UTM UNIVERSITI TEKNOLOGI MALAYSIA
INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

Performance Evaluation of Support Vector Machine Kernels in Intrusion Detection System for Wireless Sensor Network

Muhammad Amir Hamzah & Siti Hajar Othman
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia
81310 Johor Bahru, Malaysia
Email: muhammad.amir@utm.graduate.my; hajar@utm.my

Submitted: 22/8/2021. Revised edition: 26/9/2021. Accepted: 26/10/2021. Published online: 16/5/2022
DOI: https://doi.org/10.11113/ijic.v12n1.334

***Abstract*—Wireless sensor network is very popular in the industrial application due to its characteristics of infrastructure-less wireless network and self-configured for physical and environmental conditions monitoring. However, the dynamic environments of wireless network expose WSN to network vulnerabilities. Intrusion Detection System (IDS) has been used to mitigate the vulnerability issue of network. Researches towards the efficiency improvement of WSN-IDS has been extensively done because the rapid growth of technologies influence the growth of network attacks. Implementation Support Vector Machine (SVM) was found to be one of the optimum algorithms for the improvement of WSN-IDS. Yet, classification efficiency of SVM is based on the kernel function used because different kernel gives different SVM architecture. Linear classification of SVM has limitation to maximize the margin due to the dynamic environment of wireless network which consist of nonlinear data. Since maximizing the margin is the primary goal of SVM, it is crucial to implement the optimum kernel in the classification of nonlinear data. Each SVM model in this research use different kernels which are Linear, RBF, Polynomial and Sigmoid kernels. Further, NSL-KDD dataset was used for the experiment of this research. Performance of each kernel were evaluated based on the experimental result obtained and it was found that RBF kernel provides the best classification accuracy with the score of 91%. Finally, discussion based on the findings was made.**

***Keywords*—WSN-IDS, SVM, Non-Linear, Kernel Function**

I. INTRODUCTION

Wireless Sensor Network (WSN) has been widely used in many industries such as agriculture, health care and vehicle cloud. WSN is formed with a combination of two components which are sinks and sensor nodes. In nature, the main characteristic of wireless network is self-organized as it serves the dynamic environment of wireless network which makes it highly in demand. However, the characteristics stated the cause the network environment of WSN become more complex which expose it to the vulnerabilities of network. Thus, security mechanism such as Intrusion Detection System (IDS) has been extensively used to mitigate this issue. IDS functions as a tool to identify and react to any harmful and intrusive activities occur within the system's facilities [1]. Intrusion can be detected based on the two modules: Signature Based and Anomaly Based [2].

Signature-based approach detect intrusion by examining the pattern and match it with the signature that is stored in the existing database of the system [3]. Anomaly based is a detection approach of IDS that detects intrusion based on the created profile of the system. The profile is created to define the normal traffic pattern and the detection is done by matching the pattern of the traffic. If the pattern is different from the profile, hence intrusion of malicious packet is detected. Development of IDS has gone through numerous developments and many improvements have been proposed for the detection of IDS.

According to [4], the concept of Machine Learning is important for the application of WSN due to the rapid changes of all network sensor which is also used in monitoring the dynamic environment. In addition, the complicated environment in the development of WSN needs complex algorithm to solve simple mathematical model and cater the

issue of improper WSN operation in collecting new information due to unexpected errors. Machine Learning can be categorized into Supervised Learning and Unsupervised Learning. Both algorithms can also be combined which makes it a hybrid algorithm of Supervised and Unsupervised Learning. Unsupervised learning of Machine Learning use clustering technique in finding group of hidden data. Supervised learning of machine learning use two technique which are classification and regression to produce an output based on the sample of input.

In [5], the research made a comparison study to different the performances of several Supervised Machine Learning algorithm in finding anomaly based on NSL-KDD dataset. [6] also made a research in developing a Supervised Machine Learning IDS using Support Vector Machine algorithm where performance of developed model was compared with the performances of Decision Tree, Logistic Regression, Naïve Bayes and Artificial Neural Network algorithm. Based on the research stated, Support Vector Machine (SVM) algorithm was found to have the highest performance result in detecting anomaly.

