section 3 provides the clear explanation for the proposed AI based IDS framework with its overall workflow and illustrations. Section 4 validates the results of both existing and proposed anomaly detection methodologies by using various parameters and datasets. Finally, the overall paper is summarized with the findings and future scope in Section 5.

RELATED WORKS

This section presents the complete literature review of existing methodologies used for developing an IDS framework to secure WSNs. It also examines the benefits and limitations of the existing works according to their anomaly detection process and operations.

Paul, et al [17] implemented a neuro-fuzzy based IDS framework for improving the security of WSNs. The purpose of this work was to develop a lightweight security mechanism for protecting WSN against harmful networking attacks. Here, the centralized approach has been utilized to enable reliable and valid data exchange in networks. Moreover, the integration of WSN with IoT networks could be one of the most difficult tasks, due to many security challenges. Amouri, et al [18] introduced a cross-layered IDS approach for detecting anomalies in the WSN. Here, the Accumulated Measure of Fluctuation (AMoF) has been utilized to accurately classify the attacks in the network. Sarkunavathi, et al [19] presented a comprehensive analysis to examine the different types of machine learning and deep learning techniques used for developing an effective IDS framework. This paper objects to attaining an increased attack classification accuracy with reduced false positives. Typically, the IDS used in WSN is categorized into the following types:

- Anomaly detection IDS
- Misuse detection IDS
- Clustering based IDS
- Hybridized IDS
- Trust enabled IDS
- Zone-based IDS

Moreover, an efficient IDS framework should satisfy the following security parameters for ensuring better intrusion detection performance. It includes energy efficiency, accuracy, memory, and network topology. Salfaldin, et al [20] implemented an improved binary grey wolf optimization algorithm for constructing an effective IDS framework. This work mentioned that the feature selection was the most essential stage in the IDS, since it helps to obtain a high accuracy with reduced redundancy and maximized relevancy. Zhang, et al [21] implemented a Multi-Kernel Extreme Learning Machine (MK-ELM) model for strengthening the security of WSN against horrible intrusions. The original contribution of this work was to obtain an increased detection accuracy with ensured Quality of Service (QoS). This algorithm incorporates the operations of both ELM and multi-kernel SVM for increasing the robustness and detection accuracy. Alwan, et al [22] utilized a Slime Mould Algorithm (SMA) for developing a new IDS framework for WSN. Elsaid, et al [23] developed an optimized collaboration based IDS framework for increasing the security of WSNs. This paper intends to improve the robustness, detection rate, and reduce the false alarm rate of classification. Typically, the WSN is highly vulnerable to network intrusions and attacks, hence it must be protected for ensuring the security and reliability of the network. Hence, the IDS is one of the most suitable option for WSN security, which supports to spot the intrusions or unauthenticated activities in the network by analyzing the features of network and data. Pan, et al [24] deployed a lightweight and intelligent intrusion detection model for guaranteeing the privacy, security, and confidentiality of WSN. Here, the K-Nearest Neighbor (KNN) algorithm incorporated with the Sine Cosine Algorithm (SCA) was utilized to minimize the false alarm rate and increase the classification accuracy of this detection framework. Moreover, the alarm response generated by the IDS could be used to block the intrusions or attacks in the network. In addition to that, the Polymorphic Mutation Strategy (PM) has been utilized to choose the features for analyzing the characteristics of attacks. The advantages of this work are minimal computational complexity, ensured system robustness and reliability. However, the time required to train and test the data samples are increased, whi ch degrades the efficacy of the suggested system.

Gowdhaman, et al [25] used a Deep Neural Network (DNN) approach to deal with the unbalanced attacks in the WSNs. In this case, the cross correlation has been used to effectively choose the pertinent features from the datasets for correctly identifying the intrusions. Sood, et al [26] utilized a conditional Generative Adversarial Network (GAN) model for protecting the WSN against the harmful network intrusions. This research work focused on an unsupervised learning method and how it may be used to create secure IDS. Also, it generated some fictitious data to mislead the attacker. In contrast to other deep learning based IDS models, it can secure the network and transport data between the sender and the recipient. However,

it failed to prove the detection accuracy and QoS of the suggested model, which could be major limitation of this work. Masengo, et al [2] suggested an AI based anomaly detection model for an integrated Software Defined WSN (SDWSN) platform. The purpose of this paper was to analyze the efficacy and performance of various classification techniques such as DT, NB, and deep ANN for developing a computationally intelligent IDS framework for securing SDWSN. In order to analyze the efficacy of these models, the prediction time, run time, and memory size have been estimated in this work. Also, this study stated that the deep ANN model outperforms other techniques with improved performance values. Karthic, et al [27] introduced a hybrid optimized DNN for detecting intrusions in the WSNs. Here, the standard CNN model is incorporated with the LSTM framework for identifying and categorizing the class of intrusions. Moreover, an enhanced conditional random field based feature selection mechanism was also used to simplify the process of feature learning. Due to an efficient learning of features, the overall detection accuracy of the suggested framework was highly improved. However, it could be difficult to understand the system model, due to the complexity in computational operations.

Rezvi, et al [28] implemented a new data mining technique for developing an effective IDS framework. Here, the dataset preprocessing was performed to characterize the attack types into the discrete values. Then, the different types of classification models such as ANN, KNN, SVM, LR, and NB have been validated to choose the most efficient technique for an accurate intrusion detection and classification. In addition, the SMOTE analysis was performed to estimate the prediction rate of the suggested framework. Di Mauro, et al [29] suggested a Weightless Neural Network (WNN) for an effective IDS framework. Here, the attack types of were classified according to their features such as coarse grained features, flow based features, time based features, byte based features, packet based features, and flag based features. Yet, it required to minimize the classification time, which affects the performance of classifier. Halbouni, et al [30] utilized a CNN-LSTM classification technique for designing a competent IDS with increased accuracy and highest detection rate. This framework includes the operating stages of data encoding, normalization, optimization, and classification.

The survey found that existing IDS frameworks are primarily concerned with increasing detection rates, lowering false positives, and boosting learning effectiveness.

However, it has the following issues:

- The features' testing and training take a lot of time.
- Complicated feature extraction and selection processes.
- Oversampling.
- Lack of reliability.
- Difficult to implement.

Therefore, the goal of the proposed work is to create a new security paradigm that will shield WSN against damaging intrusions or abnormalities.

PROPOSED METHODOLOGY

This section provides the clear explanation for the proposed IDS framework with its overall workflow and illustrations. The original contribution of this paper is to develop a computationally intelligent IDS framework for securing WSN from network intrusions. For this purpose, a novel and efficient Intensive Binary Pigeon Optimization (IBPO) technique incorporated with a Bi-directional Long Short Term Memory (Bi-LSTM) models are implemented. The overall workflow of the proposed system is depicted in Fig. 1, which includes the following operations:

- Data preparation
- Intensive Binary Pigeon Optimization (IBiPO)
- Bi-Directional LSTM (Bi-LSTM) classification
- Performance evaluation

Here, the popular and public IDS datasets are used for system implementation, which are preprocessed at the initial stage with the data cleaning, normalization, splitting, and clustering operations. Then, the balanced dataset is used for further optimization and classification processes. The IbiPO technique is mainly used to optimally select the most relevant features for accurately predicting the intrusions with reduced false alarms. Moreover, this technique helps to simplify the process of intrusion identification and classification with high accuracy. Then, the obtained features are passed to the Bi-LSTM classifier for training and testing. This classifier predicts the normal and anomalous data based on the training features. The advantages of the proposed IbiPO + Bi-LSTM model are reduced overfitting, false alarms, time consumption, and high detection accuracy.

• Data Preparation

Before processing and evaluating the data, it is highly essential to prepare the balanced dataset. The data preparation holds the major operations of data cleaning, normalization, transformation, clustering and compression. In which, the process of eliminating redundant or irrelevant entries and addressing missing data is known as data cleaning. It is a crucial step in making sure the data is reliable, accurate, and useable. Moreover, it is more important to eliminate the duplicate entries in the given dataset to keep the classifiers from learning rare records and from being biased toward the most common records. Moreover, the transformation of symbolic data into numerical values and label transfer are considered as the additional data cleaning operations. In this work, three different datasets (NSL-KDD, CICIDS 2018 and UNSW-NB15) have the class labels with symbolic values like "normal" or "intrusion type." Then, the created IDS tries to distinguish the legitimate and malicious communications without disclosing the nature of the assault. Consequently, the process of scaling or changing each feature's data values into a proportional range is known as data normalization. Here, the given dataset is normalized into the range of 0 to 1 as represented in below:

$$DS_N = \frac{(DS - DS^{Min})}{(DS_{Max} - DS_{Min})}$$

(1)

Where, DS indicates the IDS dataset, DS_N is the normalized dataset, DS_{Min} and DS_{Max} are the minimum and maximum values of dataset. Subsequently, the data splitting and clustering operations are also performed to generate the balanced dataset, where training dataset is split into two such as training and validation. In which, the training set is to train the entire model, and the validation set is used to test the model at the time of parameter tuning. Then, the distance based clustering mechanism is applied to reduce the size of dataset. During this process, the random centroid is selected at first, and the data points are assigned to the nearest cluster according to its distance or similarity. Each cluster's centroid is calculated as the average of all the data points that belong to that cluster after all the data points have been assigned to the closest group. The procedure of assigning the data points to the new cluster's centroid is then repeated until the centroid's values remain consistent. Finally, the preprocessed dataset is generated and used for further operations.

