Support Vector Machine Model For Threat Detection And Classification In Deep Packet Inspection Applications

Ogbonnaya Irukwu Joe[1]
Department of Electrical / Electronic Engineering
University of Uyo, Akwa Ibom State

Kufre M. Udofia[2]
Department of Electrical / Electronic Engineering
University of Uyo, Akwa Ibom State

Akaninyene Bernard Obot[3]
Department of Electrical / Electronic Engineering
University of Uyo, Akwa Ibom State

***Abstract*—In this study Support Vector Machine (SVM) model for threat detection and classification in deep packet inspection applications is presented. The focus was to utilise SVM for the detection of network threat at the point of packet arrival to a network. The result obtained from the SVM model is then employed in the deep packet inspection (DPI) for detection of intrusion on the Software Defined Network (SDN) facility. Network threat dataset from Alibaba Silexscure limited and Kaggle repository was used in the SVM model development. The model evolution results showed that benign attack had 88.9% for TPR and 11.2% for FNR, DDoS was classified with 64.8% for TPR and 36.2% for FNR. Also, web based attack considering brute force recorded 99.4% for TPR and 0.6% for FNR, while SQL injection attack recorded 100% for TPR and 0% for FNR. Web based XSS attack recorded null TPR and 100% FNR, while for normal packet classification, the TPR reported 100% and 0% FNR. In all, the results showed that the model was very good in classifying certain threats such as benign, DDoS, SQL and normal packet, it was however fair in detecting brute force attack and also was not able to correctly classify XSS attack.**

Keywords— Support Vector Machine, Deep Packet Inspection, Threat Detection, Threat Classification, Network Security

1. INTRODUCTION

In recent years, researchers have revealed that virtually all sectors of the global economy fell victim to cybercrime, marking an alarming escalation in the global cyber-threat landscape [1,2,3]. Coupled with the sophistication of the attack models due to the advancement of technologies and also the inevitability of flaws in the targeted network infrastructures, this has underscored the critical importance of robust Cyber Security (CS) measures against cyber threats [4,5].

According to [6,7], a cyber-threat is a malicious act perpetrated by an attacker to damage, disrupt, steal, or compromise digital life with the motive to cause harm to an organization or individual. In other words, it is an activity tailored toward the compromise of an automated information system through unauthorized access, destruction, disclosure, modification of information, or denial of service [8,9]. The process of cyberattack begins with the scanning of network infrastructure for vulnerabilities by the threat actors, and then the identified flaws are exploited for the attack using threat features such as malware, viruses, worms, data breaches, denial of service attacks, and other attack vectors [10,11]. When this threat penetration is successful, [12,13] revealed that the implications can be devastating, with huge consequences such as disruption of critical services, financial losses, reputational damage, compromise of sensitive information, and, in some cases, even national security threats, hence the need for an urgent solution.

Accordingly, in this work, Support Vector Machine (SVM) model for threat detection and classification in deep packet inspection applications is presented [14,15]. This study is therefore, focused on modelling and demonstration of a smart deep packet inspection framework for the security of critical network infrastructure and this can be integrated with deception based machine learning technique. When integrated, the deception approach can then divert the attacker to a decoy network using approach like Honey-X technique, while the machine learning is utilized for deep packet inspection.

2. METHODOLOGY

Support Vector Machine (SVM) model can be applied for both classification and regression problems. In this work, the SVM algorithm was adopted to train and generate the deep packet inspection (DPI) model for the detection of intrusion on the Software Defined Network (SDN) facility. The SVM works by transforming the input data into higher dimensional features and then utilize decision boundaries which are trained as hyper-plane that separate the features in classes. The hyper-planes are trained using kernel function selection, optimal hyper parameter settings with quadratic programming technique to solve the optimization problem of the hyper plane decision boundary, which are then applied to make prediction [17]. The diagram in Figure 1 presented the training sequence of the SVM to generate the deep packet inspection model.