II. Support Vector Machine

Support Vector Machine (SVM) is a detection method that learns data for the identification of the pattern. Classification and regression of SVM are flexible with diverse binary classification complexity through the hyperplane construction for the representation of the boundary in the middle of two classes [7]. This approach can learn a huge pattern of datasets as well as having the better scaling as the classification of SVM are done without the dependencies of the features' dimensions.

Vapnik has introduced SVM as a model for Machine Learning that use kernel in performing classification and regression task [8]. [9] made reviews on several experiments done on SVM in the year 2009 where it shows that SVM algorithm is affected severely when imbalanced data sets is applied. However, the performance of SVM can be improved by choosing the suitable parameter [10]. One of the parameters that is significant for SVM is the kernel used. Kernel in SVM algorithm defines the architecture of the algorithm.

Classification in SVM is done by the constructing a hyperplane. Multiple numbers of hyperplane also can be used in a process of classification. However, the effectiveness of classification in SVM is unlikely based on the number of hyperplanes used, yet it is based on the maximum margin that a hyperplane can produce between two classes of data point [11]. There are two types of data that use SVM for classification, linearly separable data and non-linearly separable data. For non-linearly separable data, the data need to be presented in a form of high dimensional space and maximum margin of the hyperplane can be applied which can be achieved by implementing the kernel function [12]. In addition, the selection of the kernel used needs to be properly done as the different kernel use will construct a different architecture of SVM which will affect the performance and capability of the SVM.

III. Previous Work

Numerous research have been done in implementing SVM in the development of WSN-IDS. The implementation of machine learning for a specific type of Intrusion Detection System (IDS) such as WSN-IDS or N-IDS are similar because it generally uses the same technique, dataset and evaluation process. The only different is how it will be implemented in the real world. There are four types of kernel that are usually used for SVM. The kernels are Linear, Polynomial, Radial Basis Function (RBF) and Sigmoid. According to [13] Linear is the most simple kernel for SVM as it needs a small requirement for computation and it is unparameterized, Polynomial is a kernel for SVM that is preferably for problems that use normalized data which parameterized based on the degree, RBF is the most preferred kernel as it can give high scores of classification and Sigmoid is the least kernel preferred as it is expected to have the least score in classification accuracy which parameterized based on the curve.

[14] has proposed a SVM classifier model for IDS that implements Rough Set Theory for feature selection. The main goal of feature selection is to find a subset in a dataset which consists of only relevant attribute that can describe the dataset entirely. Rough Set Theory works by identifying unimportant feature in a dataset to form a reduct set and the set will be reducted once the identification has been done completely. In this research, it was found that only 15 features are important from NSL-KDD dataset after it undergoes the feature selection process of Rough Set Theory. Three SV kernel were used, RBF, Polynomial and Sigmoid. Based on the result, it was found that RBF achieve the highest accuracy with 97.76% of score.

[15] has evaluated kernels of SVM models for IDS where each classifier model used different kernel. This research uses three different kernels which were Linear, RBF, and Polynomial. In addition, Grid search technique was also used to find the optimum parameter for each SVM classifier model. Further, two types of datasets were used and each dataset uses different tuning on SVM classifier model. According to the researcher, the configuration was made to increase the accuracy of classification. The mentioned datasets are RRE-KDD and NSL KDD. RRE-KDD is a dataset that consist of KDD99 Test+ and KDD99Train+ dataset for testing as well as training. Based on the findings, the researcher found that RBF kernel performs better on RRE-KDD dataset with the classification accuracy score of 92.99 % while Polynomial kernel performs better on NSL-KDD dataset with the classification accuracy score of 73.54%.

