Vol. 2 No. 4 (2024)

Spectrum of Engineering Sciences

SPECTRUM OF ENGINEERING SCIENCES

Online ISSN
3007-3138
Print ISSN
3007-312X

A Survey of Software-Defined Networks Based on Advance Machine Learning Based Techniques

Fazeel Mahmood[1]

Department of Computer Science, Faculty of Computer Science & IT Superior, University Lahore, Pakistan. fazeelgujjar14@gamil.com

Muhammad Shehroz[2]

Department of Computer Science, Faculty of Computer Science & IT Superior, University Lahore, Pakistan. Muhammadshehroz741@gmail.com

Zoha Ansari[3]

Department of Computer Science, Faculty of Computer Science & IT Superior, University Lahore, Pakistan. zohaansari0930@gmail.com

Zain-ul-Abideen[4]

Department of Computer Science, Faculty of Computer Science & IT Superior, University Lahore, Pakistan. zainulabideensuperior@gmail.com

Faizan Rauf[5]

Department of Computer Science, Faculty of Computer Science & IT Superior, University Lahore, Pakistan. faizanrauf2525@gmail.com

Abdul Hanan[6]

Department of Computer Science, Faculty of Computer Science & IT Superior , University Lahore, Pakistan. hnananjutt@gmail.com

M. Masab[7]

Department of Computer Science, Faculty of Computer Science & IT Superior, University Lahore, Pakistan. jatoimasab@gmail.com

Abstract

Presently, networks have opportunities to build systems with improved and intelligent solutions which can be easily tailored for different users. In conventional networking, the software defined networking (SDN) work separately for its control plane and data plane, which makes it better in managing, more secure and affordable. Since ML is the most fundamental branch of AI when amalgamated with SDN it will provide an efficiency and effectiveness in the management of resources such as bandwidth mapping, flow control, error control and security on the network. In this paper you will find the use of network alongside ML implemented on SDN concepts in two ways. First it describes how appropriate ML algorithms are integrated with SDN based networks following assessment. On the other hand, it provides reasonable recommendations for various network applications based on SDN.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

And toward the end, it discusses the extra development needed for Machine Learning (ML) algorithms and SDN concepts. The common point of AI, Big data, computer networking and similar fields is discussed in this paper. Researchers from different professions and ages have their findings with regards to AI for various uses since it is a young and intricate area. Thus, this paper will assist these researchers to point out these main faults more accurately.

Keywords: Artificial intelligence, machine learning, network management, software-defined networking.

Introduction

Today's reality in the development of technologies and the interconnection of systems require better management of traffic and efficient handling of devices where there're networks that are complex and different. Further ahead the main achievement for network technology is to develop, systems that are, scalable, and personalized. Scholars are currently considering ways of implementing this goal through intent based networks (IBNs) or intent driven network (IDNs) which seek to translate business intent into network configuration settings. For this approach to work especially well the network setting requires flexibility to accommodate various user needs and varying quantities of data and traffic [1].

Another common strategy utility settlement, of IBNs includes the utilization of AI powered Software Defined Networking (AI SDNs). In this setup AI and machine learning technologies are used therefore for analyzing data and describing user's objectives or goals, mapping these objectives into specific network operation. In this case, therefore, software such as SDN's controllers coordinate the creation process allowing the network to respond accordingly. What this means is that for prevalence, maintenance and control of networks it is not sufficient to have the right hardware. You also require sophisticated instruments and innovations to develop network performance and cut operational costs AI intervenes play an essential part to meet these requirements because of its learning feature AI can analyze different data to optimize network efficiency and conserve computing energy Thus, attains sophisticated service management [2].

Besides, to that AI proved to be very useful as an application for network security because can support identifying of the attacks' patterns and

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

an overall improvement of security. Machine learning is a component of artificial intelligence on which it depends to predict and manage network resources with regard to datasets. The truly flexible nature concerning this technology presents possibilities for analyzation of user activity and enhances the Internet network. Although, AI and machine learning, as work, in progress in network strategies, there is an even more promising advancement which can facilitate these innovations through better network experiences dubbed Software Defined Networking (SDN). Software installed content defines Software Defined Networking (SDn) as free from limitations of hardware systems. Improves the maneuverability and address space of the control plane addressing tasks of routing and power within devices such as routers and switches [3].

It could be seen that the integration of AI and SDNs has affected some aspects of network management. For example, Desktop Network (SDNs) features a policy-based control of resources such as computing power, data storing and communications management. In addition, machine learning enhances the control of network traffic through the integrated view in the SDN to optimise traffic control and meet QoS requirements. Besides, to that's relevant to network security, there are the specifics of SDNs and machine learning potentiality that enables recognizing threats, such as DDoS attacks and unauthorized access [4].

