Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 3, Issue 10, 2025

ENHANCING INTRUSION DETECTION IN SOFTWARE-DEFINED NETWORKS THROUGH ENSEMBLE LEARNING TECHNIQUES

Kishwar Noor[*1], Chaman Saeed[2], Naeem Aslam[3], Nimra Bashir[4], Hira Saleem[5]

[*1,2,3,4,5]*Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan*

DOI:https://doi.org/10.5281/zenodo.17322284

Keywords
Ensemble learning, Intrusion detection, Software-Defined Networks, Machine learning, Network security

Article History
Received: 20 August 2025
Accepted: 30 September 2025
Published: 11 October 2025

Copyright @Author
Corresponding Author: *
Kishwar Noor

Abstract
The paper recommendations consist of a better method of intrusion detection in Software-Defined Networks (SDNs) by the joint methods of ensemble learning. The article reveals the latent issue of flexibility, scalability and robustness on network security under the foundation of diverse machine learning algorithms, Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN) separately and in combination. The three different sets of collected data used in Kaggle repositories that incorporate the general network intrusion, attack specific in the IoT, and SDN specific conditions were used to conduct experiments. The model optimization consisted of the rigorous preprocessing, the feature engineering and model tuning. The findings show that the ensemble model performance has been higher at all times, the accuracy score is over 99.5 and F1 scores are close to 0.99 under all the evaluating conditions. The comparative analysis also established that even though the good performance was convincingly achieved with the individual models, the collective approach is superior in enhancing the overall performance of the generalization as well as the false alarms especially in the heterogeneous and the high-traffic conditions such as the case with SDNs of the modern and the internet of things networks. The technical discussion shows the practicality and the practical outcomes of the ensemble learning such as its ability to accommodate the imbalance of the classes, the high dimensional feature space, and the dynamical traffic pattern. This study concludes by finding that ensemble learning offers scalable, powerful, and explainable architecture of the state-of-the-art intrusion detection. The subsequent steps are explainable AI and federated learning and ML to provide more transparency, privacy, and automation to real-time implementations of network security.

1. INTRODUCTION

The SDN is a new technology, which has transformed the life of the current network architecture and network management in the past few years. The SDN is also offered with the paradigm shift lacking in the traditional networks which relies on the inactive and distributed hardware environment and de-couples the data and the control plane and transforms the network intelligent and control to an isolated central controller. The isolation of architecture like this provides freedom of coordination and tuning network resources to the network administrators and to control traffic, as well as to adjust to the new network environment as they choose at any time as never before[1], [2]. Therefore, SDNs are at the luxury of being more programmable, easier to install, and less multiplicities in their operations thus very attractive to the modern network infrastructures. However, this centralization has resulted in a second security risk in

that, now, the SDN controller is another target attacker and he/ she can impact the network and alter the network or intrude into the network.

The complexity of network environments besides being heterogeneous on top of dynamism of SDNs have posed a major challenge to the conventional security mechanisms. The old signatures-based Intrusion Detection System (IDS) that run on the already devised attacks patterns will barely be able to detect the new or intelligent attacks that continue to evolve within the network perimeter. The weakness has led to the requirement of the anomaly-based IDS systems in the cybersecurity systems. Such mechanisms define the normative network behavior profiles and raise an alarm in the event of such violation that might otherwise be a zero-day attack or an advanced persistent threat (APT) or other novel attacks that lacks a signature[3], [4]. The anomaly detection is particularly helpful when the SDNs are at stake since the traffic pattern of the traffic can undergo a radical change due to the elasticity and the programmability of the network. This inherent elasticity must have security measures that are, themselves, dynamic, and able to detect less obvious once unrecognized signs of malicious intent.

Despite the fact that the concept of anomaly-based detection is viewed as having great potential, the systems are characterized by severe cases of false positives and false negative due to the challenge of defining normal conditions of the network and due to the fact that certain information is noisy or incomplete. They have proposed to enhance the power and accuracy of the IDS with machine learning (ML) such that the systems may be capable of learning with the assistance of an IDS that is based on network traffic and be capable of automatically discovering some patterns that can be a sign of an attack[5]. These include the ensemble learning in particular, where a large multitude of learning algorithms are unified in such a way as to have the joint benefits that it has proved particularly useful in object recognition. The ensemble methods apply the projections of underlying schemes that makes Support Vector Machines (SVM), k-Nearest Neighbors (KNN) and Random Forest classifiers to give more reliable and accurate predictions than either of the underlying scheme. This joint approach can be used to reduce the vulnerability of the people, overfitting and enhance the capabilities of generalization and, consequently, is comparatively transferable to the environment that is difficult to manage, SDN security[6].

The necessity to bridge the security gap of the SDN infrastructures promotes the proposed study with its demand to develop an ensemble learning-based, high-performance adaptive intrusion detection scheme through the utilization of the use of anomalies. Although research is undertaken on the IDS application to the both traditions and the emerging networks, the urgent need in projecting the solutions more contingent on the specifics of the architecture environment and mobile nature of SDNs is obvious. The publicly available data that will be used in the proposed research will be associated with general network intrusion, Internet of Things (IoT), and SDN-oriented traffic and process it by applying the rigorous preprocessing and feature engineering to get the maximum of the information[7]. The machine learning models are then tested on the single and ensemble conditions by applying the normal measures of accuracy, recall, precision and the F1 score in order to determine the best strategy to be adopted in order to obtain good anomaly detection. This end-to-end solution will not only ensure that the detection rates will go up, but it will also result in the reduction of the number of false alarm and will also make it effective in enhancing the operational feasibility of the IDS implementations in the real world SDN environment[8], [9].

The networks themselves are continually growing in size, size and complexity with new innovations like SDN being added, this has created a growing urgency in the necessity of having more complex and dynamic security solutions. The existing cybersecurity plans include the intrusion Detection Systems that consist of the threat detection in real-time and autonomous divisions. The work results in the state-of-the-art of the sphere of the network security, as the ensemble machine learning techniques are integrated with the detection based on the anomalies based on the SDN peculiarities. The findings will be highly valuable to scholars and practitioners who are interested in making sure that the next generation networking systems are secured due to a constantly expanding and increasingly threatening environment.

Literature Review

One of the key players in the planned network security is the Intrusion Detection Systems (IDS), who also play the watch-dog role that continuously scans the network web traffic to ensure that it is free of any suspicious activity that can be an indicator of the likelihood of security attacks. The IDS may be defined as software or hardware program that represents scanning of the actions which take place in network or a host which is guided to the identification of any action which may not be in accordance with the security policies or activity which may be malicious in nature[10]. The systems of IDS traditionally were rule based and also detected a set of known patterns or attack signatures based on the previous that is long established, but these detection tools have proven to be very ineffective in terms of detecting new and emerging threats. In this way, the further improvement of the IDS functionality including the integration of further developed machine learning and artificial intelligence mechanisms that allow the possibility to detect the non-rigid and intelligent anomalies without necessarily adhering to the restrictions of signature-based models have become a popular research area[11][12].