[16] has proposed new data scaling method for SVM model of IDS. Instead of using the MinMax normalization technique, two non-linear data scaling method were proposed. These scalers were used in the data preprocessing phase of the SVM classifier model for data representation. According to the researcher, MinMax normalization method has a problem where the value completely depends on the minimum and maximum value which will affect the accuracy of final detection during testing phase. Therefore, two non-linear scaling method has been proposed which are Logistic scaling

and Arctan scaling. Both scalers are data independent. In the research, four different kernels were employed in the SVM implementation. The kernels are Linear, RBF, Polynomial and Sigmoid. On the other hand, the dataset used for this research was raw NSL-KDD dataset. Thus, the dataset was preprocessed to well suit with the SVM classifier model. Each scaler was used in each SVM classifier model. Based on the findings, Polynomial kernel has the highest classification accuracy with 82% of score for both scalers.

TABLE I. RESEARCH DONE ON THE PERFORMANCE OF SVM KERNEL BASED ON CLASSIFICATION ACCURACY

Researcher	**Kernels Used**	**Kernel with Highest Accuracy**	**Dataset**	**Score (%)**
Reddy *et al.* (2015)	RBF, Polynomial and Sigmoid	RBF	NSL-KDD	97.76
Hasan *et al.* (2016)	Linear, Polynomial, and RBF	RBF	RRE-KDD	93.19
		Polynomial	NSL-KDD	73.54
Tang *et al.* (2018)	Linear, Polynomial, RBF and Sigmoid	Polynomial	NSL-KDD	82

Table I shows the comparison of several research in evaluating the performance of SVM kernels for WSN-IDS. Based on table above, Polynomial and RBF kernels have the highest classification accuracy score on different dataset and SVM classifier model. Therefore, different kernels has different classification efficiency depends on SVM classifier's configuration and dataset used.

IV. DATASET

Datasets has usually been used to test the performance of proposed IDS. Currently, NSL KDD is one of the most used datasets for testing the performance of IDS. NSL KDD is the extension of KDD Cup 99 datasets. This newer version of the mentioned dataset reduces the problems of duplication, redundancy and the distribution of target class is non-uniform [17]. In NSL KDD, there are four network attacks: Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R) and Probing. NSL-KDD has 41 features where 38 numeric feature and 3 non-numeric features [17]. Most of the research done focusing on the classification of N-IDS and W-IDS which include WSN-IDS and IDS for Manet used the same method and process flow as well as the dataset used which is NSL-KDD. Therefore, NSL-KDD is very suitable to test the accuracy of classification.

V. DESIGN AND IMPLEMENTATION

In IDS, several crucial processes need to be considered to have the optimal result of data classification. This research goal is to perform several SVM classifiers with different kernels.

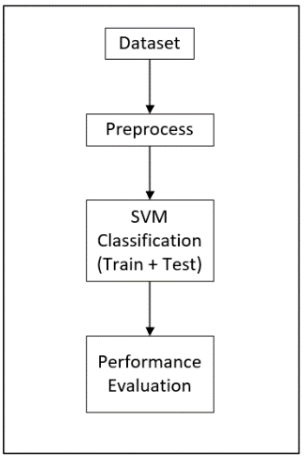

Fig. 1. Flow of SVM Process

A. NSL-KDD Dataset

Dataset used for this experiment is NSL-KDD. This dataset is available online and it usually in .crv format. This dataset is an improved version of KDD99 dataset where it removes all the redundant records. Moreover, records in NSL-KDD are more suitable for testing and training. The dataset used in this experiment has 41 features containing 125,973 samples for training and 22,544 samples for testing. The list of features is shown in Table II. Both samples consist of normal and attack samples. Nevertheless, only 14 of the features will be used in this experiment. The selection of the features will be discussed in this section.

