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

A Review of Support Vector Machine-based Intrusion Detection System for Wireless Sensor Network with Different Kernel Functions

Muhammad Amir Hamzah, Siti Hajar Othman
School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
Email: hajar@utm.my.com

Submitted: 9/3/2021. Revised edition: 19/4/2021. Accepted: 20/4/2021. Published online: 24/05/2021
DOI: https://doi.org/10.11113/ijic.v11n1.303

***Abstract*—Wireless sensor network (WSN) is among the popular communication technology which capable of self-configured and infrastructure-less wireless networks to monitor physical or environmental conditions. WSN also is the most standard services employed in commercial and industrial applications, because of its technical development in a processor, communication, and low-power usage of embedded computing devices. However, WSN is vulnerable due to the dynamic nature of wireless network. One of the best solutions to mitigate the risk is implementing Intrusion Detection System (IDS) to the network. Numerous researches were done to improve the efficiency of WSN-IDS because attacks in networks has been evolved due to the rapid growth of technology. Support Vector Machine (SVM) is one of the best algorithms for the enhancement of WSN-IDS. Nevertheless, the efficiency of classification in SVM is based on the kernel function used. Since dynamic environment of WSN consist of nonlinear data, linear classification of SVM has limitations in maximizing its margin during the classification. It is important to have the best kernel in classifying nonlinear data as the main goal of SVM to maximize the margin in the feature space during classification. In this research, kernel function of SVM such as Linear, RBF, Polynomial and Sigmoid were used separately in data classification. In addition, a modified version of KDD'99, NSL-KDD was used for the experiment of this research. Performance evaluation was made based on the experimental result obtained. Finally, this research found out that RBF kernel provides the best classification result with 91% accuracy.**

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

I. INTRODUCTION

Wireless Sensor Network (WSN) has been used in many applications such as health care, vehicle cloud and agriculture. WSN is formed with a combination of two components which are sinks and sensor nodes. Due to the exposure of vulnerabilities towards WSN cause by its dynamic environment, wireless Sensor Network Intrusion Detection System (WSN-IDS) was developed to detect intrusion that could lead to network attack and send the alarm for any intrusion signals. Signature-based detection of IDS have a better performance in terms of detecting known abnormalities but it was found inefficient when it comes to new patterns of abnormalities (Rustam and Zahras, 2018). Therefore, anomaly-based detection of IDS is preferred as it identifies changes in the pattern of the system behavior (Kwon *et al*., 2019). 0According to (Ahmad *et al*., 2018), people are mostly implemented soft computing, machine learning and data mining to detect the attitude of intruders. This improvement is influenced by the problem that occurs when dealing with a huge amount of data (Reddy *et al*., 2016).

Support Vector Machine (SVM) is one of the machines learning methods that can provide high efficiency in classification (Rustam and Zahras, 2018). SVM has been widely used to propose a possible solution for the problem of complexity in IDS (Al Mehedi Hasan *et al*., 2013). However, the performance of SVM depends on the kernel used, parameters and data representation. Since data traffic in the real world is nonlinear (Rustam and Zahras, 2018),

(Farnaaz and Jabbar, 2016). According to Mathew *et al*. (2017), SVM work efficiently with problems that are nonlinear based on the kernel function. Different kernel function will have a different SVM architecture for data classification. In SVM, the main goal of its classifiers is to have a hyperplane that can maximize its margin between data point (Kausar *et al*., 2011). This goal can be achieved by using the kernel function in SVM (Mitrokotsa and Douligeris, 2005). This research is aimed to obtain the best SVM kernel function in the classification of nonlinear data for Wireless Sensor Network Intrusion Detection System (WSN-IDS). By implementing the best kernel function, enhancement towards classification accuracy of SVM classifiers which also influenced the performance of SVM based WSN-IDS can be achieved. In the following sections of this paper, literature review, methodology, implementation, and experimental result of this research will be discussed.

II. Literature Review

Intrusion detection is a process that monitors and analyze computer networks or system against intrusion (Liao *et al*., 2013). IDSs were first developed by Anderson in 1980 (Anderson, 1980) and improved by Denning a few years later (Denning, 1987). IDS functions as a tool to identify and react to any harmful and intrusive activities that occur within the system's facilities (Ghanem and Jantan, 2018). Many approaches have been created in the development of IDS such as implementing Machine Learning approaches to enhance the effectiveness of classification. Intrusion Detection System can be categories into three classes: Network-based (NIDS), Host-Based (HIDS) and Wireless-Based (WIDS). NIDS is a system to monitor and analyze the traffic of a network to identify malicious activity or attack. HIDS monitor and analyses the entire part of a computer system to identify intruders. HIDS can analyse the incoming network packet because of the ability of a computer system to establish connections with networks. WIDS is a system like NIDS but it is specifically designed for wireless networks.