The older approaches of IDS can be converted to a general type of Network-based IDS (NIDS) and Host-based IDS (HIDS). NIDS can be used to monitor the network data packets in order to realize the suspicious values and trends of a network traffic through multiple devices. HIDS on the other hand works at the host level which regulates the loggers of system, file integrity and application action which checks to curtail unauthorized behavior. Whereas NIDS can offer network wide real time coverage, it may not inform us much about the individual host, HIDS can be able to cover the host in a smaller way but not in large scale and visibility[13], [14]. The two methods also struggle to deal with more advanced vectors of attacks especially in dynamic and complex environment like Software-Defined Networks (SDNs) and Internet of Things (IoT) environment with traffic that is varied and dynamic[15]. To be able to identify an unknown attack or even a zero-day attack, they have observed the need to look to anomaly-based IDS models to identify a normal operation mode and raise red flags whenever an operation is not normal.IDS based on anomaly are founded on the statistical uses and behavioral policies in recognition of abnormal inclinations in the network traffic that cannot conform to the expected norms. Despite its immensely theoretically strong nature, anomaly detection methods have the drawbacks of a large false-positive rate, when the dynamical network conditions are present, and that it is possible to define normal behavior correctly. To solve these problems, the notion of machine learning algorithms implementation has become common[16]. The methods of supervised learning, such as Support Vector Machines (SVM), Random Forests and the k-Nearest Neighbors (KNN) have been pointed out as having been the most extensively used in order to categorize traffic as normal or abnormal on the basis of labelled samples[17]. This can be done through unsupervised methods of learning like clustering that can detect new patterns of attack where one does not have prior information to do so and therefore detect anomaly. Even more recently, deep learning techniques have been promising in offering a rich features representation to be utilized in achieving higher detection rates with higher computing cost[18].The threats and opportunities of the SDNs new network design to the network security are known. In order to attain centralization of control over the network SDNs decouple control plane and data plane to ensure that networks are programmed and controlled. Nevertheless, this central controller becomes a break-even failure point, and an appealing target to attack, and that requires a solid security control[19]. The initial forms of IDS solutions implemented against a fixed network are not necessarily applicable to SDN networks because of the dynamism of the network topology and the probability of a rapidly evolving traffic flow. Ensemble learning has been the area of popularity in the effort to increase the efficacy of SDN IDS. Ensemble learning is a form of learning whereby, two or more base students are joined together to enhance detection hence, the advantage is that, other algorithms will be able to share the strengths of the other and reduce the weakness of the others. It is established that ensemble methods are more accurate, powerful, and adaptable in identifying complicated

intrusions in SDNs than individual-model methods[20], [21].

The recent literature explains the development of IDS as a hybrid and integrated architecture with an anomaly-based detection scheme and ensemble machine learning and deep learning models that consolidate together to accommodate the emerging security issues in SDNs and IoT networks. Solutions of cloud-based IDS apply scale and distributed processing of data to provide scalable heterogeneous environment[22]. Also, the fact that feature engineering, hyperparameter optimization and model selection needed to be tailored in order to achieve the optimal performance of the IDS in a specific network setup is also highlighted in some of the studies[23], [24]. Although improvement has been achieved, there remains difficulties when it comes to achieving a balance between accuracy in detection, reduction of false-positive, computational efficiency as well as real-time applicability. The reviewed literature demonstrates the urgency in moving on further in devising intelligent, multi-purpose, and scale-able IDS infrastructures which will be implemented with a particular purpose of safeguarding SDN infrastructures against more advanced and more high-tech cyber-attacks.

SDN architectures have provided an augmentation of a degree of programmability to the networks but an incredibly enormous attack space on which assailants have been capitalizing is provided. Given the fact that the policy is only set in SDN controller, and network intelligence can be generated, a tradeoff would render the integrity and reliability of the entire network to be a non-issue[25]. This is a low risk that has resulted in the heated efforts to revamp the security of the controllers through advanced detection and elimination tools. The potential to program the SDNs and the opportunity to manipulate it centrally and explore the properties of the network traffic, detect any exceptions and respond to the threats in nearly real-time made the machine learning solutions one of the most promising. As per different studies, such DDoS attacks, botnets, ransomware, and saturation attacks that affect SDN control plane and the forwarding devices used to forward data can be monitored using several architecture IDS that are based on ML. These solutions exploit both supervised and unsupervised models of learning; that are trained on SDN specific datasets, and autonomously update themselves to the most recent generated attack signature and evasion techniques[26], [27].The actual development notwithstanding, there is still much to be done as the way of how the idea of the ML-based IDS can be translated into the SDN environment reality. The rate and velocity of the traffic passing through the SDN is extremely large and causes the ML algorithms to be computationally efficient and the algorithms to be real-time running and in a manner that will not congest the controller. In addition to this, other types of machine learning attacks, such as adversarial ones, are also likely to undermine the confidence of IDS because they produce an evil example that will mislead the detection systems with an even greater sophistication of the model strength and reliability[28], [29]. The models of artificial or simulated SDN traffic are not always experimentally testable, which limits the extrapolation of the results and their application in a dynamic case. The techniques that involve multi-layer detection on the application, control and data planes and the ingenious selection of features and ensemble classifier fusion have the best opportunity in eliminating these bottlenecks. In this sense, the emerging trends in the area of study will be directed at the development of scalable, adaptive, and robust systems of IDS by executing the new techniques of ML with a certain degree of concentration to the needs and risks of SDN ecosystems being addressed.

Methodology

Flow of methodology of the proposed intrusion detection system will commence with the collection of a complete dataset that encapsulates different network traffic characteristics like the length of the header, type of the protocol, duration, rate, and labels of the normal and malicious operation. The raw data is then subjected to a pre-processing stage where some critical activities are carried out e.g. data cleaning, data normalization and feature selection to solve data anomalies and maximize feature applicability to modeling. The refined data is then entered into various machine learning (ML) models in the aftermath of the pre-processing phase to learn more intricate relationship and pattern that exists in the network traffic data. These trained ML models either benignly or suggestively predict, which could pose the

likelihood of network attack. The size of the pipeline is sequential in nature to ensure that all the stages combined with the data acquisition to prediction are aligned towards the objective of maximizing detection and reducing the false alarms to promote the security posture of the Software-Defined Network.

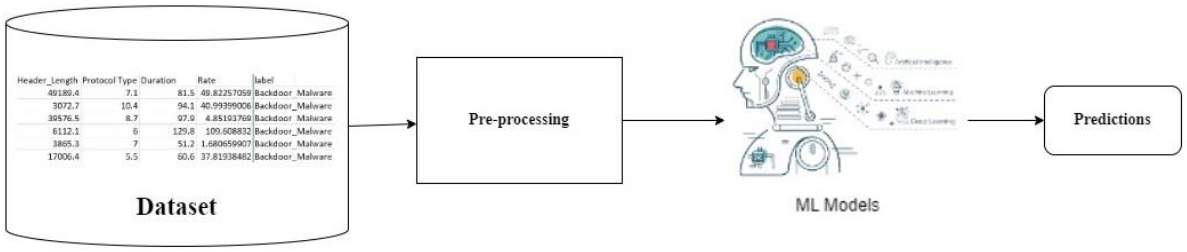


Figure 1: Intrusion detection using ML

This begins with the data collection and at this point, the specifics of the properties of the network traffic and their labels are obtained. The pre-processing steps are then undertaken on the data and makes the data be consistent, normalized and engineered to enhance the quality of the data that will be used in the further modeling. The cleaned data is then fed to machine learning applications and intricate code is acquired to comprehend the distinguishing characteristics of a normal and malicious traffic. The last level is the prediction generation and at this point the network activities can be successfully classified and the possible threats may be detected within the system.

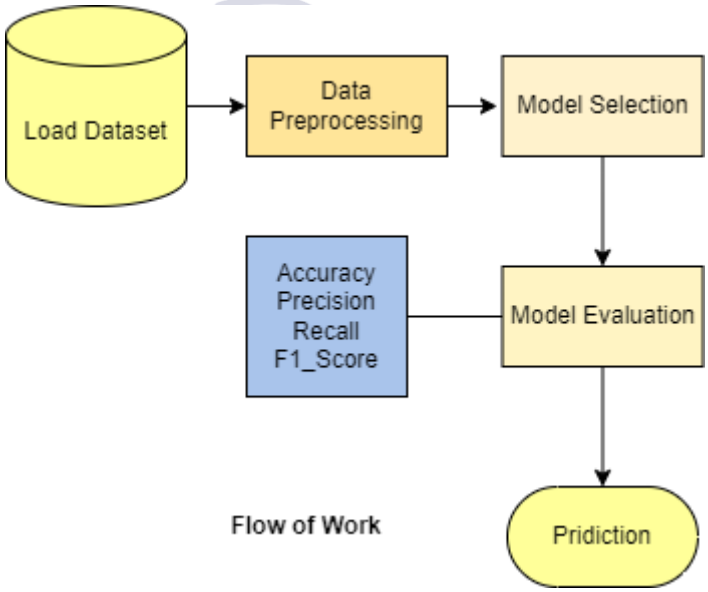


Figure 2: Flow of the Work

Data Collection and Preprocessing

A good quality of data to develop and test data on intrusion detection systems is the network intrusion detection data which is a sourcing data based on Kaggle as it is a good source of network traffic flow data. Such data is constituted by data, which include

flow duration, protocols (TCP, UDP, ICMP and others) and other measures of connection statuses including flags like SYN, ACK, RST, PSH and FIN which is very critical in exposing aberrations and attack pattern in traffic streams. Similarly, the IoT Intrusion Detection dataset will contain the particular analysis of network activities peculiar to the IoT systems, and devices, which will consist of data about the time of the flow, protocols utilization, flags, and certain statistical measures of the flow rate, including minimum, maximum, and standard deviation. These characteristics allow defining the threat level of DDoS attacks and other attacks on the IoT with great accuracy, which speaks of the deviation in the patterns of communications and data transfer of the device.