TABLE II. FEATURES OF NSL-KDD

No.	**Feature**	**No.**	**Feature**	**No.**	**Feature**
1	Duration	15	Su_attemped	29	Same_srv_rate
2	Protocol_type	16	Num_root	30	Diff_srv_rate
3	service	17	Num_file_creations	31	Srv_diff_host_count
4	flag	18	Num_shells	32	Dst_host_count
5	Src_bytes	19	Num_access_files	33	Dst_host_srv_count
6	Dst_bytes	20	Num_outbound_cmds	34	Dst_host_same_srv_rate
7	Lang	21	Is_host_login	35	Dst_host_diff_srv_rate
8	Wrong_fragment	22	Is_guest_login	36	Dst_host_same_src_port_rate
9	Urgent	23	Count	37	Dst_host_srv_diff_host_rate
10	Hot	24	Srv_count	38	Dst_host_serror_rate
11	Num_failed_logins	25	Serror_rate	39	Dst_host_srv_serror_rate
12	Logged_in	26	Srv_serror_rate	40	Dst_host_rerror_rate
13	Num_compromised	27	Srv_rerror_rate	41	Dst_host_srv_rerror_rate
14	Root_shell	28	Rerror_Rate		

B. Preprocess

Data preprocessing is the initial part of this experiment and it involves two processes. The first process is converting the character values of data samples in the NSL-KDD dataset into numerical values. Further, data is preprocessed using power transformation method. In power transformation, data is converted into 0 for mean and 1 for standard deviation. The formula is shown in [1] where x_i represent the data point, $\bar{x}$ represent the mean and s represent the standard deviation. λ ≠0 is a simple power transformation where $y^{\lambda}$is rescaled for the lamda, λ in $h(y;\lambda)$ will be converged to 0.

$$h(y;\lambda) = \frac{(y^{\lambda}-1)}{\lambda} \qquad \lambda \neq 0$$
$$= \log(y) \qquad \lambda \neq 0 \qquad [1]$$

C. SVM Classification

SVM classifier is used to classify the selected dataset. Four SVM classifiers is performed and each of the classifier implement different kernels. The first SVM classifier used linear as the kernel. Linear kernel is the simplest kernel. The Mathematical formula of Linear kernel is shown in equation [2]. The value of (x,y) will determine the slope of the hyperplane and the of C will determine the maximum margin in analyzing data.

$$\mathrm{K(x,y)} = \mathrm{(x,y)} + \mathrm{C} \qquad [2]$$

The second SVM classifier used polynomial as the kernel. Polynomial is suitable for normalized data (Drewnik & Pasternak-Winiarski, 2017). Equation [3] shows the Mathematical formula of polynomial kernel. In the equation, γ (x,y) determines the shape of the hyperplane, and C determines the maximum margin. The parameter of this kernel is determined by the degree, d which represents the curve of the hyperplane.

$$\mathrm{K(x,y)} = (\gamma\,\mathrm{(x,y)} + \mathrm{C})^{d} \qquad [3]$$

The third SVM classifier used RBF as the kernel. This kernel is the most common kernel of SVM as it can provide high classification results. Mathematical formula of RBF is represented in equation [4].

$$\mathrm{K(x,y)} = \exp(-\gamma\|x-y\|^2) \qquad [4]$$

The last SVM classifier used Sigmoid as the kernel. This kernel has the similar function of neural network, but this kernel has the least classification among the other kernels. The slope is determined by the value (x,y) . The parameter of this kernel is determined by the value of γ which will determine the value of hyperplane and exp which is the C will determine the maximum margin of each hyperplane. The Mathematical formula is formulated based on the following equation.

$$\mathrm{K(x,y)} = \tanh(\gamma\,\mathrm{(x,y)} + \mathrm{C}) \qquad [5]$$

D. Performance Evaluation

The performance of each SVM classifier in classifying the subset is evaluated in this phase. The evaluation is based on the classification Accuracy, Recall, Precision and F-measure. The formula for evaluation of each measurement is shown in the table below. Accuracy is the overall ratio score that the prediction accurateness, Recall is the sensitivity, or the detection rate based on the correct prediction, Precision is the ratio score that prediction is correctly made and F-measure is a harmonic score of recall score with precision score to determine the classification efficiency.