• Intensive Binary Pigeon Optimization (IBiPO)

In this stage, the features of the preprocessed dataset are extracted by using the IBiPO technique. It is mainly used to obtain an optimal set of features for reducing the dimensionality of dataset, which also supports to increased detection accuracy and performance of classifier. The IBiPO is a meta-heuristic optimization technique, which comprises three operators such as landmark, map, and compass. Among other optimization techniques, the key merits of using this approach are as follows: high convergence speed, reduced overfitting, and easy to deploy. In this technique, the pigeons perceive the geomagnetic fields in the map and compass operators to create a map for homing. Let consider that, the searching space having N dimensions, and i pigeons of swarms as represented in below:

$\boldsymbol{S_i} = (\boldsymbol{S_{i,1}}, \boldsymbol{S_{i,2}} \dots \boldsymbol{S_{i,N}})$ (2)

Then, the velocity of pigeon is represented based on its changing location in the N dimensional vector as shown in below:

$\boldsymbol{Q_i} = (\boldsymbol{Q_{i,1}}, \boldsymbol{Q_{i,2}} \dots \boldsymbol{Q_{i,N}})$ (3)

Similarly, the visited locations of the ith pigeons are represented as follows:

$\boldsymbol{Y_i} = (\boldsymbol{Y_{i,1}}, \boldsymbol{Y_{i,2}} \dots \boldsymbol{Y_{i,N}})$ (4)

Consequently, the global optimal location of the pigeons are considered as $(\boldsymbol{K_{i,1}}, \boldsymbol{K_{i,2}} \dots \boldsymbol{K_{i,N}})$, and all pigeons can fly in the searching space by using the following model:

$\boldsymbol{Q_i(h+1)} = \boldsymbol{Q_i(h)} \times \boldsymbol{e^{-Gh}} + \boldsymbol{a} \times (\boldsymbol{S_g} - \boldsymbol{Q_i(h)})$ (5)

$$S_i(h+1) = S_i(h) + Q_i(h+1) \quad (6)$$

Where, $S_i(h)$ denotes the present location of pigeon at time h, G denotes the map and compass factors, a is the arbitrary value ranging between 0 to 1, S_g is the global optimal solution, Q_i(h) represents the velocity at time h. By using the landmark operator, each pigeon in the searching space is ranked according to its fitness value, and the total number of pigeons are upgraded by using the following model:

$XPig$

$$(h+1) = \frac{X_{Pig}(h)}{2} \quad (7)$$

Where, X_{Pig} denotes the total number of pigeons at iteration h. Then, the location of pigeons are updated by using the following equations:

CP

$$S \quad (h+1) = \frac{\sum S_i(h+1) \times FF(S_i(h+1))}{X_{Pig} \sum FF(S_i(h+1))} \quad (8)$$

$$S_i(h+1) = S_i(h) + a \times (S_{CP}(h+1) \times S_i(h)) \quad (9)$$

Where, S_{CP} represents the location of center pigeon, and FF is the fitness function. In the proposed IBiPO, the solution is improved toward a continuous valued position in the searching space, which is the major difference between the standard and proposed pigeon optimization techniques. Finally, the fitness function is estimated for determining the optimal solution as shown in below:

G

$$FF = \Delta(\varepsilon) + \sigma^{|a|}|\beta| \quad (10)$$

Where, is the random parameter in the range of [0, 1], σ denotes the importance of reduction features,

$\Delta_G(\varepsilon)$ is the error rate of classifier, $|a|$ represents the subset size, and $|\beta|$ indicates the overall features in the dataset. By using this function, the optimal set of features are selected from the preprocessed dataset, which is used for classifier training and testing processes.

• Bi-Directional LSTM (Bi-LSTM) Classification

After feature optimization, the selected subset of features are used for classification, where the Bi- LSTM model is used to accurately spot the intrusions in WSN. Typically, the Bi-LSTM is a kind of sequence

processing mechanism, which comprises two LSTM models. In which, one LSTM will receive input going forward, and the other will receive input going backward. The effectiveness of the model is increased when the LSTM is applied twice because it changes how long-term interdependencies are learned. These dependencies can be monitored as the sequence progresses. The LSTM is made to avoid the long-term dependence issue by recollecting the data for a lengthy period of time and incorporating a memory cell. Moreover, it comprises three gates such as input gate, forget gate, and output gate, in which the input gate determines how much additional data will indeed be transferred to the memory, the output gate determines if the current value in the cell subsidizes to the output, and the forget gate determines whether to retain or discard available data. In neural networks, activation functions are used to estimate the weighted sum of inputs and biases, then it determines whether a neuron can activate or not.

A gate-like function known as an activation function, it verifies if an incoming value is greater than a threshold value. Moreover, the Bi-LSTM enhances the LSTM predecessor by integrating backward posterior probability to the already-existing forward hidden states, giving it a forward-looking capability similar to the hidden Markov model. Figure 2 depicts the structure of the Bi-LSTM, which can effectively use the temporal properties available in the contextual information to enhance the model training for network traffic.

RESULTS AND DISCUSSION

This section validates the results and performance of the proposed IDS model by using various evaluation indicators and datasets. For this assessment, the different types of network intrusion datasets have been used, which hold NSL-KDD, UNSW-NB 15, and CICIDS-2018. In which, four different types of assaults, including DoS, Probe, R2L, and U2R, are included in the NSL-KDD dataset. The remaining data is classified as typical data and only comes within the four categories. Under the four categories, 39 various attacks are grouped together, and rather than detailing each attack individually, each attack is mapped into the corresponding group. The performance metrics used in this study are computed as follows:

$$\text{Accuracy} = \frac{T_{pos}+T_{neg}}{Tpos+Tneg+Fpos+Fneg}$$

$$\text{Precision} = \frac{T_{pos}}{Tpos+Fpos}$$

$$\text{Recall or TPR} = \frac{T_{pos}}{Tpos+Fneg}$$

$$F1-\text{score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{(\text{Precision}+\text{Recall}}\right)$$

$$\text{FPR} = \frac{F_{pos}}{Fpos+Tneg}$$

$$\text{FNR} = \frac{F_{neg}}{Tpos+Fneg}$$

Where, $\mathbf{T_{pos}}$ indicates true positive, $\mathbf{T_{neg}}$ denotes true negative, $\mathbf{F_{pos}}$ indicates false positive, $\mathbf{F_{neg}}$ denotes false negative. Figure 3 compares and illustrates the detection accuracy of the proposed IDS model with that of conventional anomaly detection models such as SVM, RF, DT, and DNN [25]. In comparison to traditional algorithms, it is seen that the proposed IDS framework achieves the highest detection accuracy. Also, the proposed model classifies the intrusion more effectively than the existing methods due to optimal feature selection, effective use of hidden layers, and deep features.

The comparison of the precision values for the proposed and existing methodologies is shown in Fig. 4. From the results, it can be shown that the suggested intrusion detection model performs better than other strategies, with the other model lagging behind due to ineffective training and feature selection processes. The proposed IDS model has an average precision value of almost 99%, compared to existing models' average precision values of 70–80%.

A comparison analysis for the recall parameter is shown in Fig. 5, where the investigation shows that the proposed model has maximum recall values. The estimated results indicate that the proposed model can detect the greatest number of intrusions while traditional SVM, DT, RF, and DNN models perform less well. The proposed model has a recall percentage that is 30% higher than that of SVM, DT, RF, and DNN. Figure 6 shows the results of a comparison of F1 scores, and it shows that the proposed IDS model outperforms more traditional machine learning and deep learning based IDS methods. Moreover, Fig. 7 shows a thorough comparison of all the techniques used for the various types of attacks in the dataset. As can be seen, the suggested model's detection rate is higher than that of other approaches for all attacks.

In the proposed work, the detection rate is increased with reduced computational complexity, due to the deployment of IBPO algorithm. Since, the large dimensional dataset requires increased time to train the classifier, and also which affects the detection efficiency of IDS model. Hence, the proposed work intends to obtain the most relevant features used for classifier training and testing operations, which supports to achieve an increased detection rate with low powered devices. Table 1 provides a summary of the overall performance of the proposed model versus traditional machine learning & deep learning models. The investigation shows that the proposed IDS model has a maximum accuracy of 98.8%, which is high superior than the other techniques. Due to the best feature processing and selection, the proposed model has a better computation ability and system efficacy. The use of proposed optimization + classification model highly improves the accuracy of intrusion detection performance, but the performance of standard machine learning & deep learning algorithms declines as a result of inadequate feature selection and processing. Finally, the findings show that the suggested model may successfully identify intrusions in sensor networks.

Table 1
Performance comparative analysis

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM	65.95	67.50	69.85	56.36
DT	77.99	75.37	75.59	71
RF	81.73	80.10	74.76	68.40
DNN	95.53	94.65	91.92	92.43
Proposed	98.8	99.1	98.9	98.7

Table 2
Intrusion detection rate

Type of intrusions	SVM	ELM	MK-ELM	Proposed
Normal	97.73	97.92	99.12	99.56
DoS	96.24	97.15	98.03	99.1
Probe	93.75	94.54	95.74	98.9
R2L	55.26	65.03	76.15	99
U2R	30.73	23.02	50	98.9

Additionally, Table 3 compares the proposed IDS model with other models by using UNSW-NB 15 dataset. For this evaluation, the existing SVM, ELM, and MK-ELM [21] techniques are considered, which have

comparable accuracy. However, the accuracy of the proposed IDS model is the highest, and the accuracy of all three algorithms has decreased when compared to the proposed technique.