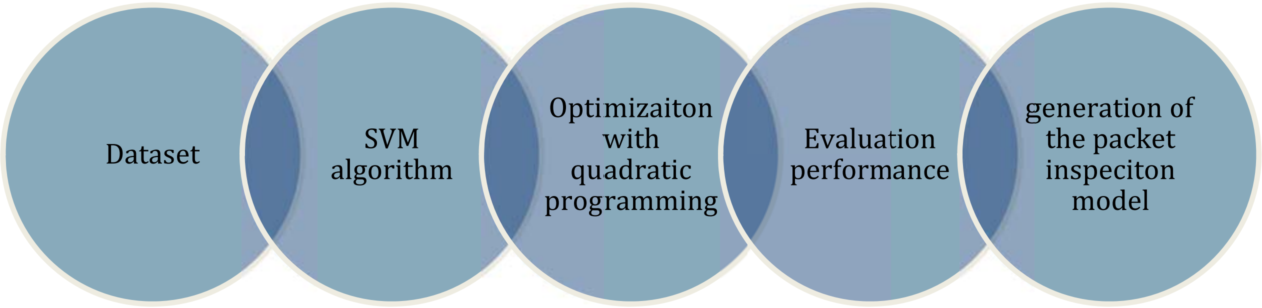
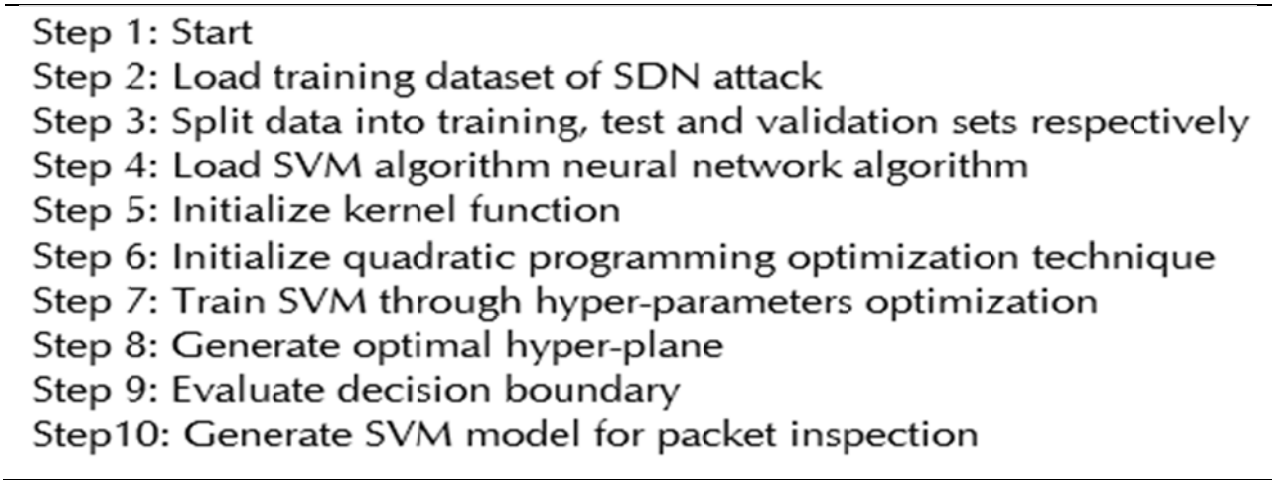

Dataset
SVM algorithm
Optimizaiton with quadratic programming
Evaluation performance
generation of the packet inspeciton model

Figure 1: Sequence for the SVM based deep packet inspection model generation

First, the collected data are imported to the SVM algorithm. This SVM converts the data into high dimensional space for separation by the decision boundary of the hyper-plane. To train the hyper-plane, quadratic programming technique is applied which adjust the hyper–parameters of the SVM such as the kernel function, and cost function objectives until the best decision classifier is generated which is the model for the detection of abnormally in the network infrastructure.