TABLE III. MEASUREMENT FORMULA FOR PERFORMANCE EVALUATION

Measure	Formula
Accuracy	$Accuracy = \frac{(TP+TN)}{(TP+FN+FP+TN)}$
Recall	$Recall = \frac{TP}{(TP+FN)}$
Precision	$Precision = \frac{TP}{(TP+FP)}$
F-measure	$Fmeasure = \frac{(2 \times Recall)}{(Precision+Recall)}$

To obtain the optimal result. Each of the value gain from each measurements are resamples using cross validation. This technique is a process to evaluate machine learning performance on a small subset. In this research, 10-fold cross validation is used in achieving the optimal result. The formula is shown below. The value is split into division of fold and prediction is fit on all points as well as evaluating error on points in each fold [18].

$$E = \frac{1}{n}\sum_{i=1}^{n}(E_i) \qquad [6]$$

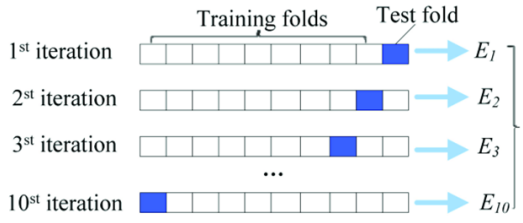

Fig. 2. 10-Fold Cross-Validation [19]

VI. EXPERIMENTAL SETUP

The experiment tools in conducting this research will be discussed. In conducting Machine Learning (ML) classifiers, several platforms coding languages such as Java, C, and Matlab can be used. However, the chosen language in

performing this experiment is Python that operates through Jupyter Notebook platform running in Windows 10 with 4GB of RAM. The selection of this tool is because the code language is much simpler compared to other coding language and it has various library such as Scikit-learn, Pandas, and NumPy that provides the algorithm for machine learning, data manipulation as well as data analyzing. In addition, many of the open-source code for machine learning available online are in Python language. This helps the model development for conducting this experiment become easier.

VII. Experimental Results and Analysis

The result of each classifier is projected using confusion matrix. Based on the result projected, the value of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) can be determined.

TABLE IV. Indication values for confusion matrix

Class		Predicted Class	
		Positive	**Negative**
Actual Class	Positive	TP	FN
	Negative	FP	TN

These values are crucially needed for performance evaluation that will be discussed in the next section. Tables below shows the results of SVM classifiers in a form of confusion matrix.

TABLE V. Confusion Matrix for SVM classifier with Linear Kernel

Class		Predicted Class	
		Positive	**Negative**
Actual Class	Positive	8927	784
	Negative	5523	7310

TABLE VI. Confusion Matrix for SVM classifier with Polynomial Kernel

Class		Predicted Class	
		Positive	**Negative**
Actual Class	Positive	8989	722
	Negative	5873	6960

TABLE VII. Confusion Matrix for SVM classifier with RBF Kernel

Class		Predicted Class	
		Positive	**Negative**
Actual Class	Positive	8940	771
	Negative	5788	7045

TABLE VIII. Confusion Matrix for SVM classifier with Sigmoid Kernel

Class		Predicted Class	
		Positive	**Negative**
Actual Class	Positive	7766	1945
	Negative	4196	8637

Performance evaluation based on Classification Accuracy, Recall, Precision and F-measure. The evaluation is calculated using the value of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). For obtaining optimal results, these values are calculated using 10-Fold Cross Validation technique. The values obtained through the validation process then are calculated using the equation mentioned in Section V to have the results of evaluation which are the Classification Accuracy, Recall, Precision and F-measure. Below are the figures projecting the result.