Needs for network enhancements today arise from technologies and gadgets; due to multifaceted multimedia applications currently in use. The last years of investigation has shown the developments taken place in the integration of the concept of ML with that of SDNs. For instance, a specific study explains the use of ML algorithms in Software Defined Networks SDNs. Provides findings. Building on this research entails integrating views from both the ML and SDNs to get a feel of the two technologies. We categorize machine learning algorithms. Describe their relevance, in software defined networking (SDN) providing a blueprint, for those who are familiar with both kinds of research [5]. A research study on technology driven networks reveals the synergy of Software Defined Networking that is known as SDGNFVD meant for enhancing decision making structures that enhance network flexibility consistent with our objective of sharing information with a number of researchers.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

More specifically this article discusses the integration between machine learning and software defined networks hence the acronym, SDN. To meet these of objectives, likely developments in this field have been forecasted and the resultant studies summarized to equip the researchers with enhanced understanding of the challenges and milestones defining the domain. From a perspective elaborated in Section II, we discuss SDNs. Evaluate, in section III of this paper, modern applications built with the aid of ML. Finally in the last section V we talk of challenges. Conclude with important points discussed in Section V.

Leveraging Machine Learning for Enhancing SDN-Based Network Systems

Since the very basic processes in ML are usually identified with terms such as "Train" and "Predict," many researchers believe that they closely reflect human cognitive processes, such as "Induce" and "Predict." This comparison is visualized as shown in figure 1, where ML principles are meant to mirror some aspects of human learning and adaptation. Instead of being extremely complicated and abstract, ML concepts reflect natural processes that humans use for reasoning and growth: We follow an inductive and synthetic but not a formal deductive approach [6].

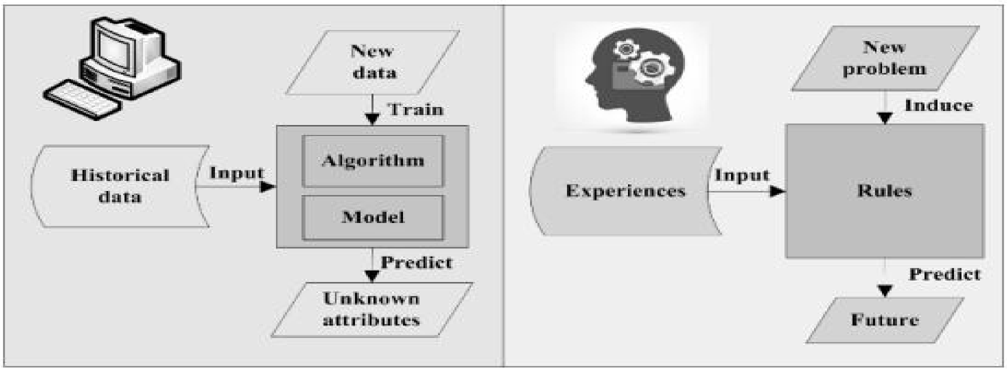

Figure 1. Comparison of ML processes with human cognitive functions [6].

There are different types of machine learning methods, depending on the task at hand. Very broadly put, the ML models can be categorized, at the high level itself, into regression models, classification models and structured learning models. What are Regression Models Regression models, also referred to as prediction models, are best suited for providing numeric outputs. In comparison, classification models are concerned with classifying data into classes and can be subdivided into binary and multi-class

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

classifications. Binary classifcation, for instance spam classifcation in which emails are classifed either as spam or as not spam, are widely used. The extension of this is for multi-class classification where data is classified in more than one class like document classification, where class will always belong to a single group [7]. Structured learning models differ from the other two in that their outputs can vary in complexity and may produce sequences or structures rather than fixed-length values. An example of a structured learning output could be the textual description generated in image-based semantic analysis, where ML is used to interpret and describe the content of an image in natural language.

Methods within machine learning also come under categorization as to how they're trained: four main types generally will suffice-supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Probably the most commonly used approach is supervised learning: this involves training a model on labeled data to teach it to recognize patterns associated with specific outcomes. Meanwhile, unsupervised learning work with unlabeled data-finding structures and patterns out, without predefined labels. Finally, semi-supervised learning acts as a middle ground when combining both labeled and unlabeled data. It quite often comes in handy in processing large datasets where it will be impossible to label all the data points. Lastly, reinforcement learning stands out by focusing on how agents learn how to interact with their environment in order to maximize a cumulative reward and, therefore, is particularly relevant to applications in game theory, control theory, and multi-agent systems [8].

ML algorithms are useful to a great extent due to their flexibleness and wide applicability areas, especially in scenarios related to classification and prediction tasks wherein consistent accurate results have always been delivered. Focusing particularly on software-defined networked systems, this work will examine in a wide variety of methods considered classical in ML applicable within this specific network type concerning the effectiveness and variety of applications [9]. To better understand the scenario, we have in Table 1 presented an exhaustive view of ML algorithm application along with their performance within the context of SDNs. Moreover, all of the abbreviations that first appeared in our tables are presented in alphabetical order in Table 2.