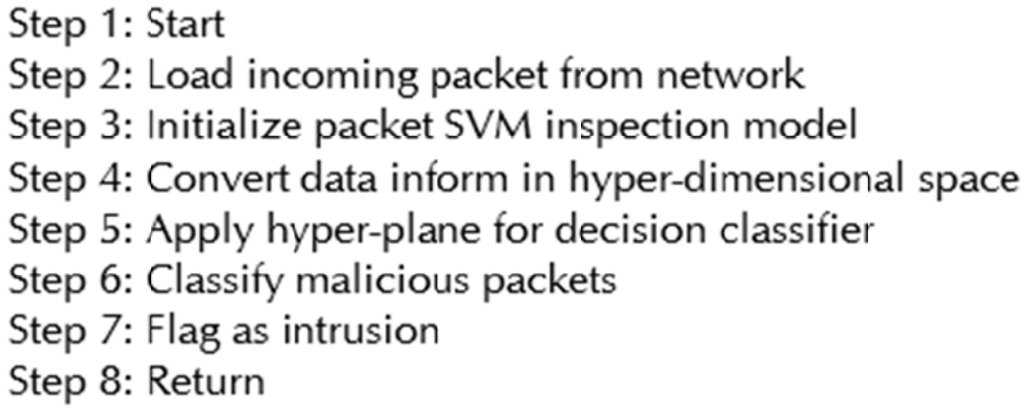
The sequence of operation for the SVM training and deep packet inspection generation is presented in Algorithm 1 and Algorithm 2. The SVM was trained using network threat dataset from Alibaba Silexscure limited and Kaggle repository. The dataset contained different threat classes out of which six were considered in this study and the include; Structured Queried Language (SQL) injection attack, brute force, distributed denial of service (DDoS), benign, and normal packet. The detailed composition of the case study threat dataset is presented in Table 1. The performance metrics adopted for the model assessment include False Discovery Rate (FDR), Positive Predictive Value (PPV), False Negative Rate (FNR), Accuracy, True Positive Rate, as well as Receiver Operating Characteristic (ROC).

Algorithm 1: The procedure for the deep packet generation with SVM



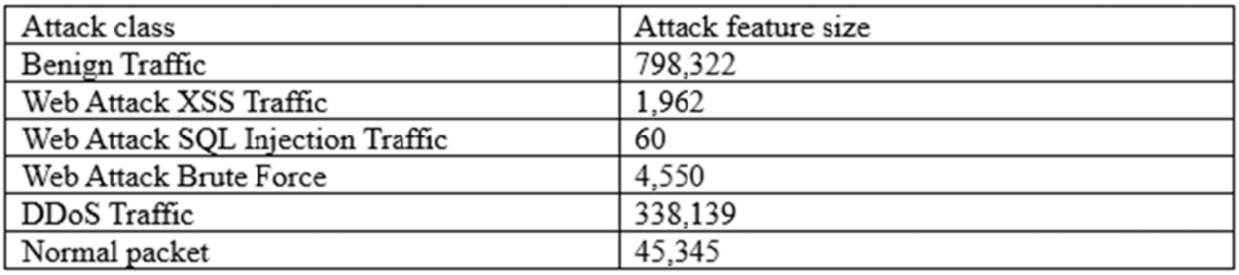
```
Step 1: Start
Step 2: Load training dataset of SDN attack
Step 3: Split data into training, test and validation sets respectively
Step 4: Load SVM algorithm neural network algorithm
Step 5: Initialize kernel function
Step 6: Initialize quadratic programming optimization technique
Step 7: Train SVM through hyper-parameters optimization
Step 8: Generate optimal hyper-plane
Step 9: Evaluate decision boundary
Step10: Generate SVM model for packet inspection
```

Algorithm 2: The SVM-based deep packet inspection operation



```
Step 1: Start
Step 2: Load incoming packet from network
Step 3: Initialize packet SVM inspection model
Step 4: Convert data inform in hyper-dimensional space
Step 5: Apply hyper-plane for decision classifier
Step 6: Classify malicious packets
Step 7: Flag as intrusion
Step 8: Return
```