A. Wireless Sensor Network Intrusion Detection System (WSN-IDS)

A wireless sensor network (WSN) is a wireless network that composes of the participation of sensor nodes and a centralized point called a sink or base station. WSN often operates in sectors like health care, climate change, room monitoring, agriculture, military, and other surveillance application. It operates the system as a group that collaborates with other neighbouring nodes and transmit data to the sink (Boubiche *et al*., 2020). Nodes in WSN are independent which can self-organize and self-healing. However, the mentioned characteristic of nodes leads the network to the exposure of network attack (Alrajeh *et al*., 2013). One of the security mechanisms that protect WSN is the integration of IDS in the network system.

B. Detection Type of Wireless Sensor Network Intrusion Detection System

Intrusion can be detected based on the two modules: Signature-Based and Anomaly Based (Bhuyan *et al*., 2014). The signature-based approach detects intrusion by examining the pattern and match it with the signature stored in the existing database of the system (Cahyo *et al*., 2016). Signature Based has great accuracy in detection. However, this approach has difficulties in detecting new attack or intrusion because the signature of the new intrusion is absence in the database.

Anomaly-based is a detection approach of IDS that detects intrusion based on the created profile of the system (Cahyo *et al*., 2016). 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 packets is detected. Thus, the anomaly-based approach has the advantage of detecting both new and old intrusion or attack (Alrajeh *et al*., 2013). Machine learning algorithm has been used in the development of IDS to have an accurate model specifically for classification, prediction and clustering (Jadidoleslamy, 2011).

C. Implementation of Machine Learning in WSN-IDS

The development of IDS has gone through numerous developments and many improvements have been are using Machine Learning (Sharma and Gupta, 2015). The implementation of Machine Learning is to enhance the capability of IDS to have efficient as well as accurate detection. According to (Baraneetharan, 2020), the concept of Machine Learning is important to solve issues in the application of WSN due to the rapid changes, complexity, and dynamic environment of the network. Machine Learning can be categorized into Supervised Learning and Unsupervised Learning. Both algorithms can also be combined which makes it Semi-Supervised Learning. Fig. 2.1 shows the categories of Machine Learning.

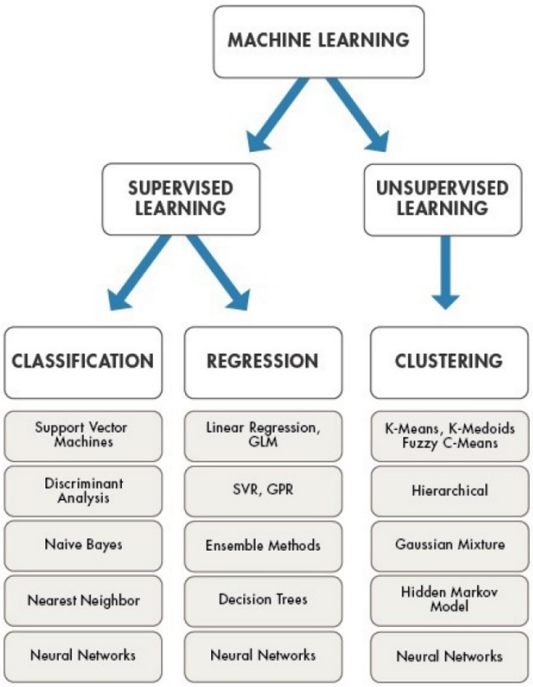

Fig. 2.1. Categories Of Machine Learning (Butun *et al*., 2013)

1) Unsupervised Learning in WSN-IDS

Unsupervised learning of Machine Learning uses the clustering technique in finding a group of hidden data. According to (Zanero and Savaresi, 2003), unsupervised learning is quite robust, and it can identify strange observation in a huge range of the event. In IDS, an unsupervised learning algorithm will learn the pattern of networks and any anomalies that occur can be detected. However, it has an issue that could have false positive alarms when detecting new patterns in a huge amount of data.

2) Supervised Learning in WSN-IDS

Supervised learning of machine learning use two techniques: classification and regression. These techniques are used to produce an output based on the sample of input. In Supervised Learning-based IDS, training is based on the training samples given. Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Naïve Bayes (NB), Artificial Neural Network (ANN) and Decision Tree (DT) are widely used for the enhancement of IDS. (Belavagi and Muniyal, 2016) has made a research using Support Vector Machine, Logistic Regression, Gaussian Naïve Bayes and Random Forest in detecting four types of attacks. (Ashwini and Manivannan, 2020) has made a research in comparing the performances of several Supervised Machine Learning algorithm in finding anomaly based on NSL-KDD dataset. (Jing and Chen, 2019) made a research in developing an IDS using Support Vector Machine algorithm where the performance of the developed model was compared with Decision Tree, Logistic Regression, Naïve Bayes and Artificial Neural Network algorithm. Based on the above-mentioned, it was found that the SVM algorithm has the highest performance result.