TABLE IX. Accuracy Score for SVM kernels

Kernel	Linear	Polynomial	RBF	Sigmoid
Accuracy	88%	90%	91%	78%

Based on the table above, it shows that RBF kernel has the highest accuracy score compared to the other three kernels. The second highest score achieves by Polynomial kernel and followed by Linear kernel. Sigmoid kernel has the lowest accuracy score compared to RBF, Polynomial and Linear kernels. Thus, RBF kernel achieves the most accurate result in classifying nonlinear data.

TABLE X. Recall Score for SVM kernels

Kernel	Linear	Polynomial	RBF	Sigmoid
Recall	91%	91%	91%	83%

Based on the table above, Linear, Polynomial and RBF kernels achieves the same score which is 91%. On the other hand, Sigmoid kernel has the lowest recall score. Thus, Sigmoid kernel has the least sensitivity or the detection rate based on the correct prediction compared to the other three kernels.

TABLE XI. Precision Score for SVM kernels

Kernel	Linear	Polynomial	RBF	Sigmoid
Precision	89%	91%	92%	78%

Based on the table above, it shows that RBF kernel has the highest Precision score compared to Linear, Polynomial and Sigmoid kernels. Sigmoid kernel achieves the lowest Precision among the other three kernels. Hence, RBF kernel has the highest prediction ratio in making a correct prediction towards the classification of nonlinear data.

TABLE XII. F-measure Score for SVM kernels

Kernel	Linear	Polynomial	RBF	Sigmoid
F-measure	90%	91%	92%	81%

The higher the score the better the performance of the kernel. A bar chart presenting the Accuracy, Precision, Recall

and F-measure scores obtained by each kernel was made to give a better view in comparing their performances in classifying nonlinear data.

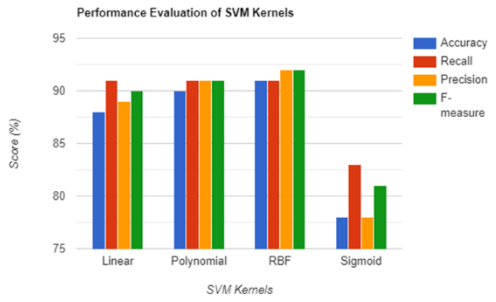

Fig. 3. Performance Evaluation of SVM Kernels

Based on observations made towards the performance evaluation and figure above, RBF kernel achieves the most efficiency in performing classification towards nonlinear data. More important findings are listed below. Further, Polynomial kernel achieves the second-best kernel in performing nonlinear data classification. Moreover, Linear kernel also performs well which makes it the third best kernel to perform classification of nonlinear data. Sigmoid kernel has the least Accuracy, Precision, Recall and F-measure scores which indicates that this kernel is inefficient in performing nonlinear data classification compared to the other three kernels.

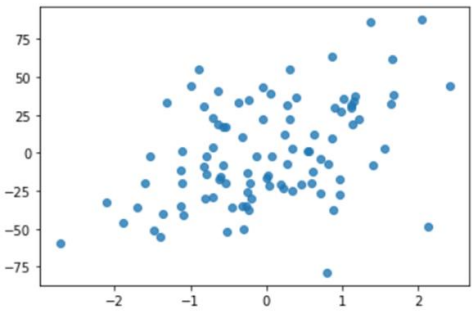

Fig. 4. Data Sample Demonstration

Figure above is the illustration of how the data used in this research was mapped on the feature space the classification of SVM. The purpose of this figure was made to discuss about the reason behind the classification performance of each kernel. Based on the discussion made in section V regarding the architecture of SVM kernels, it can be visualized that the hyperplane and margin size produced by RBF kernel is very suitable to classify nonlinear data as shown in figure above. Thus, this reason makes the kernel is the most preferable kernel in classifying data. In addition, the architecture produces by Polynomial kernel is also suitable for the classification of nonlinear data. Polynomial kernel can also achieve the best results by tuning the kernel by using optimum parameter. Moreover, Linear kernel can perform better than Sigmoid kernel in classifying nonlinear data but not as efficient as RBF and Polynomial kernel. However, the efficiency of Linear kernel could decrease to more scattered of nonlinear data.