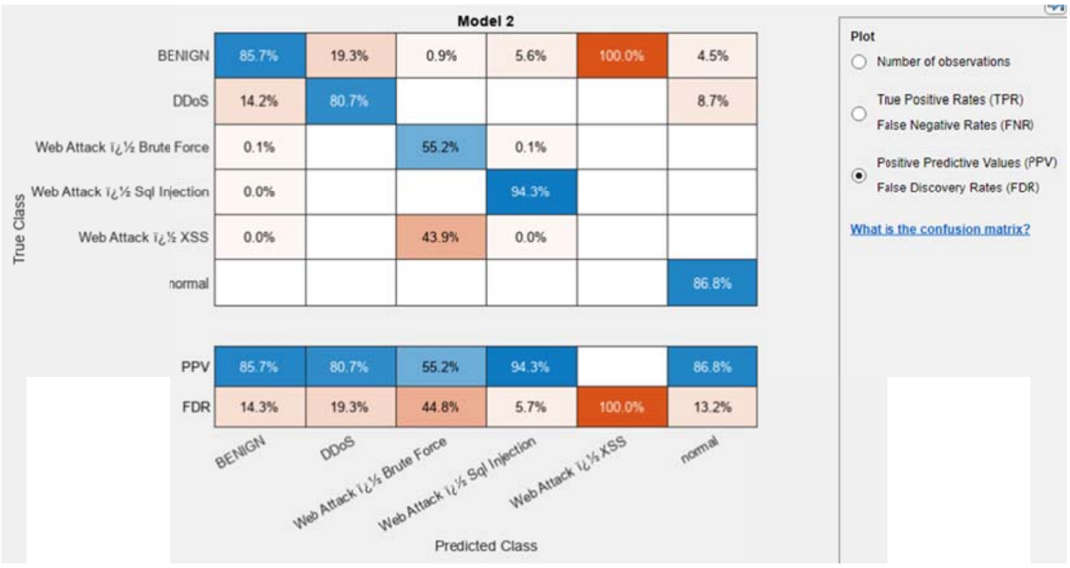
Table 1: The composition of the case study threat dataset



Attack class	Attack feature size
Benign Traffic	798,322
Web Attack XSS Traffic	1,962
Web Attack SQL Injection Traffic	60
Web Attack Brute Force	4,550
DDoS Traffic	338,139
Normal packet	45,345

3. RESULTS AND DISCUSSIONS

The model performance in terms of PPV and FDR are presented in Figure 2. The results show that the benign threat recorded 85.7% and 14.3% FDR, for DDoS the PPV reported 80.7% and FDR of 19.3%, for web based brute force attack the model recorded 56.2% PPV and 44.8% FDR, web based SQL injection attack recorded 94.7% PPV and 5.7% FDR, while normal packet recorded 86.8% PPV and 13.2% FDR. The model recorded null for web based XSS attack PPV and 100% FDR.



Model 2

True Class	BENIGN	DDoS	Web Attack ï¿½ Brute Force	Web Attack ï¿½ Sql Injection	Web Attack ï¿½ XSS	normal
BENIGN	85.7%	19.3%	0.9%	5.6%	100.0%	4.5%
DDoS	14.2%	80.7%				8.7%
Web Attack ï¿½ Brute Force	0.1%		55.2%	0.1%		
Web Attack ï¿½ Sql Injection	0.0%			94.3%		
Web Attack ï¿½ XSS	0.0%		43.9%	0.0%		
normal						86.8%

	BENIGN	DDoS	Web Attack ï¿½ Brute Force	Web Attack ï¿½ Sql Injection	Web Attack ï¿½ XSS	normal
PPV	85.7%	80.7%	55.2%	94.3%		86.8%
FDR	14.3%	19.3%	44.8%	5.7%	100.0%	13.2%