D. SVM Based WSN-IDS

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 (Chandra and Bedi, 2018). This approach can learn a huge pattern of datasets as well as having better scaling as the classification of SVM are done without the dependencies of the features' dimensions. In addition, SVM has a unique ability to update dynamically the pattern of training whenever a new pattern for the classification is found (Kuang *et al*., 2015). Vapnik has introduced SVM as a model for Machine Learning that uses kernels in performing classification and regression task (Vapnik, 1998). (Cervantes *et al*., 2020) 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 are applied. However, the performance of SVM can be improved by choosing a suitable architecture (Achirul Nanda *et al*., 2018). Kernel function in the SVM algorithm defines the architecture of the algorithm. Changing the kernel means the architecture of SVM is changed. According to (Liu *et al*., 2015), the kernel used in SVM is very crucial as it decides the performance of the algorithm classifier.

1) Kernels in SVM Based WSN-IDS

Classification in SVM is done by the hyperplane and multiple numbers of hyperplane can be used in a process of classification. However, the effectiveness of classification in SVM is unlike depending on the number of hyperplanes used, it is depending on the maximum margin that a hyperplane can produce between two classes of data point (Abd Manaf *et al*., 2011). 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 the maximum margin of the hyperplane can be applied which can be achieved by implementing the kernel function (Kumar *et al*., 2010). 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. This means that certain kernel is only effective for certain events. Following are the discussion regarding the types of kernel functions in SVM. The formula of these kernels will be discussed in the Methodology section (Section 3).

a) Linear Kernel

A linear kernel is the simplest kernel that provides the traditional classification method in SVM. Fig. 2.3 shows the visualization of how the linear kernel represents the data and how the classification of data is done.

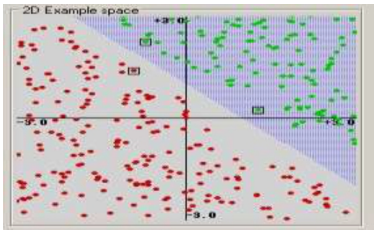

Fig. 2.3. Classification of Linear Kernel (H. Bhavsar & Panchal, 2012)

b) Radial Basis Kernel (RBF)

This kernel is the most common kernel of SVM as it can provide high classification results. Fig. 2.4 shows the visualization of the RBF kernel made by (H. Bhavsar and Panchal, 2012).

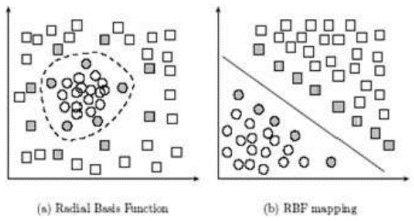

Fig. 2.4 Classification Of Rbf Kernel (H. Bhavsar And Panchal, 2012)

c) Polynomial

According to Drewnik and Pasternak-Winiarski (2017), Polynomial is suitable for normalized data. In this kernel, values of the degree will determine the flexibility of the margin used to classify data. Fig. 2.5 visualize the process of the polynomial kernel in SVM.

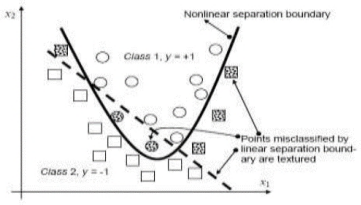

Fig. 2.5 Classification Of Polynomial Kernel (H. Bhavsar and Panchal, 2012)

d) Sigmoid Kernel

This kernel has a similar function to neural network, but this kernel has the least classification among the other kernels. Classification in this kernel is done by construction 2 decision boundaries. Data represented in between will be selected. Fig. 2.6 show the visualization of the Sigmoid kernel.

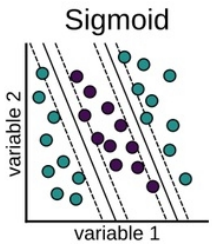

Fig. 2.6. Classification of Sigmoid Kernel

e) Limitation of SVM Classifiers

In SVM, the main limitation is to choose the best kernel for the given problem (Reddy *et al*., 2015). Data may come in many forms. However, most of the data that exists in the real world are in nonlinear form. Traditional SVM classifiers are very efficient in classifying linear or plain data (Kotpalliwar and Wajgi, 2015). However, the linear classifier of SVM may face some difficulties to classify nonlinear data due to the representation (Mathew *et al*., 2017). Nonlinear data is represented in form of high dimensional in the feature space which makes the linear classifier have difficulties to maximize its margin for decision boundary of data points. This goal can be achieved by using the kernel function in SVM (Mitrokotsa and Douligeris, 2005). According to (Al Mehedi Hasan *et al*., 2013), many of the researchers use Radial Bias Function (RBF) for this.