The best result obtained from this research which is the Accuracy score of RBF kernel was compared with the best result of previous work for further validation. Table 5.10 shows the comparison of best accuracy score obtained by this research with previous research based on the classification of NSL-KDD dataset.

TABLE XIII. COMPARISON OF ACCURACY SCORE WITH PREVIOUS RESEARCH

Researcher	Kernel	Dataset	Score (%)
Hasan *et al*. (2016)	Polynomial	NSL-KDD	73.54
Tang *et al*. (2018)	Polynomial	NSL-KDD	82
This research	**RBF**	**NSL-KDD**	**91**
Reddy *et al*. (2015)	RBF	NSL-KDD	97.76

Based on the table above, Hasan *et al*. (2016) and Tang *et al*. (2018) found out Polynomial kernel achieves the best accuracy score in classifying NSL-KDD dataset. Although the method proposed by Tang *et al*. (2018) used two types of nonlinear scaler for data scaling as discussed in section III, data scaled by the Logistic scaling and Arctan scaling are suitable with Polynomial kernel. Table XIII shows that Accuracy score obtained by RBF kernel through the research made by Reddy *et al*. (2015) achieve the highest although the research implemented the linear data scaling method which is the MinMax scaler. Yet, the SVM model in the research includes an additional process which was the feature selection process. Therefore, it was proven that feature selection can improve the classification accuracy of SVM.

VIII. CONCLUSION

The efficiency of SVM classifier is based on parameter used for the classification such as the kernel use, size of margin and data presented. In the SVM, it is important to standardize data before it can be presented in the feature space. Standardization of dataset helps to avoid any errors during the process of data classification in SVM. In addition, the efficiency of SVM classifier is influenced by the kernel function used for the classification. Nevertheless, the kernel used must be accordance to how the dataset is presented. Linear kernel can perform very well in data classification. However, it has some limitation to maximize its margin when nonlinear data is presented in the feature space. Based on the result obtained in this research, it is found that RBF is the best kernel among the other three kernels (Polynomial, Sigmoid and Linear) for classification of nonlinear which can enhance detection of attack in WSN-IDS

There is a limitation occur in conducting this research even the objective of this research has been achieved. The limitation is listed below.

i) Parameter used in the experiment such as estimator size of margin in SVM limited to minimum value as huge value can cause high computational as well as increasing the time of processing.
ii) This research focus on the efficiency of SVM kernels in classifying data without the intervention of feature selection.

In the future, the result of this experiment will be analyzed and compared between the results obtained from this work with the existing research will be done. By conducting the analysis, result obtained from this research can be validated. In addition, this would also help to find more research gap. Based on section 6.2, the limitations stated are of the research gaps found throughout this research. Below are the improvements that can be made to this research through future work.

i) Use a greater platform or computer machine in conducting experiment to achieve shorter time of execution.
ii) Includes the process of feature selection to improve the classification efficiency and increase the accuracy.
iii) Use optimum parameter that suits with the data presented on each kernel to achieve greater classification results.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Universiti Teknologi Malaysia (UTM) for funding this research under the Tier 1 Research (Vote Q.J130000.2528.14H82.