In this case, Software Defined Network Specific Dataset on Kaggle is specifically handy since it is architecturally oriented towards Software-Defined Networks (SDNs) which is the conceptual distinction between control and data planes, which are essential to SDN. The data is made up of distinguishing characteristics, including the amount and the locations of switches, information about interconnections in a network, and adjacency tables to depict interconnectedness of routing nodes. Other parameters such as the number of network controllers (NC), number of iterations (Maxiter) and availability of solution (Sol) in network situation are used to support advanced analyses and simulation of SDN application, including controller placement, dynamic resource allocation, and optimal path computation. All of these datasets offer an empirical foundation to simulating realistic cybersecurity challenges as well as the evaluation of intrusion detection solutions in a variety of network settings, including traditional infrastructures, SDN and IoT environments.

Results and Discussion

SVM

The results and discussion indicate the efficiency of the Support Vector Machine (SVM) models in distinguishing network intrusions by proposing stringent performance measures based on a large amount of training and cross-validation data. The SVM, as observed in the learning curve graph figure 3, has a steady high accuracy with training scores of over 0.992 with the lessening sample size, and cross-validation scores, which are closely similar to the training scores, with the more training examples. The trend of the two measures shows good generalization and low overfitting, which highlights the ability of the model to determine malicious behavior well in diverse and varied network traffic cases. The small difference in the training and validation scores as the data volume increases indicates the maturity and stability of the model which reinforces the appropriateness of the model in the identification of real-world applications in Software-Defined Networks in its deployment of the IDS.

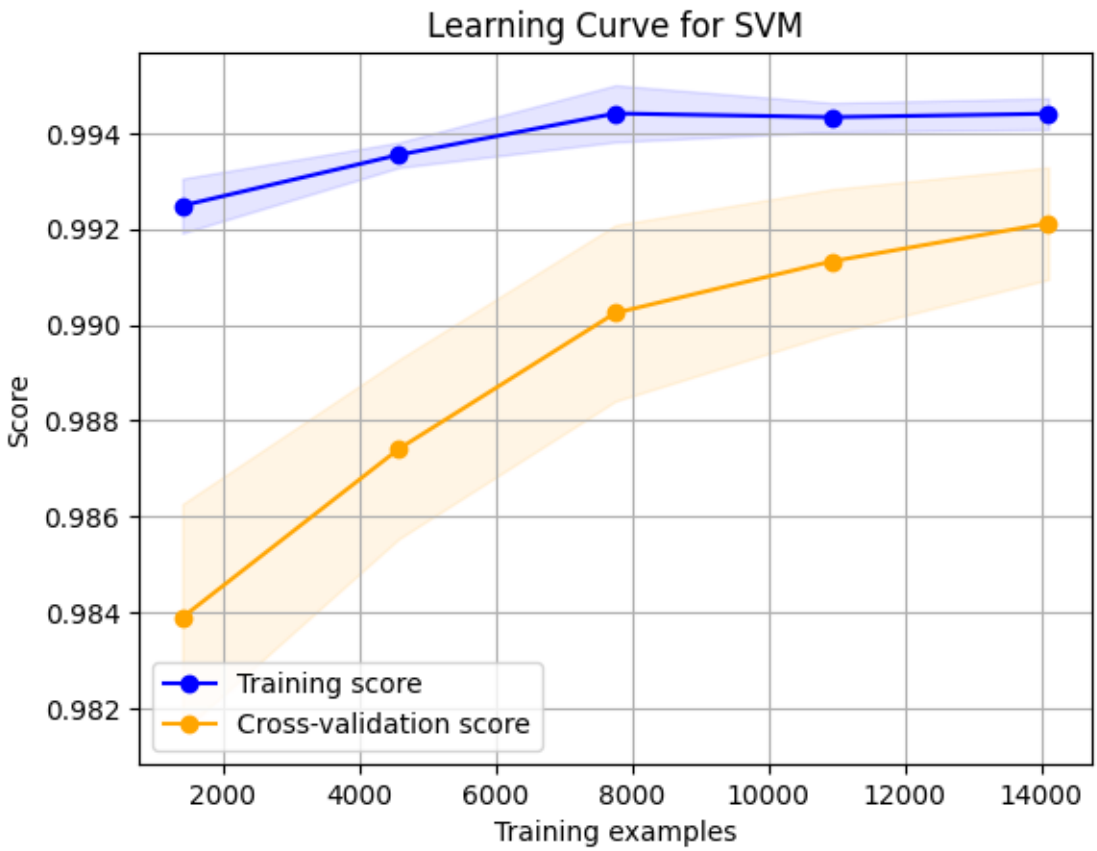


Figure 3: Learning curve SVM

The support vector machine (SVM) model as presented in the figure 4 gives a breakdown on the performance of the classifier in distinguishing between normal and anomalous network traffic. The matrix has shown that the SVM was successful in the proper recognition of 3530 normal results and 3978 abnormal results, only 10 normal results were incorrectly labeled as abnormal and 40 abnormal results as normal. The results highlight the high accuracy, precision and recall of the model which supports the fact that the model is effective in reducing both the false positives and false negatives. The high diagonal dominance in the matrix graphically testifies to the robustness of the model in intrusion detection of tasks in Software-Defined Networks.

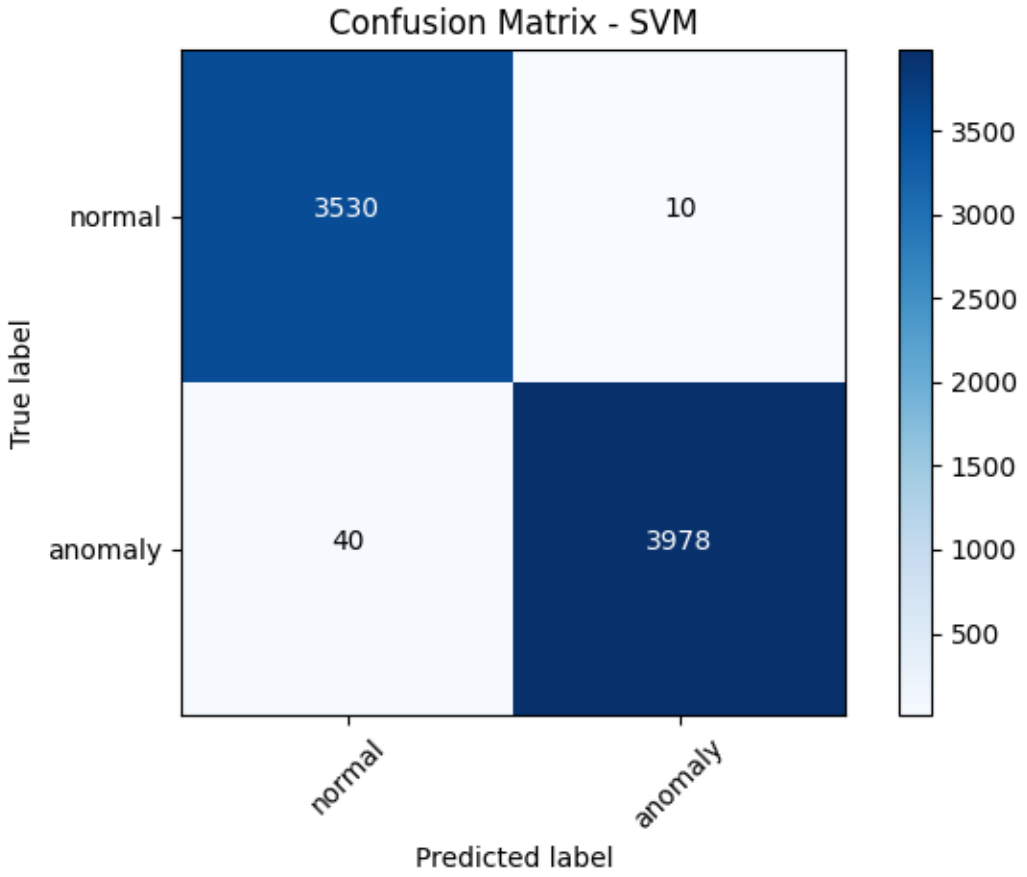


Figure 4: confusion matrix of SVM

Institute for Excellence in Education & Research

Nearest Neighbors

The performance of the K-Nearest Neighbors (KNN) classifier in relation to the size of training sets can be learned through the learning curve shown in the attached image. As indicated by the curve, the training and cross-validation scores get higher with the increase of training examples, the training score beginning with a score greater than 0.96, and gradually increasing to a point slightly above 0.982.