2) Comparative Studies of Kernel Functions in SVM for WSN-IDS

Bhavsar and Waghmare (2013) evaluated several SVM kernel functions: Polynomial, RBF and Sigmoid using NSL-KDD dataset. Based on the result, RBF was found to be the most accurate classification accuracy with 98.57%. Hasan *et al*. (2016) evaluate Linear, Polynomial, RBF and Laplace kernel functions by using RRE-KDD dataset. The research found out that RBF has the highest accuracy for the first dataset. In Tang *et al*. (2018), Linear, Polynomial, RBF and Sigmoid kernel functions were used to classify nonlinear data. On average, the research found out that the best kernel is Polynomial, and the least is Sigmoid. On average, the finding of these mentioned researches shows that RBF has the highest accuracy result compared to the other kernels. As mentioned, the effectiveness of SVM classifiers depends on the kernel used. The selection of kernel is based on the input data into the classification area of SVM. More organized data input will increase the effectiveness of kernels in SVM. However, organizing data will be difficult as it usually

comes in a huge amount which contains redundancy of data that can lead to faulty classification.

E. Data Preprocessing Using Nonlinear Feature Scaler

Data preprocessing is a very crucial process for Machine Learning as it standardizes data to avoid any prediction error. The prediction error is caused by the nature of Machine Learning in distinguishing data based on the value without properly identify it. Usually, there are two primary methods for feature scaling which are standardization and normalization. Standardization is a process to convert the value of data that have a center of 0 and a standard deviation of 1. Normalization is a process to convert the value of data into a minimum value of 0 and a maximum value of 1. However, the technique has limitation for the classification of data for Supervised Machine Learning in WSN-IDS because the value is only dependence on the maximum and minimum value (Tang *et al*., 2018). This means that the normalization technique will produce plain data or linear data. Since this research is focusing on nonlinear data, data standardization will be used.

F. Dataset for WSN-IDS

Datasets have usually been used to test the performance of the 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 the target class is non-uniform (Islabudeen & M K, 2020). 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 (Islabudeen & M K, 2020). 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.

III. Methodology of Research

This section presents the methodology of this research. Figure 3.1 is the depiction of the research framework. This research has three phases. The phases define the objectives of this research. Phase 1 and 2 of this research are adopted from the method used by Tang *et al*. (2018) while phase 3 was adopted from (Jing and Chen, 2019).

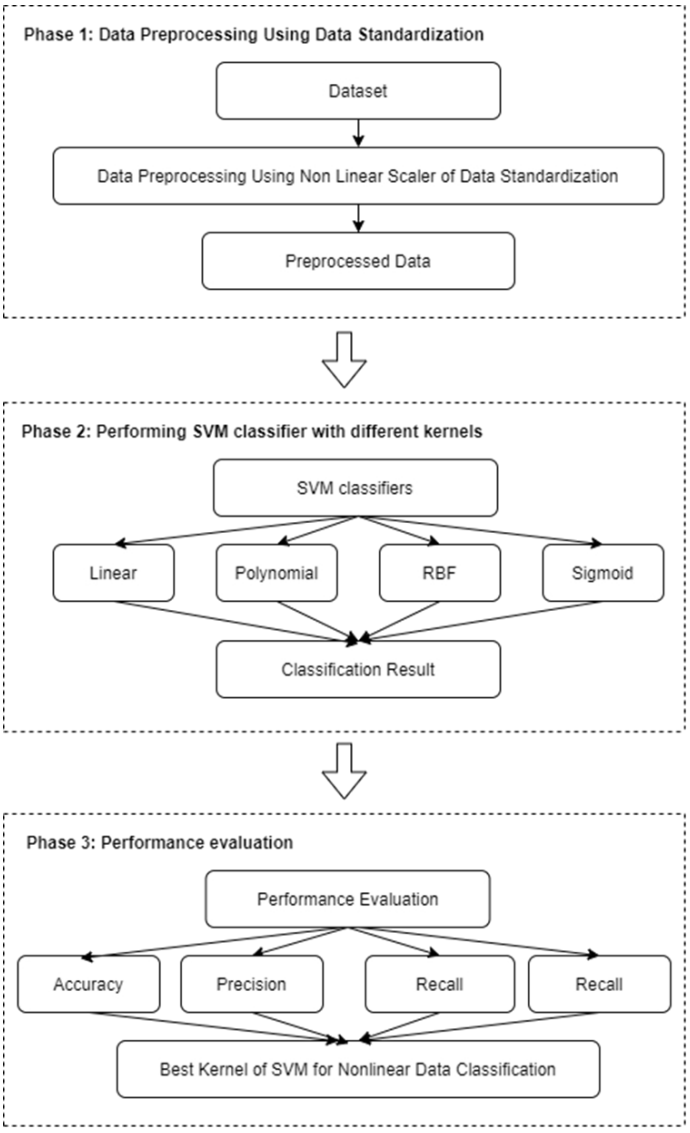

Fig. 3.1. Research Framework

1) Dataset Used.