Table 3
Performance evaluation using UNSW-NB 15 dataset

Parameters	SVM	ELM	MK-ELM	Proposed
Accuracy	88.20	87.20	92.10	99
True positive rate	83.73	83.84	89.42	98.9
False positive rate	2.34	3.76	2.37	1.21
False negative rate	16.27	16.16	10.58	2.54

Table 4 and Fig. 9 compare the average detection rate of the existing and proposed anomaly detection mechanisms with respect to different types of attacks in the NSL-KDD dataset. Here, the average detection rate for three algorithms running 50 times, when the training dataset and test dataset together have a size of 14,000. The obtained results indicate that the SVM algorithm's detection rate is comparable to that of the fundamental ELM algorithm. Then, the proposed IDS model has the highest detection rate, which is around 3% higher than the detection rates of the other three algorithms. Moreover, the testing data1000, 2000, 4000, 8000, and 14,000 are chosen for comparison with the existing methods.

Table 4
Average detection rate with respect to different types of intrusions

Type of intrusions	SVM	ELM	MK-ELM	Proposed
Normal	97.73	97.92	99.12	99.5
DoS	96.24	97.15	98.03	98.9
Probe	93.75	94.54	95.74	99.1
R2L	55.26	65.03	76.15	99
U2R	30.73	23.02	50	99.1

For validating the time consumption, the chosen experiments include 5000, 10,000, 15,000, 20,000, and 25,000 test data. In terms of time consumption, the findings shown in Fig. 11 and Table 5 show that the fundamental MK-ELM outperforms ELM and SVM. Additionally, the proposed IDS model beats SVM in terms of speed and accuracy due to the performance degradation of the existing models. This implies that the proposed technique exhibits greater scalability than the other algorithms, when classifying multiclass traffic for intrusion detection.

Table 5
Time consumption (s)

Amount of Data	SVM	ELM	MK-ELM	Proposed
5000	10	5	8	4
10000	12	6	7	5
15000	18	8	12	6.5
20000	38	12	20	8
25000	78	25	40	15

Table 6 and Fig. 11 compares the detection accuracy of existing and proposed anomaly detection models by using UNSW-NB15 dataset. Here, ten classifications are tested together with the vectorization of the UNSW- NB 15 dataset's label into ten categories. The testing set and preprocessed training dataset are chosen at random to contain 10,000, 15,000, 20,000, and 25,000 bits of data, respectively, in varied amounts. Overall, the obtained results indicate that the proposed IDS framework outperforms the other techniques with highly improved results, which shows the overall efficacy and increased attack detection performance of the proposed model.

Table 6
Accuracy

Amount of Data	SVM	ELM	MK-ELM	Proposed
10000	83	86	90	99.2
15000	83	87	88	99
20000	81	84	89	98.9
25000	78	85	89	98.9

Execution time (s)

Energy Consumption (J)

Table 7
Overall expenses

Components/Parameters	Specifications
Sensor	Raspberry Pi 4
Processor	Cortex-A72 64-bit
Cores	4
Memory	4GB
Storage	16 GB
Connectivity	Bluetooth: v5.0 & WiFi

Consequently, Fig. 12 depicts the execution time analysis of the proposed security model with respect to varying amount of data. Then, the energy consumption of the proposed framework is evaluated and depicted. According to the results, it is estimated that the proposed IBiPO + Bi-LSTM technique requires reduced execution time and energy consumption by effectively predicting the intrusions based on proper training and testing operations. Table 7 presents the overall expenses analysis of the proposed IDS framework. Table 8 presents the overall comparative analysis of the existing and proposed anomaly detection methodologies based on the parameters of false alarm rate, detection rate, accuracy, and execution time. Based on the study, it is determined that the proposed model outperforms other approaches with better prediction results.

Table 8
Overall analysis

IDS Methods	False alarm rate	Detection rate	Accuracy	Execution time
Trust based IDS	Low	Very High	Low	NA
ARIMA – traffic anomaly detection	Low	Very High	Low	High
Lightweight IDS	Very Low	Very High	High	NA
Sensor anomaly Detection	Low	Very High	High	NA
PSO-IDS	High	Low	High	NA
Evolutionary NN-IDS	Very High	Very High	Very High	NA
Proposed IBiPO + Bi-LSTM	Very Low	Very High	Very High	Very Low

CONCLUSION

In this paper, a novel and computational efficient IDS framework, named as, IBiPO + Bi-LSTM model is proposed for securing WSN. The original contribution of this paper is to highly protect the network from the harmful intrusions or anomalies. In this context, the system is implemented using the well-known and open IDS datasets, which were initially preprocessed using data cleaning, normalization, splitting, and clustering processes. The balanced dataset is then applied to additional procedures of optimization and classification. The IBiPO method is mainly used to choose the most pertinent information best in order to anticipate intrusions effectively with less false alarms. Additionally, this method aids in the high accuracy detection and simplification of intrusion classification. Following that, the Bi-LSTM classifier receives the collected features for training and testing. Based on the learning features, this classifier predicts the normal and anomalous data. The suggested IBiPO + Bi-LSTM model has the advantages of low overfitting, quick processing, and excellent detection accuracy. During performance analysis, the system

is assessed by contrast against recent relevant works in terms of DR, FPR, accuracy, precision, recall, and time consumption. Moreover, three well-known IDS datasets (NSL-KDD, CICIDS 2018 and UNSW-NB15) were utilized in all experiments for evaluation. Using the three aforementioned datasets, the suggested system performs better than the existing models.

In future, the present work can be enhanced by implementing an IDS framework for the smart application systems. Moreover, it targets to deploy the proposed IDS framework for medical sciences, industrial applications, and military services in order to ensure the high level of security.

DECLARATIONS

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Faisal Nabi reports statistical analysis was provided by Mohammad Ali Jinnah University. Faisal Nabi reports a relationship with Mohammad Ali Jinnah University that includes: employment. Faisal Nabi has patent n/a pending to n/a. we both worked hard in this paper.

REFERENCES

1. S. Shakya, "Modified Gray Wolf Feature Selection and Machine Learning Classification for Wireless Sensor Network Intrusion Detection," IRO Journal on Sustainable Wireless Systems, vol. 3, pp. 118-127, 2021.
2. S. Masengo Wa Umba, A. M. Abu-Mahfouz, and D. Ramotsoela, "Artificial Intelligence-Driven Intrusion Detection in Software-Defined Wireless Sensor Networks: Towards Secure IoT-Enabled Healthcare Systems," International Journal of Environmental Research and Public Health, vol. 19, p. 5367, 2022.
3. K. Hussain, Y. Xia, A. N. Onaizah, T. Manzoor, and K. Jalil, "Hybrid of WOA-ABC and Proposed CNN for Intrusion Detection System in wireless sensor networks," Optik, p. 170145, 2022.
4. G. Singh and N. Khare, "A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques," International Journal of Computers and Applications, vol. 44, pp. 659-669, 2022.
5. M. A. Hamzah and S. H. Othman, "Performance Evaluation of Support Vector Machine Kernels in Intrusion Detection System for Wireless Sensor Network," International Journal of Innovative Computing, vol. 12, pp. 9-15, 2022.
6. G. Sadineni, M. Archana, and R. C. Tanguturi, "Improved Practical Enabled Component Analysis for Intrusion Detection in Wireless Sensor Networks," Journal of Optoelectronics Laser, vol. 41, pp. 232-240, 2022.
7. V. Ravi, R. Chaganti, and M. Alazab, "Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system," Computers and Electrical Engineering, vol. 102, p. 108156, 2022.
8. M. Maheswari and R. Karthika, "A novel QoS based secure unequal clustering protocol with intrusion detection system in wireless sensor networks," Wireless Personal Communications, vol. 118, pp.

1535-1557, 2021.