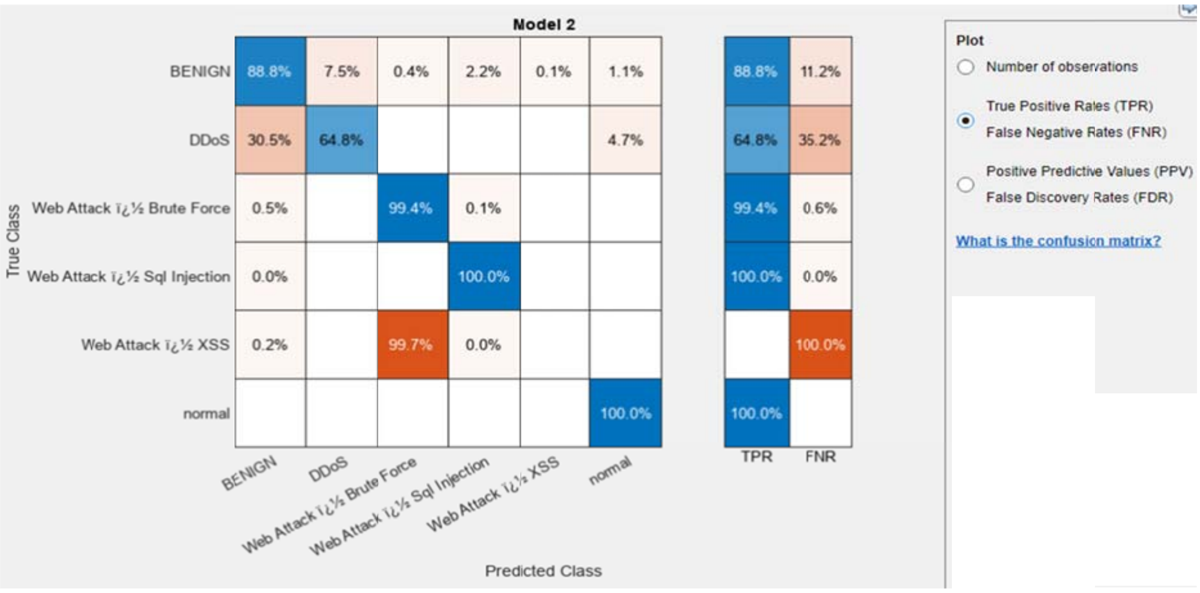
Predicted Class

Plot
○ Number of observations
○ True Positive Rates (TPR) False Negative Rates (FNR)
● Positive Predictive Values (PPV) False Discovery Rates (FDR)
What is the confusion matrix?

Figure 2: PPV and FDR for SVM-based DPI model

The Figure 3 presents the DPI model performance considering the TPR and FNR respectively. From the result, it was noticed that benign attack had 88.9% for TPR and 11.2% for FNR, DDoS was classified with 64.8% for TPR and 36.2% for FNR. Also, web based attack considering brute force recorded 99.4% for TPR and 0.6% for FNR, while SQL injection attack recorded 100% for TPR and 0% for FNR. Web based XSS attack recorded null TPR and

100% FNR, while for normal packet classification, the TPR reported 100% and 0% FNR.



Model 2

True Class	BENIGN	DDoS	Web Attack ï¿½ Brute Force	Web Attack ï¿½ Sql Injection	Web Attack ï¿½ XSS	normal	TPR	FNR
BENIGN	88.8%	7.5%	0.4%	2.2%	0.1%	1.1%	88.8%	11.2%
DDoS	30.5%	64.8%				4.7%	64.8%	35.2%
Web Attack ï¿½ Brute Force	0.5%		99.4%	0.1%			99.4%	0.6%
Web Attack ï¿½ Sql Injection	0.0%			100.0%			100.0%	0.0%
Web Attack ï¿½ XSS	0.2%		99.7%	0.0%				100.0%
normal						100.0%	100.0%	