The dataset used for this experiment is NSL-KDD. 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. This dataset is an open source dataset which can be obtained online and available in *.crv* format.

2) Phase 1: Data Preprocessing Using Nonlinear Scaler of Data Standardization.

Data standardization has several methods and this research use power transformation method. Power transformation is a nonlinear scaling method. In power transformation, data is converted into 0 for mean and 1 for standard deviation. The formula is shown in equation 1

where x_i represent the data point, x represent the mean and s represent the standard deviation. $\lambda \neq 0$ is a simple power transformation where y^λ is rescaled for the lamda, λ in h(y;λ)will be converted to 0.

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

3) Phase 2: Performing SVM with Linear, Polynomial, RBF and Sigmoid.

In this phase, an SVM classifier is used to classify the selected dataset. Four SVM classifiers are performed and each of the classifiers implements a different kernel function. The first SVM classifier used linear as the kernel. A linear kernel is the simplest kernel. The equation of the Linear kernel is shown in equation 2. The value of (x,y) will determine the slope of the hyperplane and the value of constant, C will determine the maximum margin in analyzing subset.

$$\mathrm{K(x,y)} = \mathrm{(x,y)} + \mathrm{C} \tag{2}$$

The second SVM classifier used polynomial as the kernel. Polynomial is suitable for normalized data (Drewnik and Pasternak-Winiarski, 2017). The equation of the polynomial kernel shown in equation 3. In the formula, ^{x,y} determine the slope of the SVM hyperplane, the gamma value, $^{\gamma}$ will determine the curve of the hyperplane and constant, C determine the maximum margin. The parameter of this kernel is determined by the degree, d which helps to add more curve to the hyperplane.

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

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. The equation of RBF is represented in equation 4. Based on the formula, x and y will determine the hyperplane and the value of gamma, -γ will determine the flexibility of it.

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

The last SVM classifier used Sigmoid as the kernel. This kernel has a similar function to neural network, but this kernel has the least classification among the other kernels. The formula is shown below.

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

4) Phase 3: Performance Evaluation.

The performance of each SVM classifier is evaluated in this phase. The evaluation is based on the classification Accuracy, Recall, Precision and F-measure. The formula for the 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 the ratio score.

TABLE 3.1. Measurement Formula for Performance Evaluation

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

To obtain the optimal result. Each of the value gains from each measurement is 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. The value is split into the division of fold and prediction is fit on all points as well as evaluating error on points in each fold (Panda *et al*., 2012).

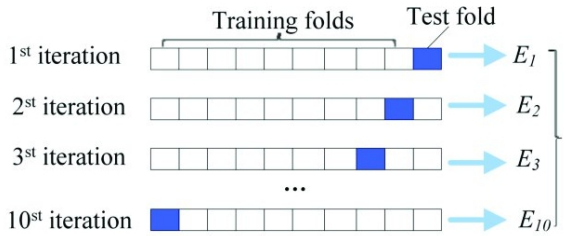

Fig. 3.2. 10-Fold Cross-Validation (Niu *et al*., 2018)

IV. Implementation and Experimental Result

In this section, the setup for the experiment based on the brief discussion made in the previous section will be discussed. This experiment used Jupyter Notebook running in Windows 10 with 4GB of RAM. Libraries available on this platform such as Scikit-learn, Pandas, and NumPy were also used. The execution of this experiment is discussed as follows.

A. Data Preprocessing Using Nonlinear Scaler of Data Standardization.

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. The dataset was loaded into Python Jupyter Notebook and the values of data samples were converted using two steps. The first step was converting the character value of NSL-KDD sample into binary using Label Encoder and the second step was organizing all the values into their feature category. Fig. 4.2 shows the data that has been converted into numerical based on their feature category.

.

```
  protocol_type   service flag
0           tcp  ftp_data   SF
1           udp     other   SF
2           tcp   private   S0
3           tcp      http   SF
4           tcp      http   SF
--------------------
   protocol_type  service  flag
0              1       20     9
1              2       44     9
2              1       49     5
3              1       24     9
4              1       24     9
```