REFERENCES

[1] W. A. H. M. Ghanem and A. Jantan. (2018). Hybridizing Artificial Bee Colony with Monarch Butterfly Optimization for Numerical Optimization Problems. *Neural Comput. Appl.*, 30(1), 163-181. Doi: 10.1007/s00521-016-2665-1.
[2] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita. (2014). Towards an Unsupervised Method for Network Anomaly Detection in Large Datasets. *Comput. Informatics*, 33(1), 1-34.
[3] A. N. Cahyo, R. Hidayat, and D. Adhipta. (2016). Performance Comparison of Intrusion Detection System based Anomaly Detection Using Artificial Neural Network and support Vector Machine. AIP Conference Proceedings, 1755(1), 70011.
[4] E. Baraneetharan. (2020). Role of Machine Learning Algorithms Intrusion Detection in WSNs: A Survey.
[5] B. A. Ashwini and S. S. Manivannan. (2020). Supervised Machine Learning Classification Algorithmic Approach for Finding Anomaly Type of Intrusion Detection in Wireless Sensor Network. Opt. Mem. Neural Networks *(Information Opt.*, 29(3), 2443-256. Doi: 10.3103/S1060992X20030029.
[6] D. Jing and H. Chen. (2019). SVM Based Network Intrusion Detection for the UNSW-NB15 Dataset. *2019 IEEE 13th International Conference on ASIC (ASICON)*, 1-4. Doi: 10.1109/ASICON47005.2019.8983598.
[7] A. Majid, S. Ali, M. Iqbal, and N. Kausar. (2014). Prediction of Human Breast and Colon Cancers from Imbalanced Data using Nearest Neighbor and Support Vector Machines. *Comput. Methods Programs Biomed.*, 113(3), 792-808. Doi: https://doi.org/10.1016/j.cmpb.2014.01.001.
[8] V. Vapnik. (1998). The Support Vector Method of Function Estimation. Nonlinear Modeling, Springer, 55-85.
[9] J. Cervantes, F. García-Lamont, L. Rodríguez, and A. Lopez-Chau. (2020). A Comprehensive Survey on support Vector Machine Classification: Applications, Challenges and Trends. *Neurocomputing*, 408. Doi: 10.1016/j.neucom.2019.10.118.
[10] D. R. Amancio *et al.* (2014). A Systematic Comparison of Supervised Classifiers. *PLoS One*, 9(4), e94137.
[11] A. Abd Manaf, S. Sahibuddin, R. Ahmad, S. M. Daud, and E. El-Qawasmeh. (2011). Informatics Engineering and Information Science. *Conference Proceedings ICIEIS*, 42.
[12] G. Kumar, K. Kumar, and M. Sachdeva. (2010). The Use of Artificial Intelligence based Techniques for Intrusion Detection: A Review. *Artif. Intell. Rev.*, 34(4), 369-387. Doi: 10.1007/s10462-010-9179-5.
[13] M. Drewnik and Z. Pasternak-Winiarski. (2017). SVM Kernel Configuration and Optimization for the Handwritten Digit Recognition.
[14] R. R. Reddy, Y. Ramadevi, and K. V. N. Sunitha. (2015). Anomaly Detection using Feature Selection and SVM Kernel Trick. *Int. J. Comput. Appl.*, 975, 8887.
[15] M. A. M. Hasan, S. Xu, M. M. J. Kabir, and S. Ahmad. (2016). Performance Evaluation of Different Kernels for Support Vector Machine used in Intrusion Detection System. *Int. J. Comput. Networks Commun.*, 8(6), 39-54.
[16] X. Tang, S. X.-. Tan, and H. Chen. (2018). SVM Based Intrusion Detection Using Nonlinear Scaling Scheme. *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), Oct. 2018*, 1-4. Doi: 10.1109/ICSICT.2018.8565736.
[17] M. Islabudeen and K. D. M. K. (2020). A Smart Approach for Intrusion Detection and Prevention System in Mobile Ad Hoc Networks Against Security Attacks. *Wirel. Pers. Commun.*, 112. Doi: 10.1007/s11277-019-07022-5.
[18] M. Panda, A. Abraham, and M. R. Patra. (2012). A Hybrid Intelligent Approach for Network Intrusion Detection. *Procedia Eng.* 30, 1-9z. Doi: https://doi.org/10.1016/j.proeng.2012.01.827.
[19] M. Niu, Y. Li, C. Wang, and K. Han. (2018). RFAmyloid: A Web Server for Predicting Amyloid Proteins. *Int. J. Mol. Sci.*, 19. Doi: 10.3390/ijms19072071.