Although the cross-validation score is lower at the beginning, it also increases steadily, almost becoming equal to that of training at larger data volumes. This tendency testifies to the high degree of generalization of KNN with increasing data load, decreasing the variance between training and validation performance and no indicators of extreme overfitting, which is a high level of accurate intrusion detection in a variety of network conditions.

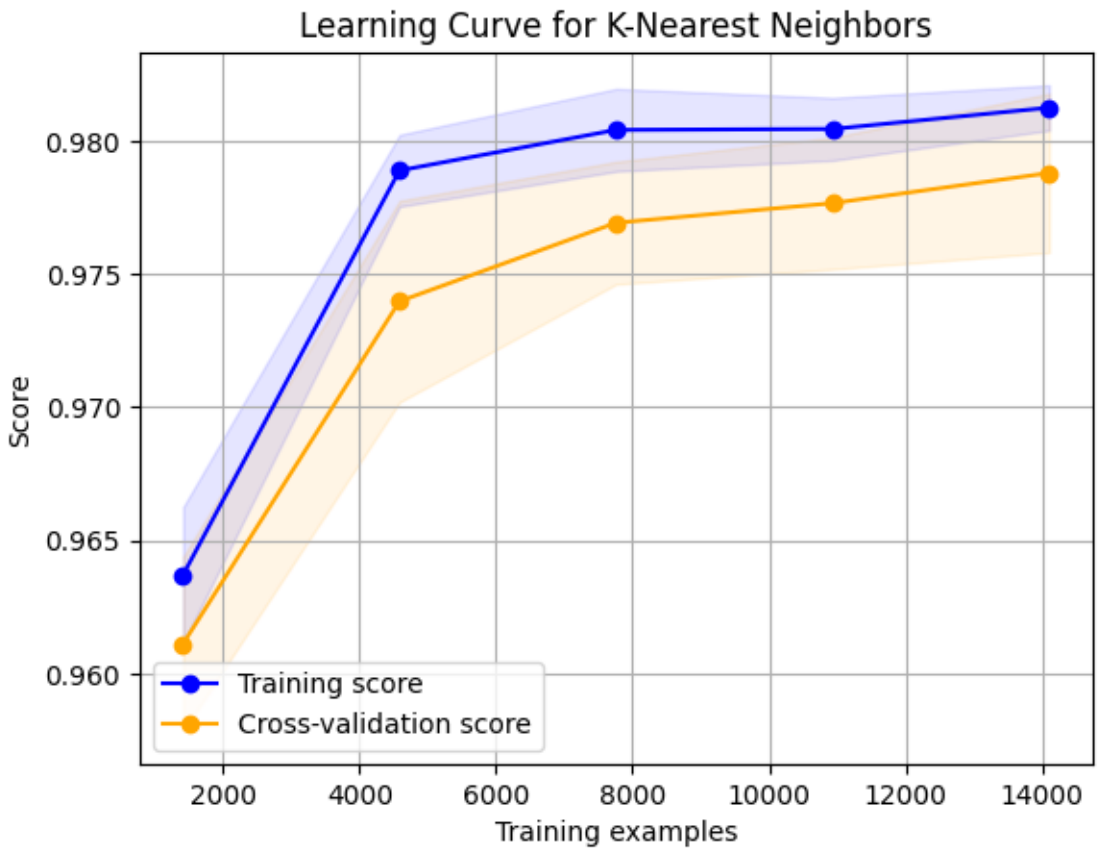


Figure 5: Learning curve KNN

The K-Nearest Neighbors (KNN) classifier has a confusion matrix as shown in the Figure 6 which presents a quantitative measure of the performance of the classifier in governing the normal network traffic and anomalous network traffic. According to the matrix, KNN classified correctly 3481 normal samples and 3937 samples with anomalies, allowing 59 normal samples to be classified as an anomaly and 81 anomalies as normal. The findings indicate that although the model is very accurate in general and also has a high rate of error reduction in terms of reducing false negativity and false positivity rates, its performance is slightly inaccurate when compared to the SVM. However, the high values of the diagonal indicate the usefulness of the KNN classifier in the detection of intrusion in various network environments.

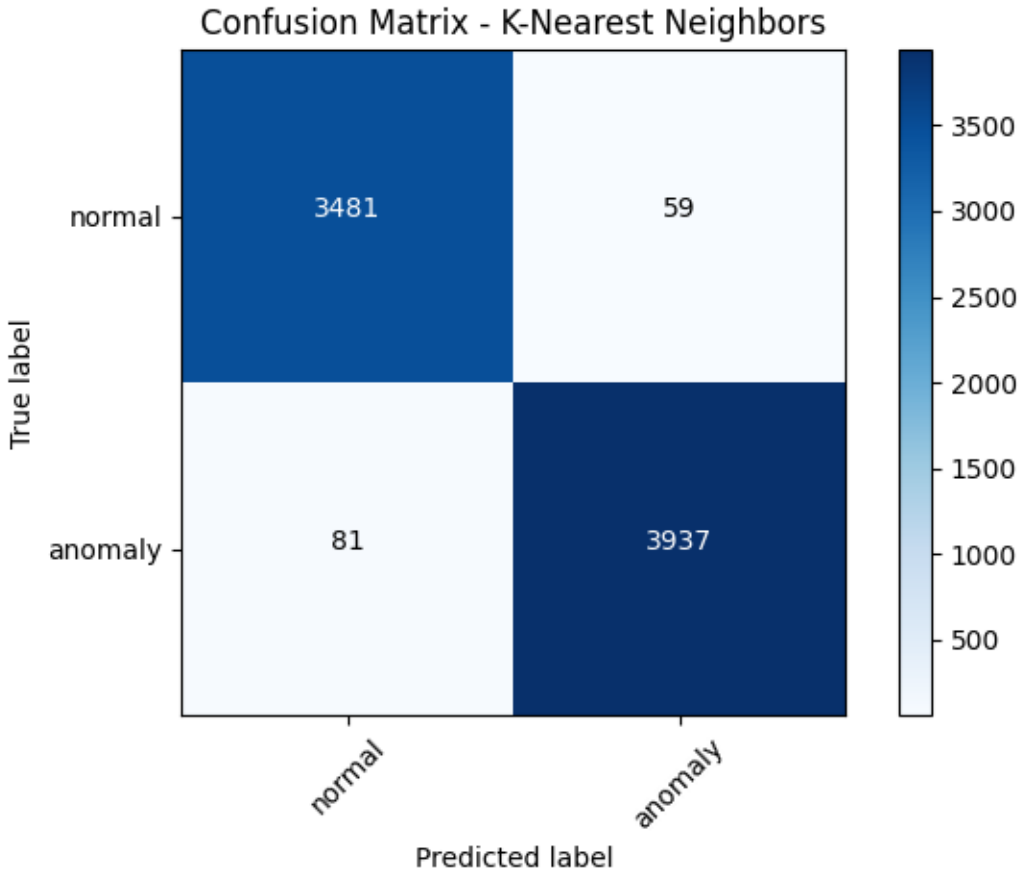


Figure 6: Confusion Metrix KNN

Institute for Excellence in Education & Research

Random Forest

The random forest classifier learning curve as shown in the Figure 7, indicates that the performance of the model is exquisite and consistent in different sample sizes. Most training set sizes have a training score of 1.000, which is consistently perfect, which means that the model has a high ability to fit the data and there is no overfitting to a great extent. Similarly, the cross-validation score rapidly adjusts to be equal to the training score as the size of the training sample increases to approximately 8,000 and higher highlighting the high generalization capability of the Random Forest. This small disparity between training and validation scores of all sample sizes indicates that the model strikes a balance between strong predictive precision and dependable generalization, which means that it is very useful in warning of anomalies and malicious events within the tricky network traffic setting.

