1. Preliminary

Background

Software Defined Network (SDN) is a new technology in the concept of computer networking, where SDN can separate the data plane and control plane. However, SDN still has potential for Denial of Service (DoS) attacks. This happens because SDN performs network control centrally through a controller [6]. Denial of Service (DoS) attacks are among common attacks on the network. DoS attacks prevent users from normal service access, due to the excessive consumption of network resources, memory, processor, etc. The most common DoS attacks are when attackers “flood” the network with many request at the same time, making the server unavailable to answer to that many request[1]. A lot of research has conducted the detection and mitigation of DoS attacks, in which research on DoS attack detection is carried out using machine learning models or algorithms trained by the NSL-KDD dataset. This dataset is often used for DoS detection because some of the attributes in the dataset are closely related to DoS. Research that has been done often concludes the results on the accuracy of the model obtained after being trained without validating it. therefore, it is necessary to validate the model used and has been trained by datasets taken from the SDN network, so that the accuracy results of the model can be said to be more precise.

In [2] that research was conducted to detect DoS attacks using the NSL-KDD dataset. The research detects DoS attacks using Probabilistic Neural Network (PNN) algorithm as a model that has been trained with NSL-KDD dataset, the accuracy of that model is 98.06%. That means based on the research, the NSL-KDD can be used to train the model. in order to detect DoS attacks. in [3] that research was conducted to detect DoS attacks on SDN networks using the NSL-KDD dataset. In this research, the models used were Decision Tree and Nave Bayes, where in this research was conducted using SDN and using the NSL-KDD dataset as the dataset. The accuracy score of the models used include 99.0% for the decision tree and 97.0% for nave Bayes. Research was also conducted [10], where in this research the model used was Deep Neural Network (DNN). The result of this model is 75.75%. Based on these studies, it can be concluded that the NSL-KDD dataset can used to detect DoS attacks on SDN networks. Because in [[2],[3],[10]] used NSL-KDD for the dataset to detect DoS on the SDN and the features on the NSL-KDD can classify DoS attacks and dataset in this research from Canadian Institute for Cybersecurity on UNB.

In [4] conducted a research using data train results from the SDN network data generation that had been made and classified using algorithms without validating the model. Meanwhile, one way that we can see how good the model we choose is by validating it. So from this research it can be stated that the model used has not been validated so it cannot be said that the model is good for classification. Based on these problems, this research validates a model that is used to detect DoS attacks. This validation will use machine learning validation, namely K-Fold Cross Validation. This method was chosen because match with the model to be used and the dataset. This research will use the dataset resulting from capturing traffic on the Software Defined Network (SDN) network and will also use the NSL-KDD dataset as a reference, so that it can be used by later modeling. The purpose of this research is to prove how well a model is used for DoS attack detection through the results obtained and find out how to validate the model so that it can be used for other models.

Problem Statement and Problem Identification

As discussed in the previous section, the validation process has never been carried out on a SDN network and also these research as a reference only used accuracy from the model to assess the algorithm perform. While the validation process is one way to find out how well the model is used. So with that, in this research validation is carried out on the SDN network and using K-Fold Cross Validation to know algorithm perform, this method is a validation method commonly used in machine learning and also this method is compatible with the machine learning algorithm and the dataset that used in this research. This research will use datasets obtained from the Software Defined Network (SDN). The dataset will also be shaped like the NSL-KDD, this is because that dataset has been used in several research to conduct training on DoS detection models on Software Defined Network (SDN). In this research, there are 12 features of NSL-KDD dataset that are used as a reference. This research will only carry out the validation process. where the results of the validation process will show a good model used for detection DoS.

Purposes

The purpose of this research is to find out how to carry out the validation process and also prove how well the model is used by looking at the results of the validation values using the K-Fold Cross Validation method. The relationship between objectives, testing and conclusions can be seen in Table 1.

Tabel 1. The relationship between objectives, testing and conclusions

No	Objectives	Testing	Conclusion
1	Knowing how to carry out the validation process and also proving how well the model is used to detect DoS by validating it using K-Fold Cross Validation,	The model used is SVM, Naïve Bayes and Decision Tree. the model is trained with datasets taken from the SDN network and validates using the K-Fold Cross Validation method.	results of that model in validation process are 99.79% for SVM, 99.84% for Decision Tree and 96.84% for Naïve Bayes. these results prove that the model is good for detecting. Although no the highest, it's because the SVM algorithm still has shortcomings in the optimization process.

Writing Sections

The next part is the research methodology which will discuss the things that support this research. After that section is the system built where in this section it will be explained about how to validate this research to get the results. The output of the system built section will be explained and analyze further in the evaluation section and the last section is conclusion which will conclude the results of this research.