Fig. 4.1. Original Value of Dataset

Further, data was preprocessed using a power transformation method. Scikit learns of Jupyter Notebook has stored this algorithm in the library. Based on the figure, features that have values were scaled using the power transformer which performs data non-linear data scaling. This can be done by making the data as an input variable and the function of the power transformer was called to process it. Fig. 4.2 shows the train and test data that has been converted using the Power Transformer.

```
print (X_Df)
[[-0.29346719  1.01250105 -0.90023028 ... -0.01972622  0.82515007
  -0.04643159]
 [-0.29346719  0.65422036 -0.90023028 ... -0.01972622  0.82515007
  -0.04643159]
 [-0.29346719 -1.11931079 -0.90023028 ... -0.01972622 -1.21190076
  -0.04643159]
 ...
 [-0.29346719  1.42738433  0.95088181 ... -0.01972622  0.82515007
  -0.04643159]
 [-0.29346719 -1.11931079 -0.90023028 ... -0.01972622 -1.21190076
  -0.04643159]
 [-0.29346719  0.66447796 -0.90023028 ... -0.01972622  0.82515007
  -0.04643159]]

print (X_Df_test)
[[-0.42411645 -1.21835318 -1.07964865 ...  0.          0.
   0.        ]
 [-0.42411645 -1.21835318 -1.07964865 ...  0.          0.
   0.        ]
 [ 1.89562852  1.68349851 -1.07964865 ...  0.          0.
   0.        ]
 ...
 [-0.42411645  2.00184152  1.48159865 ...  0.          0.
   0.        ]
 [-0.42411645  0.12408878  0.13357583 ...  0.          0.
   0.        ]
 [-0.42411645 -1.21835318 -1.07964865 ...  0.          0.
   0.        ]]
```

Fig. 4.2. Non-Linear Scaling Using Power Transformer.

B. Performing SVM Classifiers

Four basic SVM model with different kernels each are developed in Jupyter Notebook and other parameters for all the classifiers were set to default. The parameters used are shown in the table below.

TABLE 4.3. Default Parameters for Each Svm Classifier with Different Kernel

SVM Kernel	Constant, C	Gamma	Degree, d
Linear	1.0	scale	-
Polynomial	1.0	scale	3
RBF	1.0	scale	-
Sigmoid	1.0	scale	-

C. Analysis and Discussion

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, the 10-Fold Cross-Validation technique is involved in the calculation. The tables below show the result of performance evaluations.

TABLE 4.9. Performance Result for Kernels

Kernel	Accuracy	Recall	Precision	F-measure
Linear	88	91	89	90
Polynomial	90	91	91	91
RBF	91	91	92	92
Sigmoid	78	83	78	81

Based on the performance evaluations, it is found that RBF has the most accurate result where classification accuracy score is 91%, Recall score 91%, the Precision score is 92% and F-measure score is 92%. The shape of the hyperplane and size of margin of RBF kernel classifies nonlinear data efficiently. In contrast, Sigmoid kernel has the least performance compared to the others kernel with accuracy score of 78%, Recall score 83%, the Precision score is 78% and F-measure score is 81%. Classification made by the decision boundaries in Sigmoid kernel is insufficient to cater nonlinear data which influenced the result shows in Table 4.9. Fig. 4.4 is a bar chart that indicates the performance evaluation scores of each kernel, and it was made for giving a better view regarding the results.

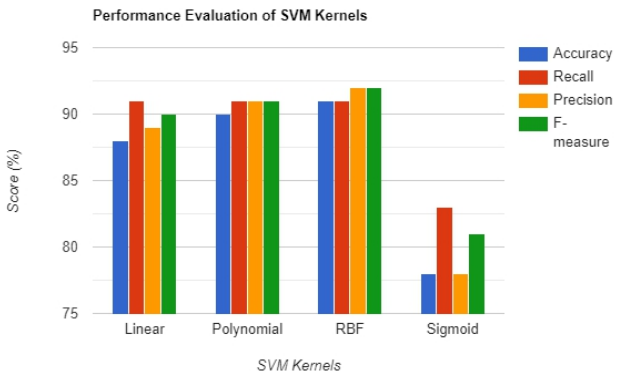

Fig. 4.4. Performance Evaluation of SVM Kernels

To validate the result of this research, the classification accuracy result was compared with the previous works cited in the literature review. Table 4.10 shows the comparison of RBF kernel results. Based on the results, it was found that 91% accuracy of RBF kernel in this research is higher than the result made by Tang *et al.* (2018) but slightly lower than Hasan *et al.* (2016).

TABLE 4.10. The comparison of RBF kernel results

Research	Kernel	Accuracy (%)
Hasan *et al.* (2016).	RBF	91.4
This Research	**RBF**	**91**
Tang *et al.* (2018)	RBF	77.8

V. Conclusion

The efficiency of SVM kernel for classification is based on the parameter used and how dataset is presented. In SVM, data can be presented using data standardization and it is important to standardize the data before classification to avoid errors. This research is focusing on the performance of Linear, RBF, Polynomial and Sigmoid kernel in classifying nonlinear data. Thus, dataset was standardize using a nonlinear scaler. 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 data with 91% accuracy, 91% recall, 92% precision and 92% F-measure. The performance of RBF kernel in was influenced by architecture of the classifiers that affect the margin size which makes data classification more efficient. Therefore, it can enhance detection of attack in WSN-IDS.