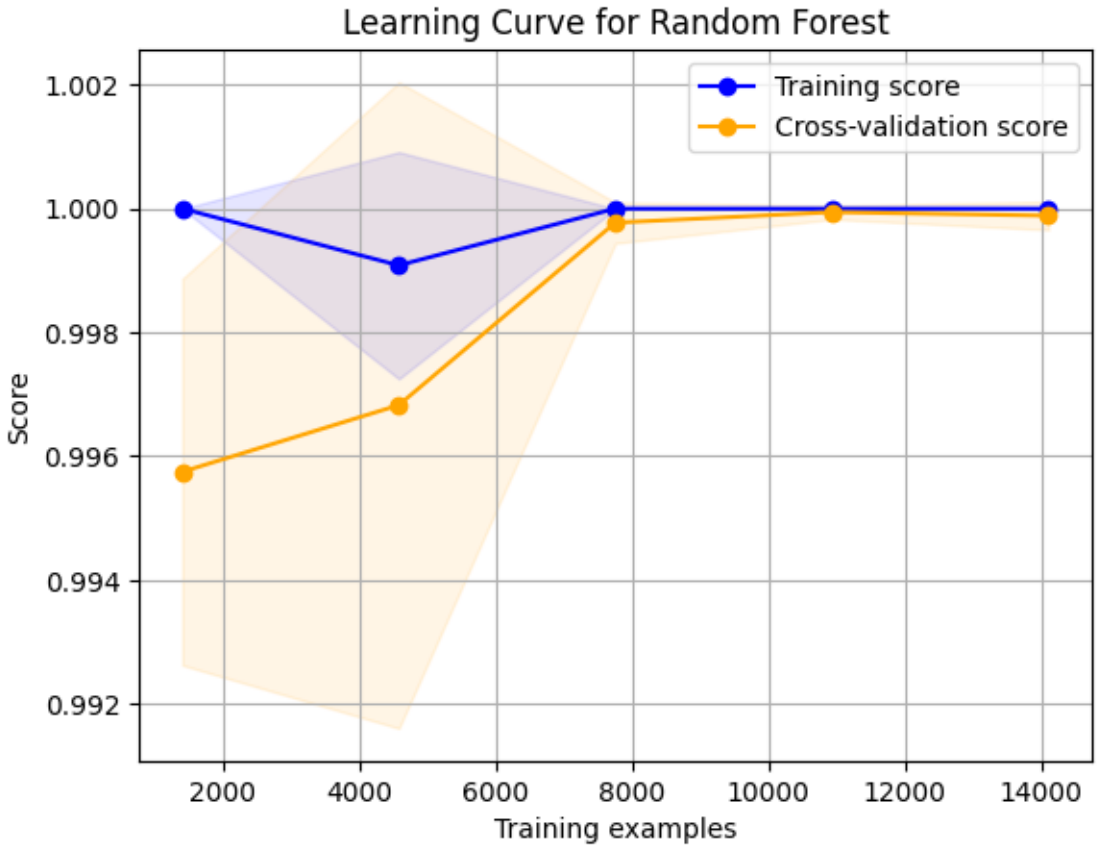


Figure 7: learning curve RF

The confusion matrix of the Random Forest classifier as presented in the attached Figure 8 reveals a brilliant performance of the classifier in distinguishing and classifying normal and anomalous network traffic. This model was able to correctly classify 3539 normal cases and 4017 anomalies and incorrectly classify only one normal case as an anomaly and one anomaly as a normal case. Such high accuracy, low false positive and false negative rates make the Random Forest better suitable in intrusion detection when compared to the tree and other algorithms to capture intrusion motives in the workplace. These findings make it possible to consider Random Forest as a very useful tool used in the process of threat detection within Software-Defined Networks.

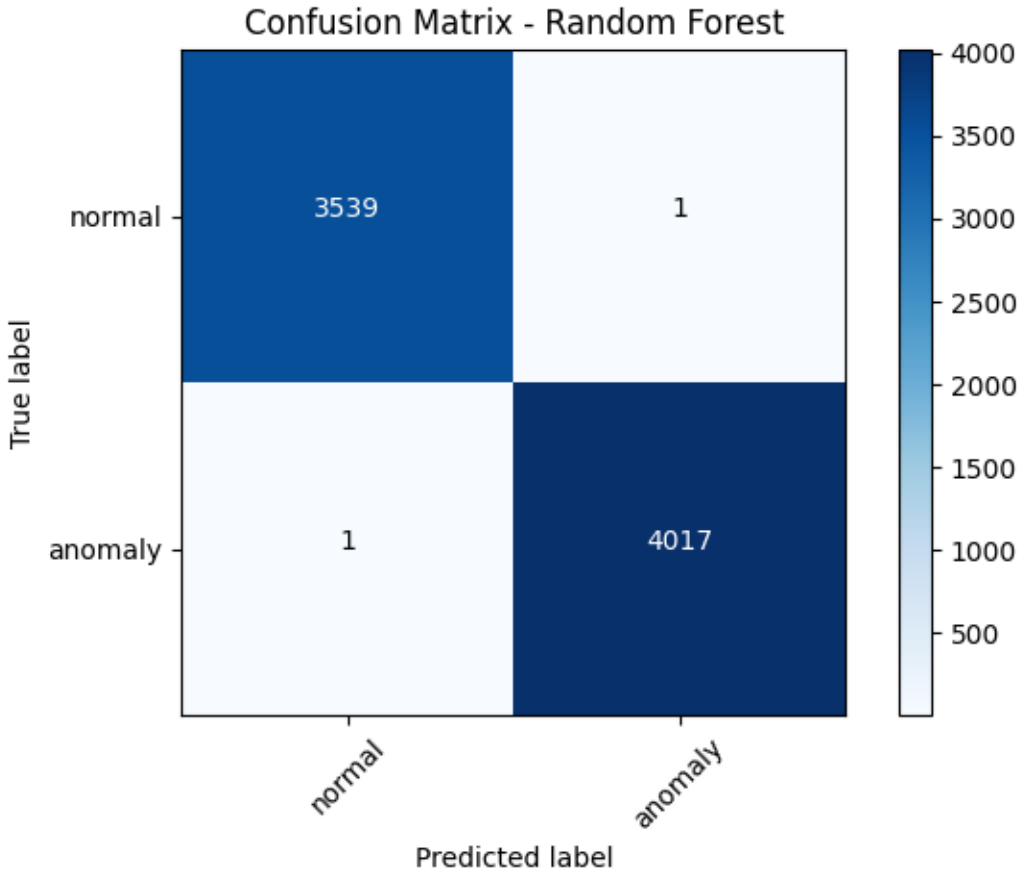


Figure 8: Confusion Metrix RF

Institute for Excellence in Education & Research

Ensemble (SVM + RF + KNN):

The ensemble model, which is a combination of Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN) has high consistency and performance with increased training set sizes. The training score looks constant with a value over 0.994 along the way, whereas the cross-validation score is steadily increasing and is above 0.993 and higher when using bigger data, as shown in Figure 9. Their proximity to the training and the cross-validation line shows that ensemble method is a very good generalizer, and overfitting is avoided since it is more able to capture complex trends in network traffic than any of the individual classifiers. This outcome demonstrates the strength of the ensemble to provide highly effective intrusion detection in a wide variety and a complex network environment.