9. A. Sarkunavathi, V. Srinivasan, and M. Ramalingam, "Dense Net RNN–An Intrusion Prevention System to Mitigate DoS Attacks in Wireless Sensor Networks," 2022.
10. R. Yadav, I. Sreedevi, and D. Gupta, "Augmentation in performance and security of WSNs for IoT applications using feature selection and classification techniques," Alexandria Engineering Journal, 2022.
11. M. Srivastava, S. S. Yadav, and J. Dheeba, "A Novel Secured Wireless Sensor Network with Ensemble based Intrusion Detection System and Middleware Architecture," in 2022 International Conference on IoT and Blockchain Technology (ICIBT), 2022, pp. 1-6.
12. S. Karthic, S. Manoj Kumar, and P. Senthil Prakash, "Grey wolf based feature reduction for intrusion detection in WSN using LSTM," International Journal of Information Technology, pp. 1- 6, 2022.
13. S. Ifzarne, H. Tabbaa, I. Hafidi, and N. Lamghari, "Anomaly detection using machine learning techniques in wireless sensor networks," in Journal of Physics: Conference Series, 2021, p. 012021.
14. A. Jamalipour and S. Murali, "A taxonomy of machine learning based intrusion detection systems for the internet of things: A survey," IEEE Internet of Things Journal, 2021.
15. M. A. Hamzah and S. H. Othman, "A Review of Support Vector Machine-based Intrusion Detection System for Wireless Sensor Network with Different Kernel Functions," International Journal of Innovative Computing, vol. 11, pp. 59-67, 2021.
16. T. Moulahi, S. Zidi, A. Alabdulatif, and M. Atiquzzaman, "Comparative performance evaluation of intrusion detection based on machine learning in in-vehicle controller area network bus," IEEE Access, vol. 9, pp. 99595-99605, 2021.
17. A. Paul, S. Sinha, R. N. Shaw, and A. Ghosh, "A neuro-fuzzy based IDS for internet-integrated WSN," in Computationally Intelligent Systems and their Applications, ed: Springer, 2021, pp. 71- 86.
18. A. Amouri, S. D. Morgera, M. A. Bencherif, and R. Manthena, "A cross-layer, anomaly-based IDS for WSN and MANET," Sensors, vol. 18, p. 651, 2018.
19. A. Sarkunavathi, V. Srinivasan, and M. Ramalingam, "Comprehensive Analysis of Intrusion Prevention and Detection System and Dataset used in WSN using Machine Learning & Deep Learning," Mathematical Statistician and Engineering Applications, vol. 71, pp. 638-657, 2022.
20. M. Safaldin, M. Otair, and L. Abualigah, "Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks," Journal of ambient intelligence and humanized computing, vol. 12, pp. 1559-1576, 2021.
21. W. Zhang, D. Han, K.-C. Li, and F. I. Massetto, "Wireless sensor network intrusion detection system based on MK-ELM," Soft Computing, vol. 24, pp. 12361-12374, 2020.
22. M. H. Alwan, Y. I. Hammadi, O. A. Mahmood, A. Muthanna, and A. Koucheryavy, "High Density Sensor Networks Intrusion Detection System for Anomaly Intruders Using the Slime Mould Algorithm," Electronics, vol. 11, p. 3332, 2022.

23. S. A. Elsaid and N. S. Albatati, "An optimized collaborative intrusion detection system for wireless sensor networks," Soft Computing, vol. 24, pp. 12553-12567, 2020.
24. J.-S. Pan, F. Fan, S.-C. Chu, H.-Q. Zhao, and G.-Y. Liu, "A Lightweight Intelligent Intrusion Detection Model for Wireless Sensor Networks," Security and Communication Networks, vol. 2021, 2021.
25. V. Gowdhaman and R. Dhanapal, "An intrusion detection system for wireless sensor networks using deep neural network," Soft Computing, vol. 26, pp. 13059-13067, 2022.
26. T. Sood, S. Prakash, S. Sharma, A. Singh, and H. Choubey, "Intrusion Detection System in Wireless Sensor Network Using Conditional Generative Adversarial Network," Wireless Personal Communications, pp. 1-21, 2022.
27. S. Karthic and S. M. Kumar, "Hybrid Optimized Deep Neural Network with Enhanced Conditional Random Field Based Intrusion Detection on Wireless Sensor Network," Neural Processing Letters, pp. 1-21, 2022.
28. M. A. Rezvi, S. Moontaha, K. A. Trisha, S. T. Cynthia, and S. Ripon, "Data mining approach to analyzing intrusion detection of wireless sensor network," Indonesian J. Electric. Eng. Comput. Sci, vol. 21, pp. 516-523, 2021.
29. M. Di Mauro, G. Galatro, and A. Liotta, "A WNN-Based Approach for Network Intrusion Detection," in International Symposium on Intelligent and Distributed Computing, 2022, pp. 79-88.
30. A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, "CNN- LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System," IEEE Access, vol. 10, pp. 99837-99849, 2022.