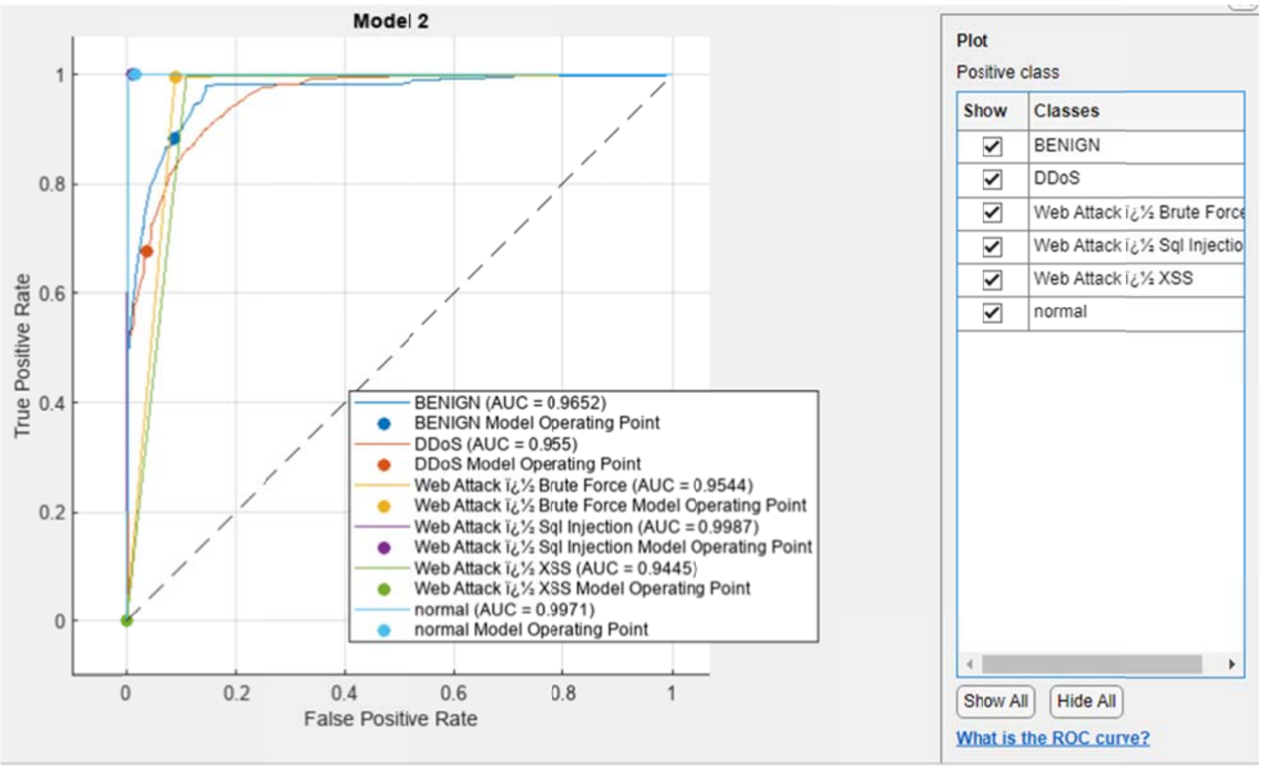
Predicted Class

Plot
○ Number of observations
● True Positive Rates (TPR)
False Negative Rates (FNR)
○ Positive Predictive Values (PPV)
False Discovery Rates (FDR)

What is the confusion matrix?

Figure 3: SVM based DPI result considering TPR and FNR

The results in Figure 4 presents the ROC for the SVM based DPI model generated. The results showed that the AUC for benign classification is 0.9662, DDoS recorded 0.955, brute force attack reported 0.9554, SQL injection attack reported 0.9997, XSS attack reported 0.9445, and normal packet recorded 0.9971, while the accuracy is 89.9% respectively.



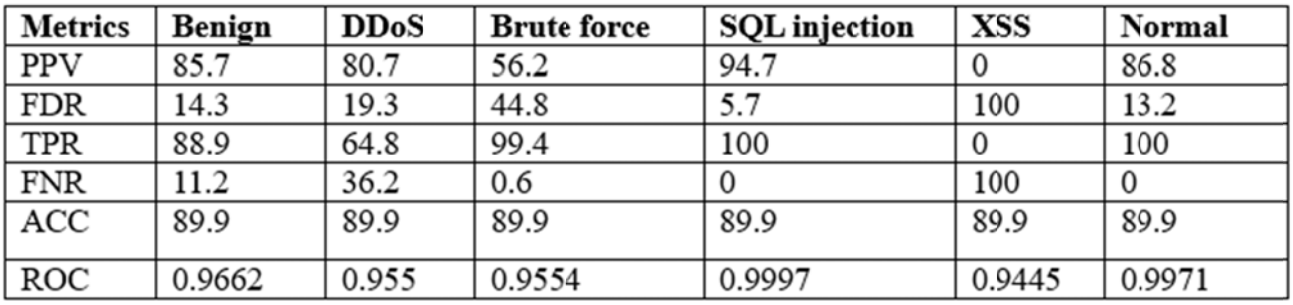
Plot
Positive class

Show	Classes
☑	BENIGN
☑	DDoS
☑	Web Attack ï¿½ Brute Force
☑	Web Attack ï¿½ Sql Injectio
☑	Web Attack ï¿½ XSS
☑	normal

Figure 4: ROC for the SVM based DPI model

The performance evaluation results is summarized for the SVM based DPI model and presented in Table 2. From the result it as observed that while the model was very good in classifying certain threats such as benign, DDoS, SQL and normal packet, it was fair in detecting brute force attack and also was not able to correctly classify XSS attack.