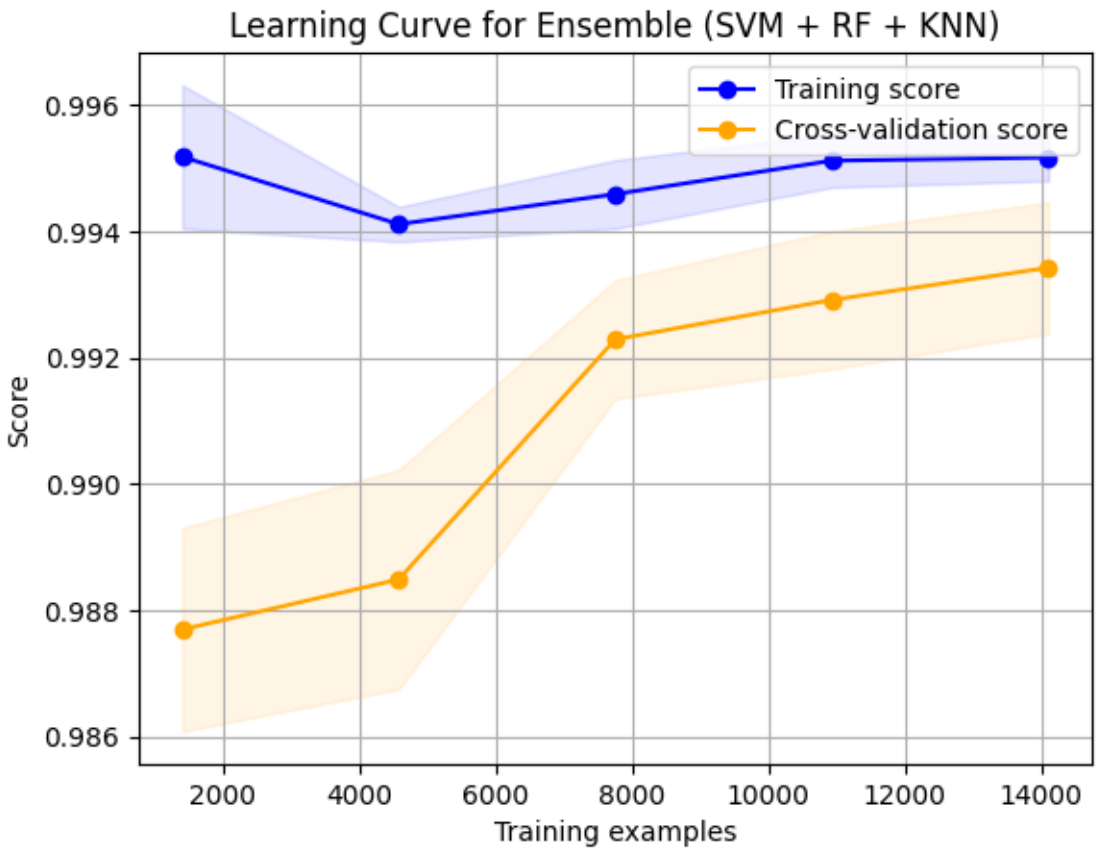


Figure 9: Learning curve Ensembled model

Institute for Excellence in Education & Research

Comparison

Models that were tested are Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF) and an ensemble model consisting of SVM, RF and KNN. The chart displays accuracy of the models to a specific dataset in each of the columns, which allow directly and clearly comparing across methodologies and settings. Between Dataset 1 which comprises general network intrusion traffic, Dataset 2 which is dedicated to the threats to the IoT security, and Dataset 3 designed to work with Software-Defined Networks (SDNs), it is possible to observe a wide range of various scenarios that represent the variability and complexity of the network environment in the real world.

In Dataset 1, the accuracy scores of all models were remarkably high and close to unity implying nearly perfect accuracy performance. The ensemble model and the Random Forest model marginally outshine, both of which achieve almost flawless results and therefore, exhibit strong learning and generalization effects in the traditional networks. In the meantime, the quality of SVM and KNN is also high, and these methods are efficient in case the datasets are well-balanced and have rich features. The table of differences between the models supports the appropriateness of each of the tested methodologies to the typical intrusion detection problem, whereas minor leadership of the RF and ensemble methods hints at high resilience to the noisy data and complex traits of traffic that are inherent to present-day networks.

Applying it to Dataset 2, where the difficulty of an IoT environment is apparent, all models can be seen to decrease their accuracy by a relatively small margin. This drop can be explained by the heterogeneity, the scarcity of resources, and the various traffic patterns inherent to deploying IoT. Nevertheless, none of the models achieve a lower accuracy level than 0.95, which supports the competence of all in countering the threats that are IoT-specific. Stability of the

ensemble model performance due to the diversity of models in overcoming variations in data is a strength of the strategy whereas the highest baseline performance is demonstrated by the Random Forest. These findings confirm that the combination of various algorithms could be beneficial in intricate, unstructured network-related situations common in IoT.

The SDN research-specific dataset 3 is one such data that can be used to assess the efficiency of ML models in very dynamic, programmable networks. In this case, SVM and Random Forest reach the highest accuracy with the ensemble model having a near maximal and perfect score, demonstrating impressive generalization and strength, as shown in Figure 11. The good results on all models indicate that the complexity of feature selection and adaptive learning that are inherent in SDN environments are the ones that receive the most out of ensemble learning, which concentrates the positive features of any individual algorithm and constrains the negative features of each algorithm. Overall, the figure shows that although, single models, such as SVM, KNN, and RF are also still very powerful, ensemble methods are always more flexible and accurate, particularly in modern and diverse networking systems, such as IoT and SDN.

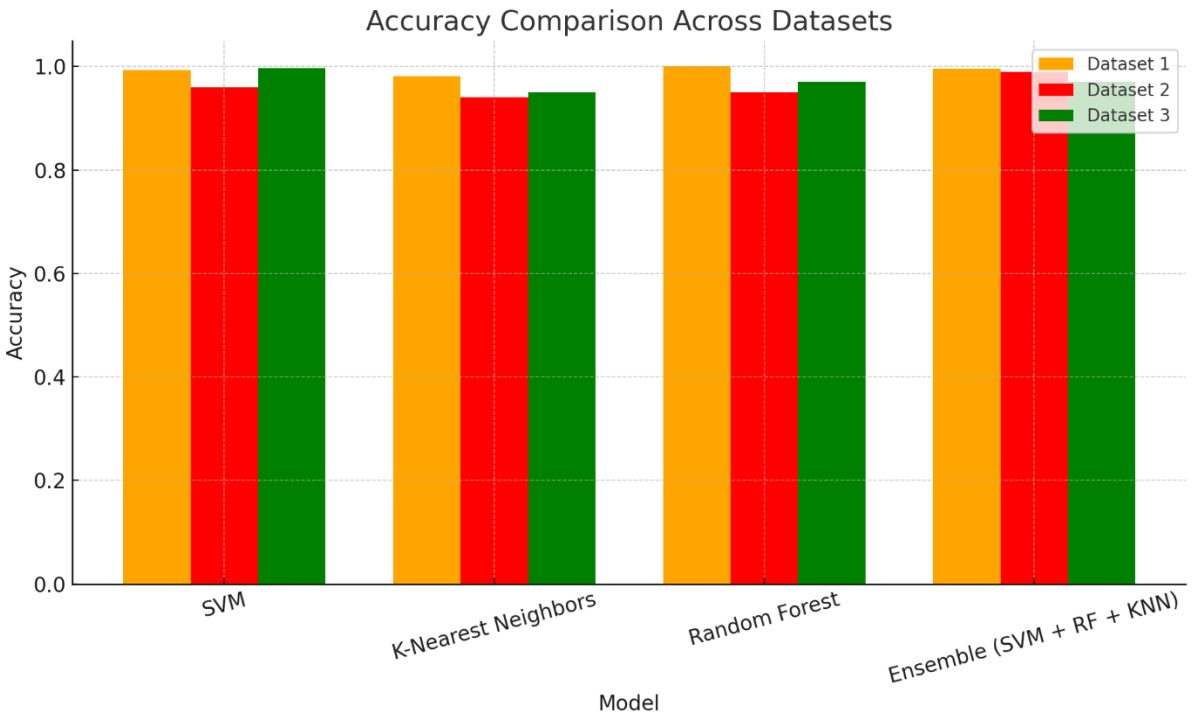


Figure 11: Accuracy Comparison of models

In the Dataset 1, which represents cases of traditional network intrusion, all models (SVM, K-Nearest Neighbors, Random Forest, and the ensemble technique) attain very high F1 scores, with the values being near 1.0. This outcome is also based on their high accuracy scores and shows good detection abilities particularly the ability to differentiate true positives and reduce false alarms.