Figures

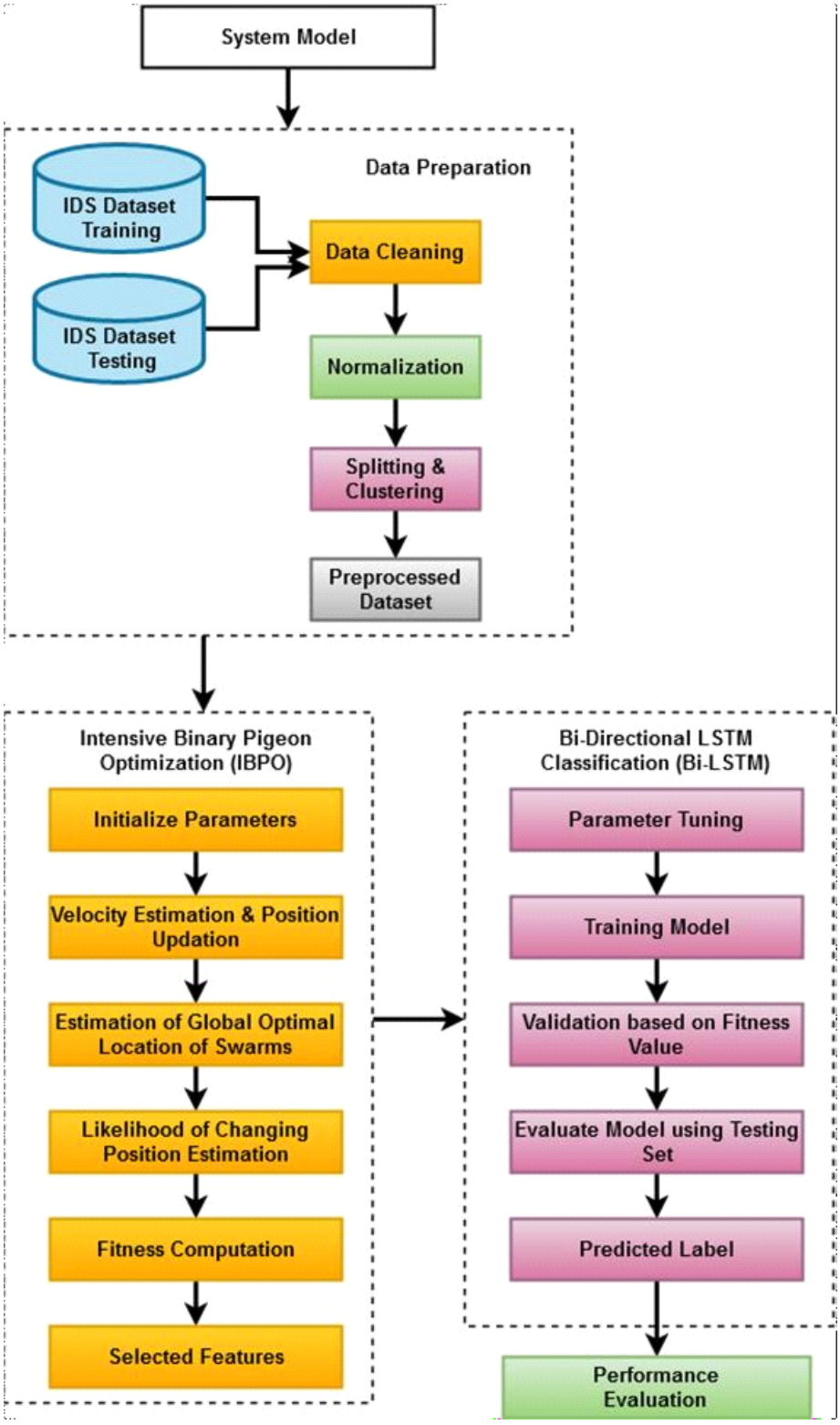


Figure 1

Work flow of the proposed system

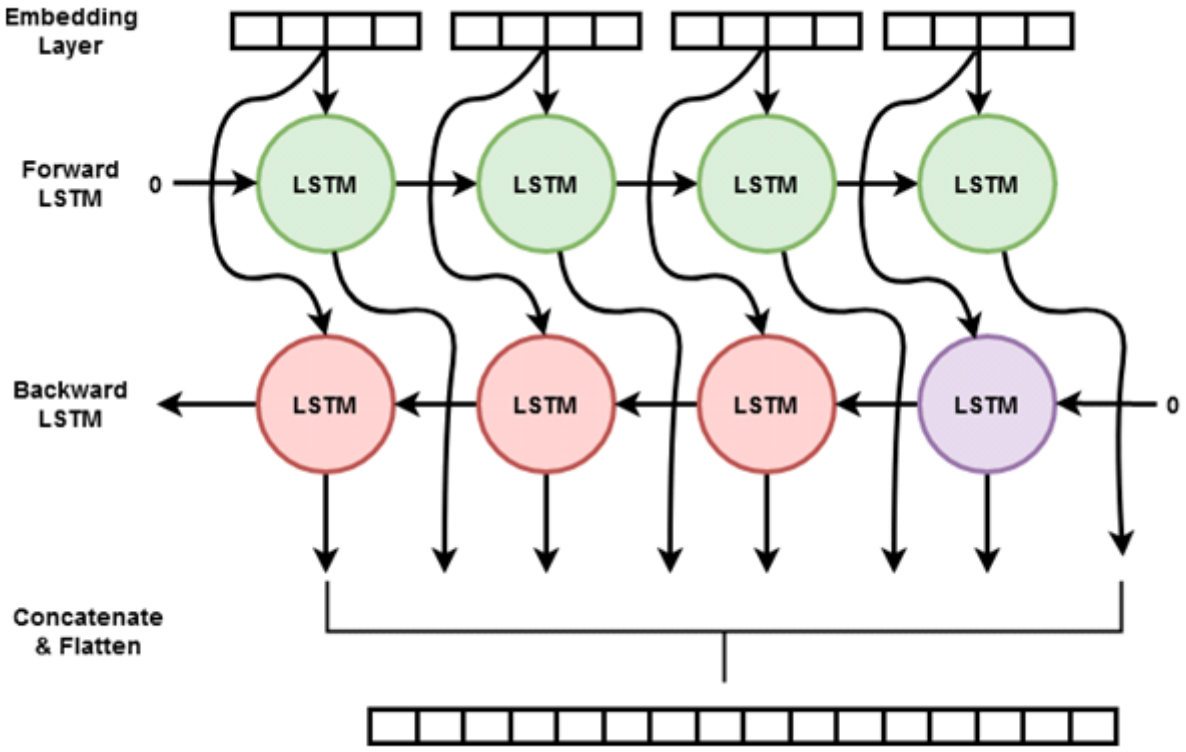


Figure 2

Bi-LSTM architecture

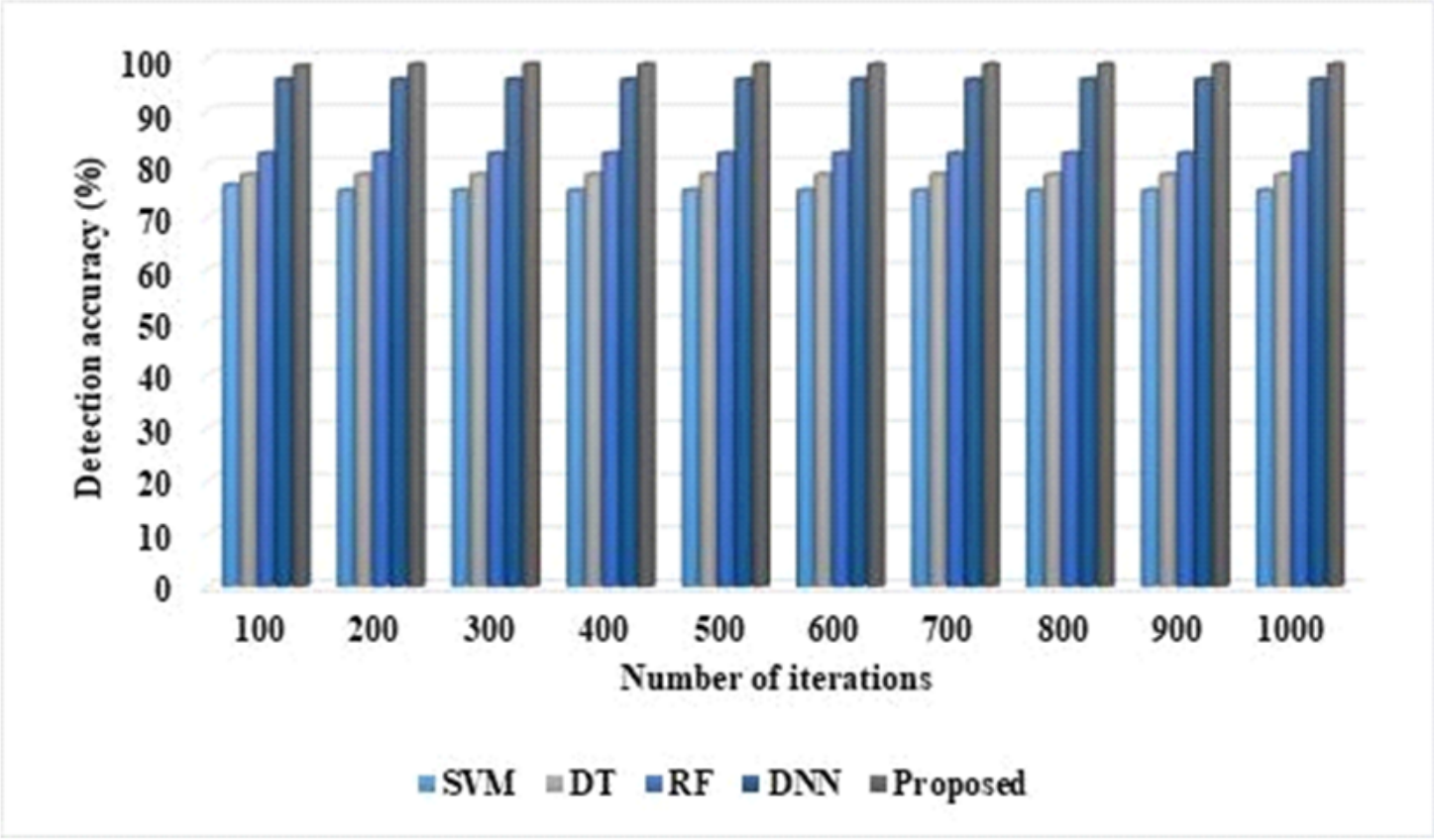


Figure 3

Detection accuracy Vs Number of iterations

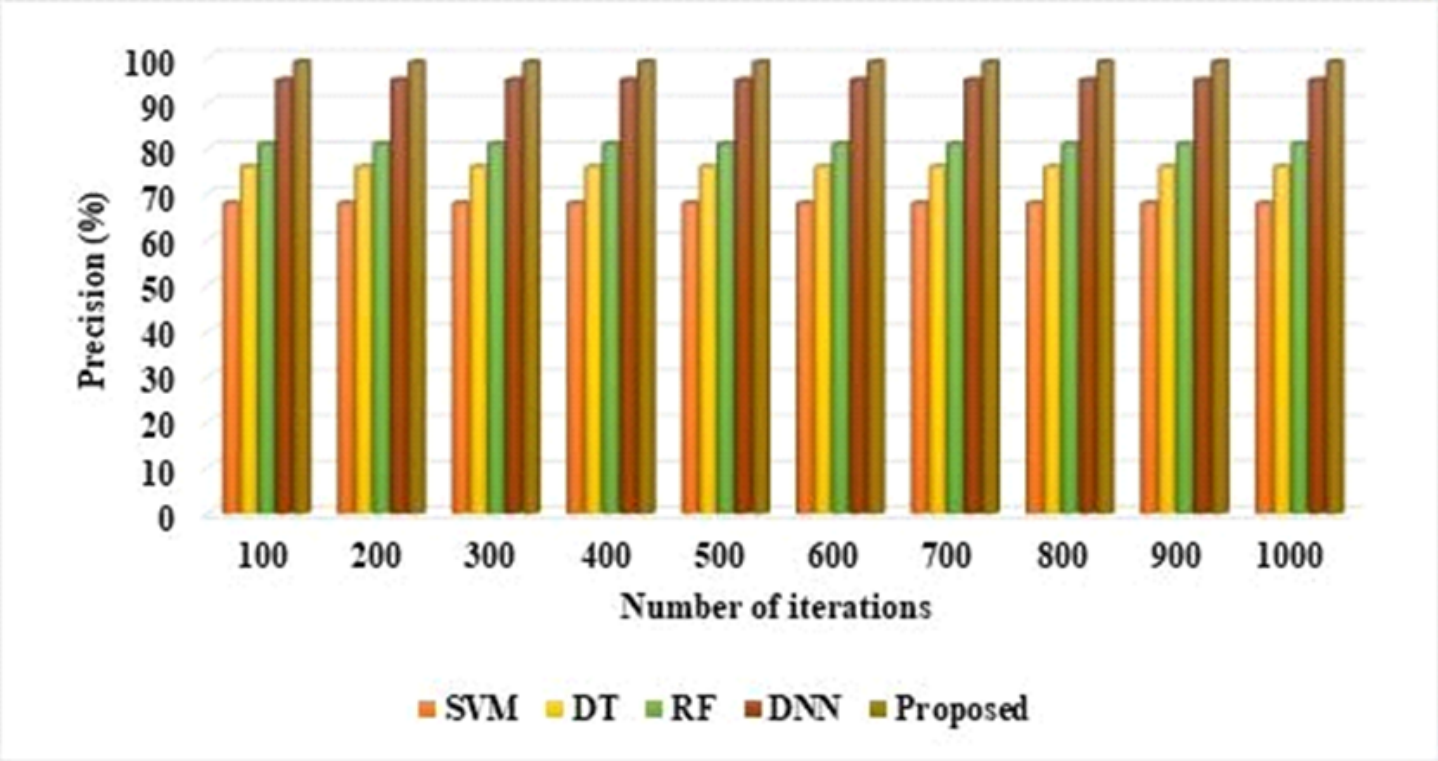