In the future, the result of this experiment will be analyzed further and a comparison between the result obtained from this work with the existing research will be done. By conducting the analysis, the result obtained from this research can be validated. In addition, this would also help to find more research gap.

Acknowledgement

First of all, praise Allah for giving me strength with determination and grant me the knowledge that allows me to finish this project. I would like to express my greatest gratitude to my supervisor Ts. Dr Siti Hajar Binti Othman for her assistant and support as well as the knowledge that gave to me throughout the whole process of this project. Besides, I would like to express my appreciation to my parents that have sacrificed to provide me with the best in everything and constantly give moral support which allows me to survive until this stage. I also would like to thank my family and friend for the endless support. Finally, I must express my appreciation and gratitude to Nur Amirah Khairina Binti Khairil Anwar that has helped and teach me in surviving this adventurous journey as a postgraduate student.

References

[1] Abd Manaf, A., Sahibuddin, S., Ahmad, R., Daud, S. M., & El-Qawasmeh, E. (2011). Informatics Engineering and Information Science. *Conference Proceedings ICIEIS*, 42.

[2] Achirul Nanda, M., Seminar, K., Nandika, D., & Maddu, A. (2018). A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection. *Information*, 9, 5.

[3] Ahmad, B., Jian, W., & Anwar Ali, Z. (2018). Role of Machine Learning and Data Mining in Internet Security: Standing State with Future Directions. *Journal of Computer Networks and Communications*, 6383145.

[4] Al Mehedi Hasan, M., Nasser, M., & Pal, B. (2013). On the KDD'99 Dataset: Support Vector Machine Based Intrusion Detection System (ids) with Different Kernels. *Int. J. Electron. Commun. Comput. Eng*, 4(4), 1164-1170.

[5] Alrajeh, N. A., Khan, S., & Shams, B. (2013). Intrusion Detection Systems in Wireless Sensor Networks: A Review. *International Journal of Distributed Sensor Networks*, 9(5), 167575.

[6] Anderson, J. P. (1980). Computer Security Threat Monitoring and Surveillance, James P. *Anderson Co., Fort Washington, PA*.

[7] Ashwini, B. A., & Manivannan, S. S. (2020). Supervised Machine Learning Classification Algorithmic Approach for Finding Anomaly Type of Intrusion Detection in Wireless Sensor Network. *Optical Memory and Neural Networks (Information Optics)*, 29(3), 244-256.

[8] Baraneetharan, E. (2020). *Role of Machine Learning Algorithms Intrusion Detection in WSNs: A Survey*.

[9] Belavagi, M. C., & Muniyal, B. (2016). Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection. *Procedia Computer Science*, 89, 117-123.

[10] Bhavsar, H., & Panchal, M. H. (2012). A Review on Support Vector Machine for Data Classification. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 1(10), 185-189.

[11] Bhavsar, Y., & Waghmare, K. (2013). *Intrusion Detection System Using Data Mining Technique : Support Vector*

Machine.

[12] Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2014). Towards an Unsupervised Method for Network Anomaly Detection in Large Datasets. *Computing and Informatics*, 33(1), 1-34.

[13] Boubiche, D. E., Athmani, S., Boubiche, S., & Toral-Cruz, H. (2020). Cybersecurity Issues in Wireless Sensor Networks: Current Challenges and Solutions. *Wireless Personal Communications*, 1-37.

[14] Butun, I., Morgera, S., & Sankar, R. (2013). A Survey of Intrusion Detection Systems in Wireless Sensor Networks. *IEEE Communications Surveys & Tutorials*, *PP*, 266-282.

[15] Cahyo, A. N., Hidayat, R., & Adhipta, D. (2016). Performance Comparison of Intrusion Detection System Based Anomaly Detection Using Artificial Neural Network and Support Vector Machine. *AIP Conference Proceedings*, 1755(1), 70011.

[16] Cervantes, J., García-Lamont, F., Rodríguez, L., & Lopez-Chau, A. (2020). A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends. *Neurocomputing*, 408.

[17] Chandra, M. A., & Bedi, S. S. (2018). Survey on SVM and Their Application in Image Classification. *International Journal of Information Technology*, 1-11.

[18] Denning, D. E. (1987). An Intrusion-Detection Model. *IEEE Transactions on Software Engineering*, SE-13(2), 222-232.

[19] Drewnik, M., & Pasternak-Winiarski, Z. (2017). SVM Kernel Configuration and Optimization for the Handwritten Digit Recognition. *CISIM*.

[20] Farnaaz, N., & Jabbar, M. A. (2016). Random Forest Modeling for Network Intrusion Detection System. *Procedia Computer Science*, 89, 213-217.