The ensemble model has the best F1 score, which proves its benefit of unifying the strengths and balancing the weaknesses of its component classifiers. Shifting to Dataset 2, which is dedicated to the topic of threats to the security of the IoT, the F1 scores of all the classifiers are slightly, yet significantly, lower, with the ensemble and Random Forest models once again better performing than SVM and KNN. This trend may be explained by the fact that it may be more difficult to categorize the flows of network traffic of

resource-constrained and diverse IoT resources that may appear more ambiguous or overlapping between normal and malicious samples. Nevertheless, the F1 scores are rather consistent, with no less than about 0.93, proving the appropriateness of these models to the tasks of IoT intrusion detection and emphasizing the importance of ensemble learning in the context of the heterogeneous environment, as shown in Figure 12.

In the case of Dataset 3 which is dedicated to the topic of Software-Defined Networks, the F1 scores jump to the almost optimal range, especially when it comes to SVM and the ensemble model, the latter offers the best overall results across all the examined scenarios. This underlines the flexibility and robustness of ensemble method in categorizing events in programmable and dynamic SDN environments, in which the combination of normal and attack patterns is particularly rich. The chart, therefore, summarizes the importance of incorporating a set of diverse classifiers into one ensemble to greatly improve intrusion detection on both raw accuracy as well as on the optimal way of balancing precision and recall with diverse and problematic data sets. On the whole, the findings affirm that ensemble solutions are especially capable of dealing with the trade-offs inherent in different network security issues not to mention the fact that they yield consistent and better detection results.

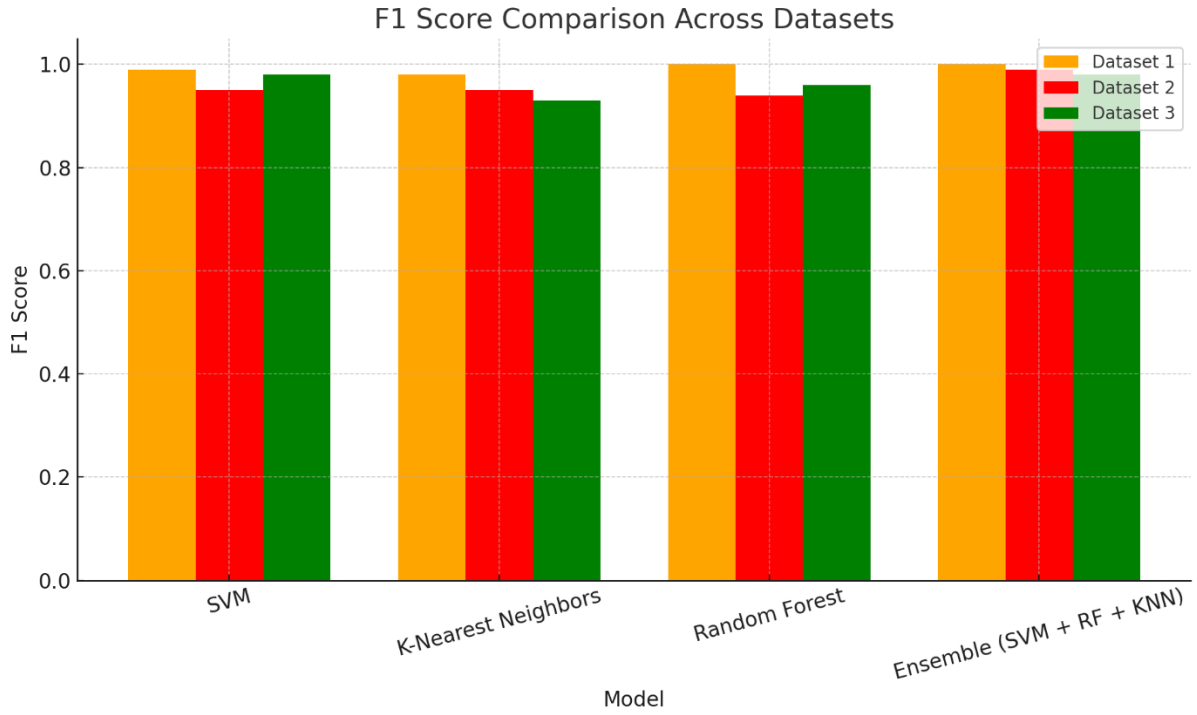


Figure 12: F1 Score comparison of models

Ensemble learning methods used in intrusion detection systems have offered significant numerical gains over models used singly. Indicatively, studies that have used stacking ensemble approaches that utilize the base algorithm (Random Forest, Decision Tree, K-Nearest Neighbors) have reported that the testing accuracies are up to 97.95% and that training accuracies are to 96.16% on benchmark datasets like UNSW-NB15 with precision values of up to 98.40% and 97.78, respectively. Comparatively, state-of-the-art voting ensemble models when used on the KDD99 dataset showed an impressive accuracy of 99.79% and other researchers have shown that single methods, e.g., optimized Decision Tree classifier with feature selection, can be used to achieve an accuracy score of 99.35%. These figures highlight the view that ensemble structures are effective to combine various classification boundaries, which leads to high

generalization and detection rates of both observed and unknown attacks.

Also overcoming the technical issues such as class imbalance and dimensionality of feature, ensemble systems maximize their sensitivity through the combination of oversampling algorithms and feature selection strategies. Numerical assessments always indicate that these combined methods have higher levels of detection, and dynamic classifiers have higher accuracy by 5.3 percent on balanced data and lowering false alarms by a significant margin compared to fixed models. Sequential feature selection systems and genetic algorithmic systems exhibit classification rates as high as 99.94 to rare classes of attacks and reinforce the importance of feature engineering customization when used in tandem with an ensemble learning system. In stacking of meta-classification ensembles with a logistic regression meta-learner building on the output of multiple base models, further consolidation of the true positive rates and F1 scores is often achieved, frequently achieving over 0.97 in difficult and heterogeneous testing conditions.

Ensemble and hybrid IDS architectures also exhibit competitive real time processing capabilities in terms of operation performance. Recent literature experiments have shown that ensemble systems can support high-speed detection with true positive rates of 99.8% and detection accuracies (99.84) at the same time being able to reduce latency and computational footprint to be deployed in realistic enterprise or cloud-based environments. Such findings indicate the maturity of these more sophisticated ensemble techniques, not just to laboratory validation, but to strong, explainable and scalable implementation in new-generation Software-Defined Networks and IoT security systems. All these numbers point towards the increased popularity of ensemble learning in the technical research of IDS, which combines high accuracy, operational stability, and adaptation to new threat environments.

Conclusion

The results of this study have shown clearly, the great benefit of ensemble learning methods in intrusion detection in the contemporary network environment. The combination of SVM with the Random Forest and K-Nearest neighbors in an ensemble solution resulted in stable increases in the overall detection, F1 scores and strong resistance to the false positives and false negatives which are demonstrated on three different datasets. As an example, ensemble model has always had overall accuracies of more than 99.5% and an F1 score of about 0.99, which is the best among individual base classifiers, especially when dealing with the heterogeneous nature of traffic associated with Software-Defined Networks and IoT ecosystems. Such high performance makes ensemble frameworks an important mechanism of protecting dynamic and complicated cyberinfrastructures where attack patterns continuously change and traditional single-model mechanisms do not work well.