Figure 4

Precision Vs Number of iterations

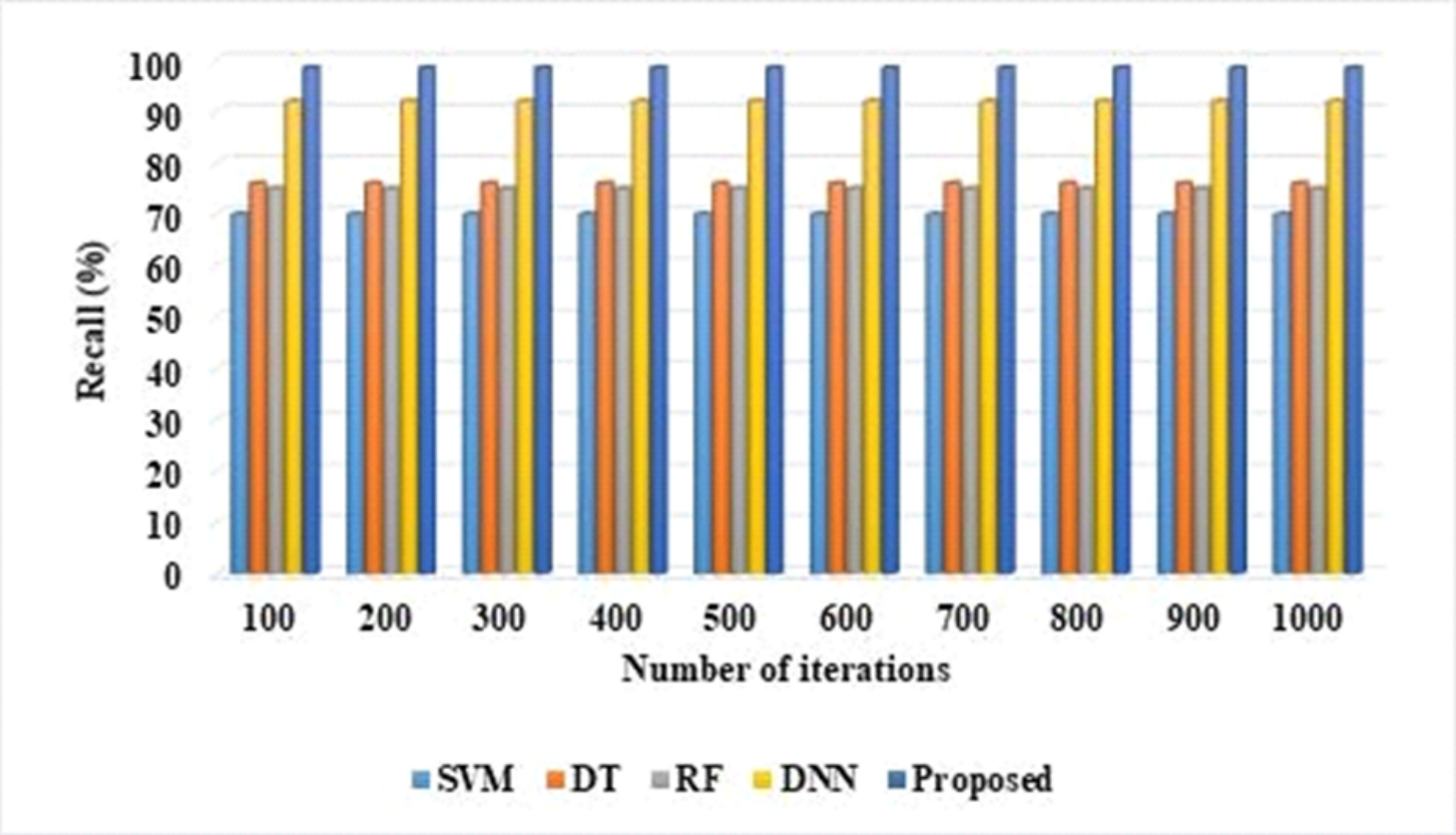


Figure 5

Recall Vs Number of iterations

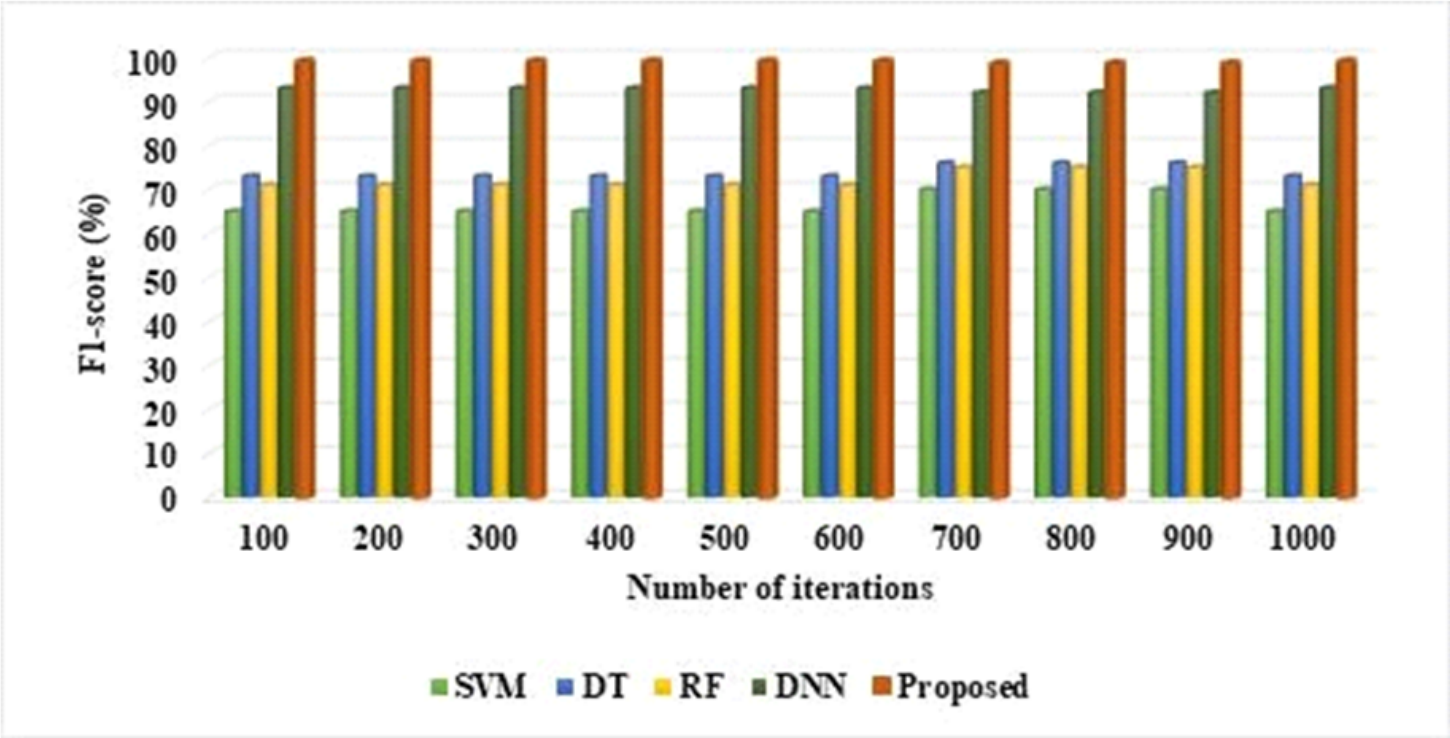


Figure 6

F1-score Vs Number of iterations

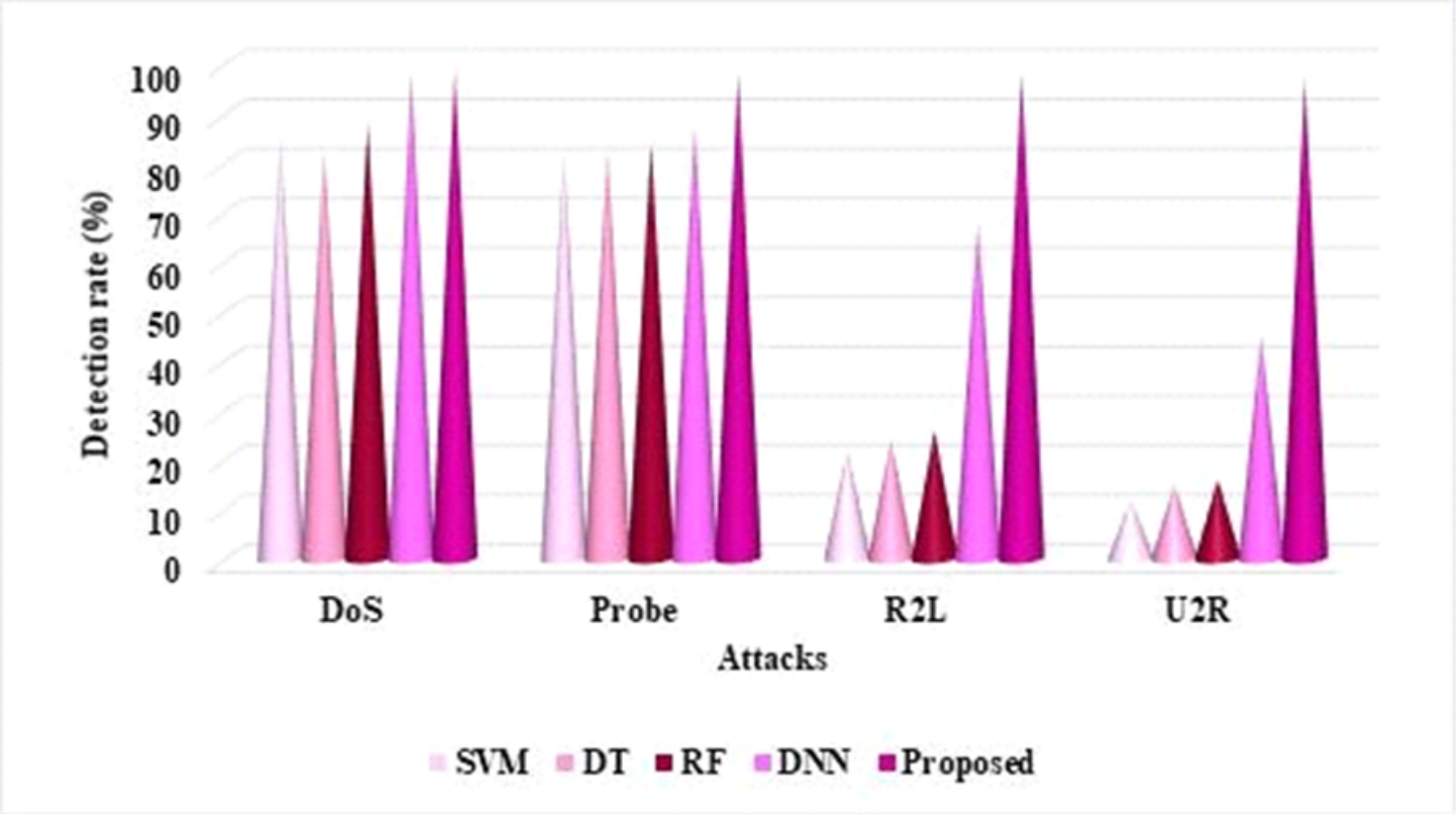


Figure 7

Comparative analysis with respect to different types of attacks