Table 2: The performance evaluation results summary for the SVM based DPI model



Metrics	Benign	DDoS	Brute force	SQL injection	XSS	Normal
PPV	85.7	80.7	56.2	94.7	0	86.8
FDR	14.3	19.3	44.8	5.7	100	13.2
TPR	88.9	64.8	99.4	100	0	100
FNR	11.2	36.2	0.6	0	100	0
ACC	89.9	89.9	89.9	89.9	89.9	89.9
ROC	0.9662	0.955	0.9554	0.9997	0.9445	0.9971

4. CONCLUSION

Threat detection in network traffic using Support Vector Machine (SVM) machine learning model is presented. The SVM model is meant to be utilised by deep packet inspection for the detection of intrusion on the Software Defined Network (SDN) facility. The model was trained evaluated based on a threat dataset, with consideration of only six different threat categories in the dataset, where the threat categories are; Structured Queried Language (SQL) injection attack, brute force, distributed denial of service (DDoS), benign, and normal packet. The results showed that the model was very good in classifying certain threats such as benign, DDoS, SQL and normal packet, it was however fair in detecting brute force attack and also was not able to correctly classify XSS attack.

REFERENCES

1. Afaq, S. A., Husain, M. S., Bello, A., & Sadia, H. (2023). A critical analysis of cyber threats and their global impact. In *Computational Intelligent Security in Wireless Communications* (pp. 201-220). CRC Press.
2. Fatima, A. (2023). Cybercrime: Investigating the Growing Threat of Online Crime. *Social Science Review Archives*, *1*(1), 10-17.
3. Mphatheni, M. R., & Maluleke, W. (2022). Cybersecurity as a response to combating cybercrime: Demystifying the prevailing threats and offering recommendations to the African regions. *International journal of research in business and social science*, *11*(4), 384-396.
4. Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. *Electronics*, *12*(6), 1333.
5. Abdullayeva, F. (2023). Cyber resilience and cyber security issues of intelligent cloud computing systems. *Results in Control and Optimization*, *12*, 100268.
6. Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. *Electronics*, *12*(6), 1333.
7. Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports*, *7*, 8176-8186.
8. Admass, W. S., Munaye, Y. Y., & Diro, A. A. (2024). Cyber security: State of the art, challenges and future directions. *Cyber Security and Applications*, *2*, 100031.
9. Ajala, O. A., Okoye, C. C., Ofodile, O. C., Arinze, C. A., & Daraojimba, O. D. (2024). Review of AI and machine learning applications to predict and Thwart cyber-attacks in real-time. *Magna Scientia Advanced Research and Reviews*, *10*(1), 312-320.
10. Ebelogu, C. U., Prasad, R., Bisallah, H. I., Hammawa, B. M., & Musa, I. (2025). Investigation of Cybersecurity Vulnerabilities and Mitigation Strategies in Nigeria's Oil and Gas Industry. *ABUAD Journal of Engineering Research and Development (AJERD)*, *8*(1), 140-150.
11. Altulaihan, E., Almaiah, M. A., & Aljughaiman, A. (2022). Cybersecurity threats, countermeasures and mitigation techniques on the IoT: Future research directions. *Electronics*, *11*(20), 3330.
12. Alqudhaibi, A., Krishna, A., Jagtap, S., Williams, N., Afy-Shararah, M., & Salonitis, K. (2024). Cybersecurity 4.0: safeguarding trust and production in the digital food industry era. *Discover Food*, *4*(1), 2.
13. Perera, S., Jin, X., Maurushat, A., & Opoku, D. G. J. (2022, March). Factors affecting reputational damage to organisations due to cyberattacks. In *Informatics* (Vol. 9, No. 1, p. 28). Multidisciplinary Digital Publishing Institute.
14. Ioannou, C., & Vassiliou, V. (2021). Network attack classification in IoT using support vector machines. *Journal of sensor and actuator networks*, *10*(3), 58.
15. Chu, W. L., Lin, C. J., & Chang, K. N. (2019). Detection and classification of advanced persistent threats and attacks using the support vector machine. *Applied Sciences*, *9*(21), 4579.
16. BibalBenifa J., Krishnann S., Long H., Kumar R., &Taniar D., (2021) Performance Analysis of Machine Learning and Pattern Matching Techniques for Deep Packet Inspection in Firewalls. Research Square: https://doi.org/10.21203/rs.3.rs-260788/v1