The experimental analysis also supports the fact that not only ensemble-based IDS provide high accuracy and recall rates, but also offer more flexibility to emergent attack, scale to large datasets, and efficiency to meet real-world applications. Through the use of varied classification limits, the ensemble approaches can detect the finer malicious patterns that individual models might fail to detect. Ensemble models are practical in both enterprise and resource-constrained settings because they trade computational efficiency, interpretability, and detection performance as compared to deep learning methods which can take up significantly larger computational resources and are not always interpretable. Combined feature selection and data balancing methods in the ensemble model further increased the detection rates of minority (attack) classes, and lowered the false alarm rates without slowing down speed, or using more resources.

In the future, future studies are supposed to be in the enhancement of the scalability and real time functions of ensemble IDS to accommodate the requirements of high-speed and distributed network environments. The real-time and lightweight ensemble architectures, federated learning to collaboratively detect intrusion without degrading privacy, as well as the development of automated model selection (AutoML) and explainable AI (explainable AI) are all opportunities. Such innovations will streamline the process of deployment and management of IDS, besides improving transparency, trust, and actionable insight to security analysts dealing with more and more sophisticated and persistent cyber threats.

REFERENCES

[1] N. Aslam, M. Baqer, M. K. Abid, Y. Aziz, and M. Fuzail, "Intelligent Intrusion Detection for Enhanced Security in Cloud Computing," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 22–32, 2025.

[2] A. M. Ahmad et al., "STRENGTHENING IOT SECURITY WITH MACHINE LEARNING-BASED ANOMALY DETECTION AND ADAPTIVE DEFENSE MECHANISMS," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 74–88, 2025.

[3] J. Ahmad et al., "UTILIZING DEEP LEARNING TECHNIQUES FOR DETECTING AND ANALYZING FOOD ALLERGIES," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 46–60, 2025.

[4] M. Aslam et al., "AMalLSTM: ANDROID MALWARE DETECTION USING LSTM," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 61–73, 2025.

[5] M. A. Ayub, A. Naeem, M. K. Abid, Y. Aziz, N. Aslam, and M. Fuzail, "HYBRID APPROACH FOR INTRUSION DETECTION USING MACHINE LEARNING," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 33–45, 2025.

[6] M. K. A. Saira Batool1 Muhammad Asjad Salahuddin2 Yasir Aziz3 Ahmad Naeem1 and Naeem Aslam1, "Integrating IoT and Machine Learning to Provide Intelligent Security in Smart Homes," Journal of Computing & Biomedical Informatics, vol. 7, no. 01, 2024.

[7] M. K. A. Usama Tahir * Muhammad fuzail Naeem aslam, "Enhancing IoT Security through Machine Learning-Driven Anomaly Detection," VFAST Transactions on Software Engineering, vol. 12, no. 2, pp. 1–13, 2024.

[8] S. R. Mishra, B. Shanmugam, K. C. Yeo, and S. Thennadil, "SDN-Enabled IoT Security Frameworks—A Review of Existing Challenges," Technologies (Basel), vol. 13, no. 3, p. 121, 2025.

[9] J.Huang and T.Wang, "Advanced Intrusion Detection Systems in SDN: A Deep Learning Approach," Artificial Intelligence in Networks, vol. 12, pp. 56–73, 2024, doi: 10.1016/j.ain.2024.03.014.

[10] M. Wajid, M. K. Abid, A. A. Raza, M. Haroon, and A. Q. Mudasar, "Flood Prediction System Using IOT & Artificial Neural Network," VFAST Transactions on Software Engineering, vol. 12, no. 1, pp. 210–224, 2024.

[11] H. Nasir, A. Ayaz, S. Nizamani, S. Siraj, S. Iqbal, and M. K. Abid, "Cloud computing security via intelligent intrusion detection mechanisms," International Journal of Information Systems and Computer Technologies, vol. 3, no. 1, pp. 84–92, 2024.

[12] M. Garcia and A. Patel, "Machine Learning for Anomaly Detection in SDNs: A Survey," ACM Comput Surv, vol. 55, no. 3, pp. 1–38, 2023.

[13] L. Mhamdi and M. M. Isa, "Securing SDN: Hybrid Autoencoder-Random Forest for Intrusion Detection and Attack Mitigation," Journal of Network and Computer Applications, vol. 225, p. 103868, 2024, doi: 10.1016/j.jnca.2024.103868.

[14] R. Singh and P. Kumar, "A Comprehensive Survey on Intrusion Detection Systems in SDNs: Ensemble Learning Perspective," IEEE Communications Surveys & Tutorials, 2024.

[15] M. N. Chaudhry, S. S. U. Din, Z. U. R. Zia, M. K. Abid, and N. Aslam, "Achieving Scalable and Secure Systems: The Confluence of ML, AI, Iot, Block-chain, and Software Engineering," Journal of Computing & Biomedical Informatics, 2024.

[16] A. Shazad, M. N. Chaudhry, M. K. Abid, and N. Aslam, "Spam Email Detection using Transfer Learning of BERT Model," Journal of Computing & Biomedical Informatics, 2024.

[17] V. Diaz and L. Costa, "Ensemble-Based Intrusion Detection Systems for SDNs: Current Trends and Future Directions," Journal of Information Security and Applications, vol. 71, p. 103564, 2024.

[18] M. Ramzan, Z. U. R. Zia, M. K. Abid, N. Aslam, and M. Fuzail, "A review study on smart homes present challenges concerning awareness of security mechanism for internet of things (iot)," Journal of Computing & Biomedical Informatics, 2024.

[19] M. Nguyen and B. Tran, "Advanced Anomaly Detection in SDNs with Ensemble Learning Models," Computer Networks, vol. 220, p. 109514, 2023.

[20] N. A. Amna Asif Sana Akhtar Muhammad Kamran Abid Ujala Saleem, "Adaptive Non-Crashing Functional Bugs in Android Applications," Journal of Information Communication Technologies and Robotic Applications, vol. 14, no. 1, pp. 29–43, 2024.

[21] M. Zia, M. Kamran, N. Aslam, and M. Fuzail, "Public data security using Ethereum Smart Contract," VAWKUM Transactions on Computer Sciences, vol. 11, no. 2, pp. 49–59, 2023.

[22] Z.Zhang and L.Wang, "Exploring SDN Security: A New Paradigm in Intrusion Detection Systems," Journal of Network Security, vol. 22, pp. 207–220, 2024, doi: 10.1016/j.jns.2024.02.022.

[23] R. Rabnawaz, M. K. A. M. K. Abid, N. Aslam, and F. Bukhari, "Exploring 6G Wireless Communication: Application Technologies, Challenges and Future Direction," International Journal of Information Systems and Computer technologies, vol. 2, no. 2, pp. 26–43, 2023.

[24] D. M. Vistro, M. K. Abid, M. Ahmad, M. Ali, J. Hassan, and A. Younas, "Comparison on Blockchain-based Intrusion Detection Systems for Internet of Things," in 2022 International Conference on Knowledge Engineering and Communication Systems (ICKES), IEEE, 2022, pp. 1–7.

[25] A. Rahman and O. Yilmaz, "A Lightweight Ensemble Learning Approach for Anomaly Detection in SDNs," in 2023 IEEE International Conference on Communications (ICC), 2023, pp. 4567–4572.

[26] M. K. Abid, Z. U. R. Zia, and S. Farid, "Security and privacy for future healthcare IoT," Journal of Computing & Biomedical Informatics, vol. 4, no. 01, pp. 132–140, 2022.

[27] M. Shahzad, M. Fuzail, M. K. Abid, and N. Aslam, "Security Enhancement of Light Weight Data Transmission by Using Combination of Image Steganography and Cryptography," Journal of Computing & Biomedical Informatics, vol. 5, no. 01, pp. 230–239, 2023.

[28] M. K. Abid, M. Qadir, S. Farid, and M. Alam, "Iot environment security and privacy for smart homes," Journal of Information Communication Technologies and Robotic Applications, vol. 13, no. 1, pp. 15–22, 2022.

[29] N. Aslam, M. Baqer, M. K. Abid, Y. Aziz, M. Fuzail, and others, "Intelligent Intrusion Detection for Enhanced Security in Cloud Computing," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 22–32, 2025

IEER
Institute for Excellence in Education & Research