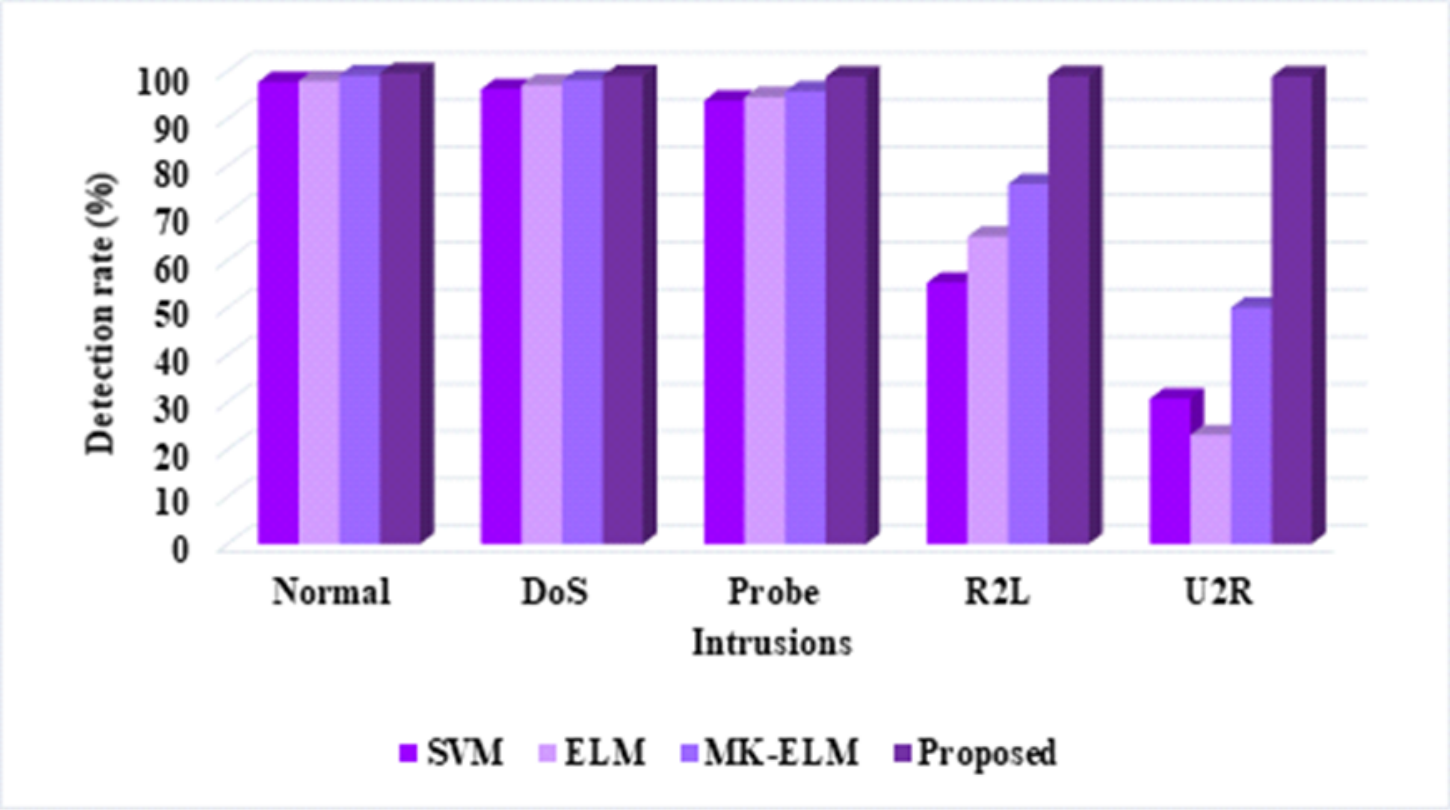


Figure 8

Analysis of intrusion detection rate

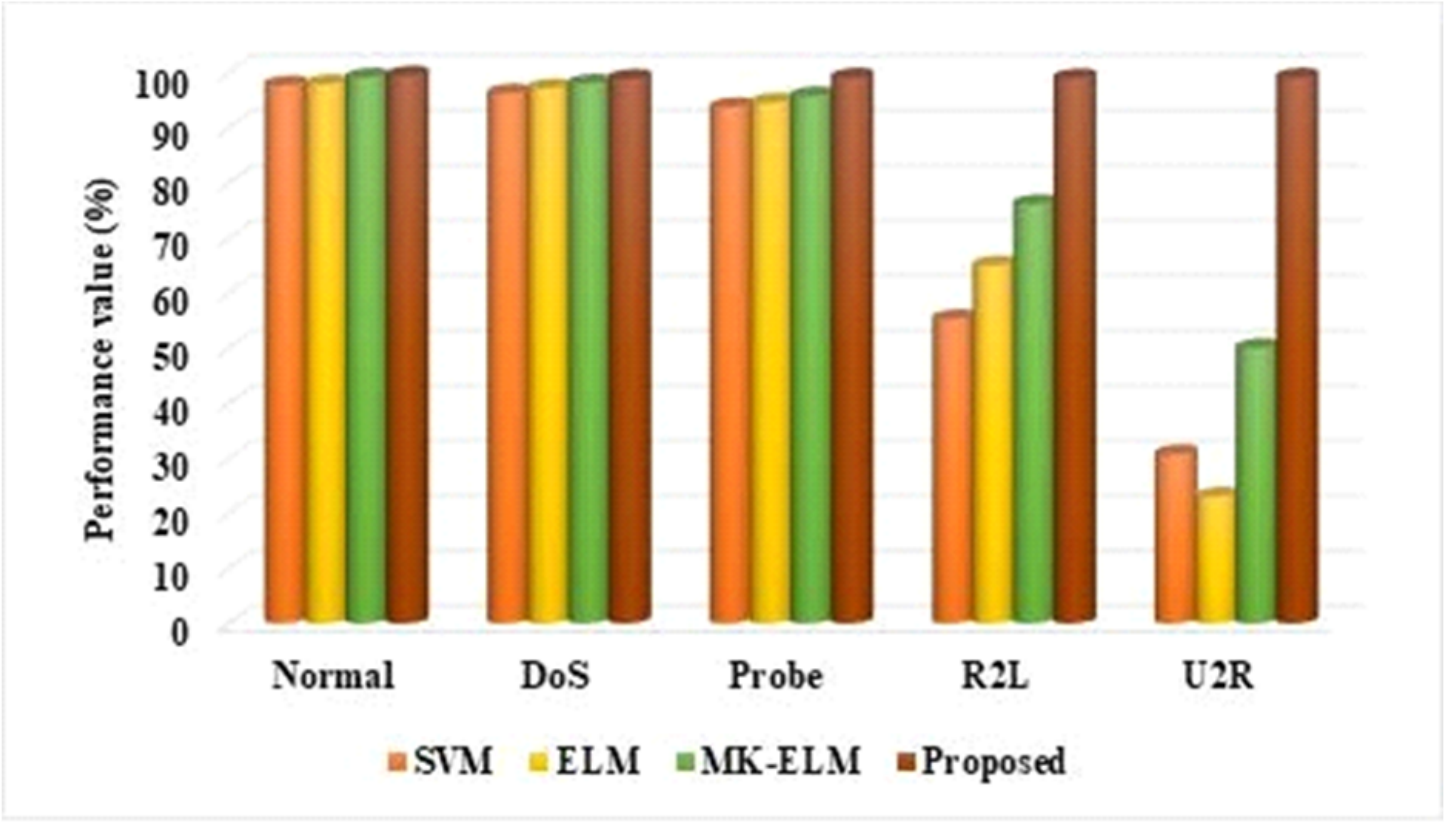


Figure 9

Detection rate Vs different types of attacks

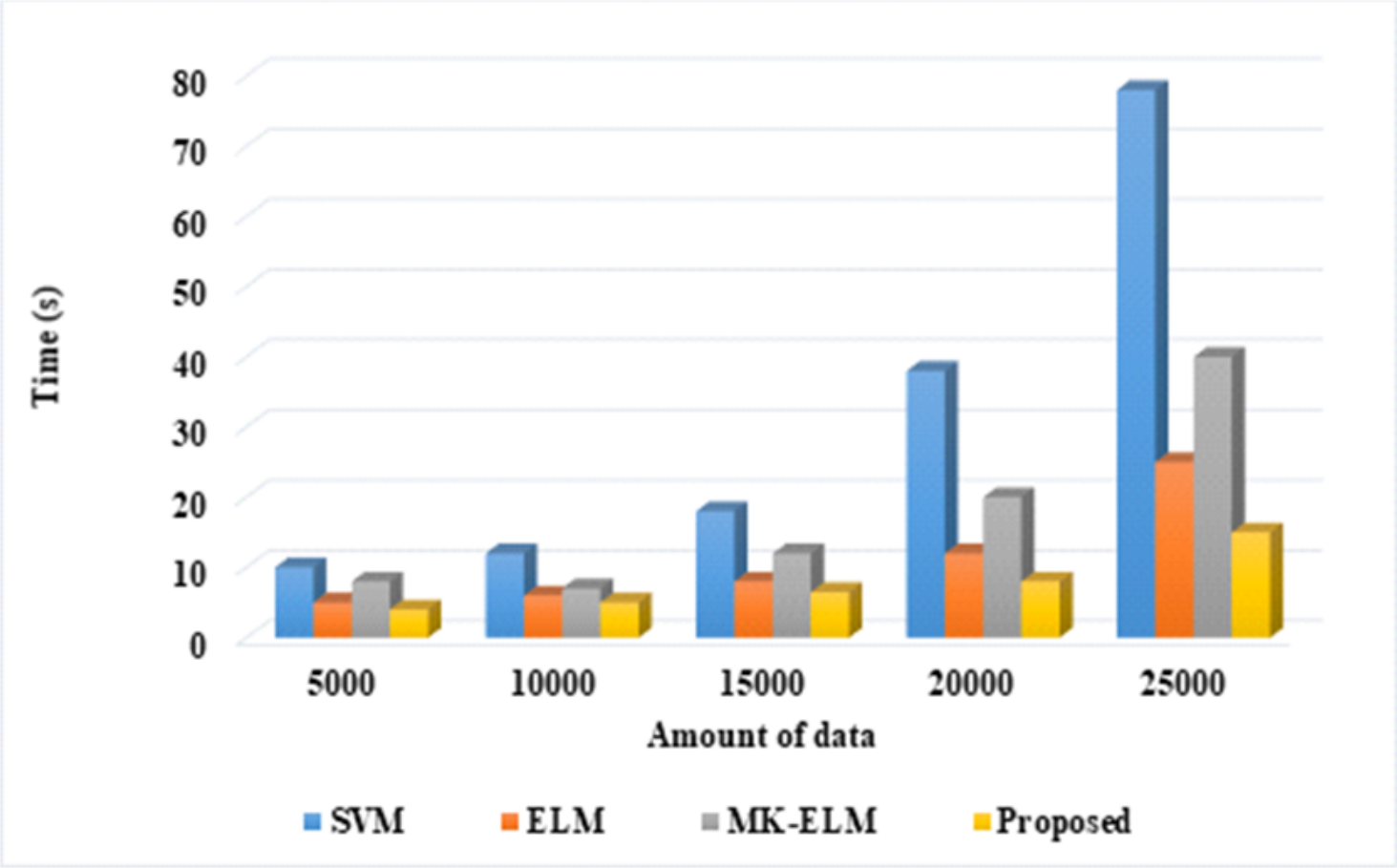
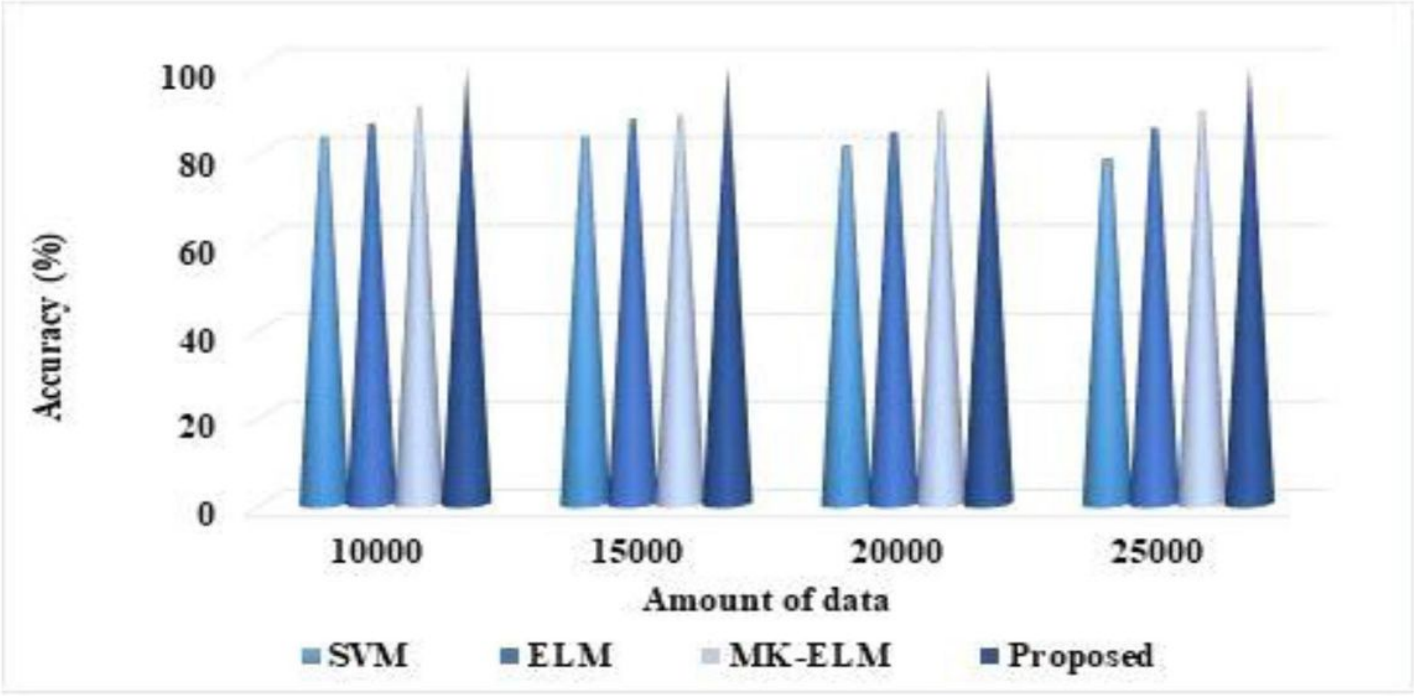


Figure 10