[21] Ghanem, W. A. H. M., & Jantan, A. (2018). Hybridizing Artificial Bee Colony with Monarch Butterfly Optimization for Numerical Optimization Problems. *Neural Computing and Applications*, 30(1), 163-181.

[22] Hasan, M. Al, Nasser, M., Ahmad, S., & Molla, M. K. (2016). Feature Selection for Intrusion Detection Using Random Forest. *Journal of Information Security*, 07, 129-140.

[23] Islabudeen, M., & M K, K. D. (2020). A Smart Approach for Intrusion Detection and Prevention System in Mobile Ad Hoc Networks Against Security Attacks. *Wireless Personal Communications*, 112.

[24] Jadidoleslamy, H. (2011). A Hierarchical Intrusion Detection Architecture for Wireless Sensor Networks. *International Journal of Network Security & Its Applications*, 3(5), 131.

[25] Jing, D., & Chen, H. (2019). SVM Based Network Intrusion Detection for the UNSW-NB15 Dataset. *2019 IEEE 13th International Conference on ASIC (ASICON)*, 1-4.

[26] Kausar, N., Brahim Belhaouari, S., Abdullah, A., Ahmad, I., & Hussain, M. (2011). A Review of Classification Approaches Using Support Vector Machine in Intrusion Detection. *Communications in Computer and Information Science*, 253.

[27] Kotpalliwar, M. V, & Wajgi, R. (2015). Classification of Attacks Using Support Vector Machine (SVM) on KDDCUP'99 IDS Database. *2015 Fifth International Conference on Communication Systems and Network Technologies*, 987-990.

[28] Kuang, F., Zhang, S., Jin, Z., & Xu, W. (2015). A Novel SVM by Combining Kernel Principal Component Analysis and Improved Chaotic Particle Swarm Optimization for Intrusion Detection. *Soft Computing*, 19(5), 1187-1199.

[29] Kumar, G., Kumar, K., & Sachdeva, M. (2010). The Use of Artificial Intelligence Based Techniques for Intrusion Detection: A Review. *Artificial Intelligence Review*, 34(4), 369-387.

[30] Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., & Kim, K. J. (2019). A Survey of Deep Learning-based Network Anomaly Detection. *Cluster Computing*, 22(1), 949-961.

[31] Liao, H.-J., Lin, C.-H. R., Lin, Y.-C., & Tung, K.-Y. (2013). Intrusion Detection System: A Comprehensive Review. *Journal of Network and Computer Applications*, 36(1), 16-24.

[32] Liu, Z., Zuo, M., Zhao, X., & Xu, H. (2015). An Analytical Approach to Fast Parameter Selection of Gaussian RBF Kernel for Support Vector Machine. *Journal of Information Science and Engineering*, 31, 691-710.

[33] Mathew, J., Pang, C. K., Luo, M., & Leong, W. H. (2017). Classification of imbalanced Data by Oversampling in Kernel Space of Support Vector Machines. *IEEE Transactions on Neural Networks and Learning Systems*, 29(9), 4065-4076.

[34] Mitrokotsa, A., & Douligeris, C. (2005). Detecting denial of service attacks using emergent self-organizing maps. *Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology, 2005*, 375-380.

[35] Niu, M., Li, Y., Wang, C., & Han, K. (2018). RFAmyloid: A Web Server for Predicting Amyloid Proteins. *International Journal of Molecular Sciences*, 19.

[36] Panda, M., Abraham, A., & Patra, M. R. (2012). A Hybrid Intelligent Approach for Network Intrusion Detection. *Procedia Engineering*, 30, 1-9.

[37] Reddy, R. R., Ramadevi, Y., & Sunitha, K. V. N. (2015). Anomaly Detection using Feature Selection and SVM Kernel Trick. *International Journal of Computer Applications*, 975, 8887.

[38] Reddy, R. R., Ramadevi, Y., & Sunitha, K. V. N. (2016). Effective discriminant function for intrusion detection using SVM. *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1148-1153.

[39] Rustam, Z., & Zahras, D. (2018). Comparison between Support Vector Machine and Fuzzy C-Means as Classifier for Intrusion Detection System. *Journal of Physics: Conference Series*, 1028, 12227.

[40] Sharma, S., & Gupta, R. (2015). Intrusion Detection System: A Review. *International Journal of Security and Its Applications*, 9, 69-76.

[41] Tang, X., Tan, S. X.-, & Chen, H. (2018). SVM Based Intrusion Detection Using Nonlinear Scaling Scheme. *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, 1-4.

[42] Vapnik, V. (1998). The Support Vector Method of Function Estimation. *Nonlinear Modeling* (pp. 55–85). Springer.

[43] Zanero, S., & Savaresi, S. (2003). Unsupervised Learning Techniques for an Intrusion Detection System. *Proceedings of the ACM Symposium on Applied Computing*, 1.