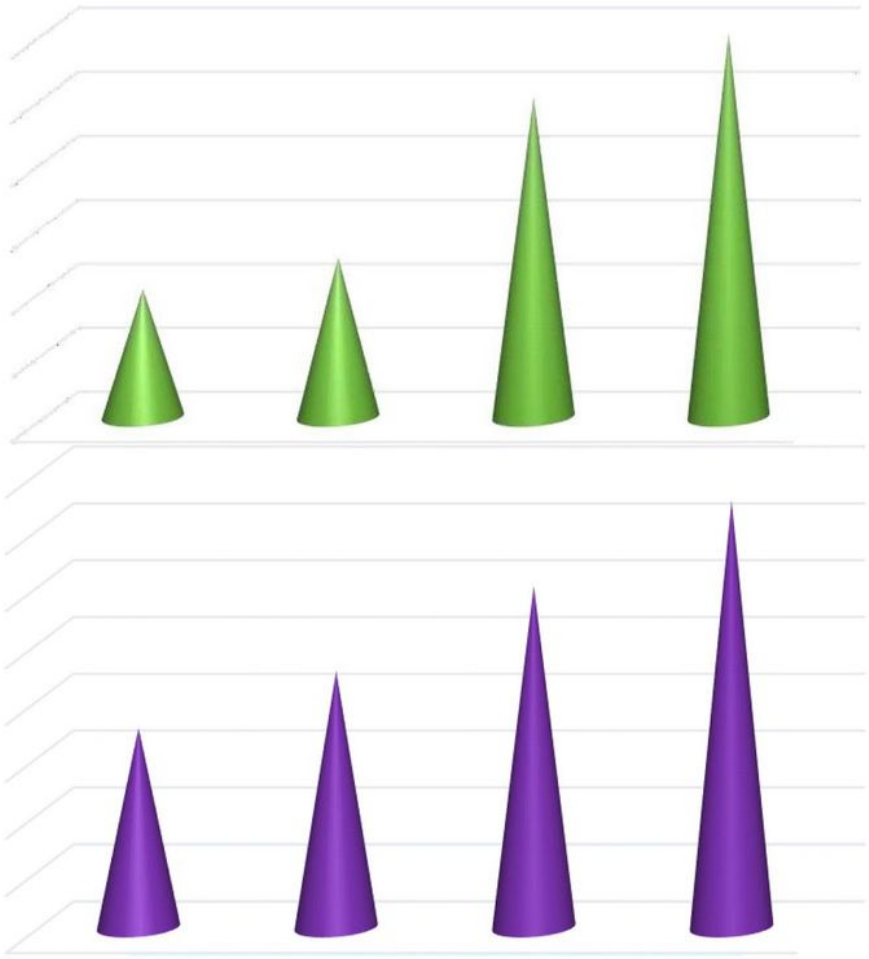
Time consumption (s)



0.12				
0.1				
0.08				
0.06				
0.04				
0.02				
0	10000	15000	20000	25000

Figure 11

Accuracy Vs amount of data



1.6				
1.4				
1.2				
1				
0.8				
0.6				
0.4				
0.2				
0	10000	15000	20000	25000

Figure 12

Execution time Vs amount of data