Supervised learning over networks based on the SDN framework

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

Supervised learning methods draw functional conclusions from fine-tuning the classifiers' parameters on a data set with labeled categories. It is widely used within numerous applications, such as speech recognition, spam filtering, and object detection. The primary objective of supervised learning is to predict values of one or more output variables based on a vector of input variables.

All classification and regression algorithms are supervised learning. The two categories differ on the basis of the type of output variables. Regression methods are used to generate quantitative outputs, which are also known as continuous variable predictions. Classification methods are used to generate qualitative outputs, which are also referred to as discrete variable predictions. Input variables can be either continuous or discrete, and the classifier itself is derived from the data that it analyzes [10]. When discussing regression algorithms in the context of SDN, they can be employed to forecast the response time for executing queries based on traffic flow within the SDN architecture. For instance, multiple linear regression models can establish a significant relationship between key performance indicators (KPIs) of applications and various network metrics. However, the use of regression algorithms in SDN remains relatively uncommon at this stage. Therefore, this paper will primarily focus on exploring classification algorithms within the framework of SDN-based networks [11].

The main function of classification algorithms is to categorize data accurately into appropriate groups. Some of the most widely used classification algorithms include K-Nearest Neighbors (KNN), Logistic Regression, Support Vector Machines (SVM), Decision Trees, and Naive Bayes.

Table 1. Utilization of SDN-Based Networks and Assessment of Machine Learning Algorithm Performance [12]

Algorithm Category	**Algorithm**	**Application**	**Performance Analysis**
Supervised Learning	KNN	Predicts multiple types of attacks, including DDoS.	Easy to implement with high accuracy.
	SVM	Predicts link failures and classifies	Not sensitive to outliers and

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

		packets.	calculates features easily.
	Decision Trees	Facilitates inductive inference and flow classification.	Suitable for multiclass classifications but may be time-consuming with large datasets.
	Ensemble Learning (including variants)	Enhances accuracy in recognizing attacks and reduces false alarm rates.	Combines strengths of various algorithms, improving overall performance.
Unsupervised Learning	SVM + K-Means	Detects anomalies and classifies data flows.	High accuracy but less effective in complex environments.
	C4.5 + K-Means	Applied in traffic classification and security policy definition.	Errors may increase rapidly with too many categories.
	DT-MCP-PCM	Aids in discovering flow rules and solves controller placement challenges.	High-speed operation with easy data preparation.
Integration of Supervised and Unsupervised Learning	Laplacian SVM	Used in various network management tasks.	Provides efficient data processing, although typically tested in controlled settings.
Semi-Supervised Learning	RL (Reinforcement Learning)	Supports cognitive network management and adaptive	Used for similar applications as supervised learning; however,

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

		multimedia traffic control.	it's more complex.
Reinforcement Learning	DRL (Deep Reinforcement Learning)	Aims to approximate the Q-value action function for dynamic resource management.	Promotes resilience and scalability in network operations.
Deep Q Learning	Deep Q Learning	Utilized for optimizing resource orchestration.	High accuracy of inference with very low false-positive rates.

K-Nearest Neighbors (KNN) in SDN Networks

KNN is a simple algorithm of classification, which shows the way data points are connected with each other based on how far apart their feature values are. The intuitive central idea of KNN is simple: if the majority of the K closest samples in a feature space fall into a certain category, then a new sample will probably fit into the same category. This is intuitive and effective because it depends on a few nearest neighbors. One of the most important benefits of KNN is that it is simple to implement. In addition, it is sensitive to outliers and usually attains high accuracy, which makes it a very popular choice for multiclass classification tasks in various fields [13].

However, there are some disadvantages of KNN. The implementation of this algorithm relies on a process called linear scanning, wherein it calculates the distance of the test data from each instance in the training set. It then sorts these distances to determine the nearest K samples. As such, when training sets are large, it takes quite a long time. Despite being one of the most basic machine learning algorithms, KNN's efficiency in implementation and feature calculation keeps it popular among users. However, its performance can be challenged when dealing with vast datasets, leading to significant processing times [14].

Literature Review

SVM is a strong generalized linear classifier that is known to classify data into two classes in a supervised learning setup. The core of SVM contains a decision boundary referred to as the maximum margin hyperplane, which can

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

optimize the separation between various classes of data points. SVM is stable since it tries to solve an optimization problem that minimizes both empirical risk and structural risk. It is worth mentioning that SVM is designed only for the binary classification. Still, when facing multiple classification challenges, these can be transformed into several binary problems, enabling SVM to maintain its effectiveness. Since SVM may be pretty effective for certain applications of network security, numerous studies have proved that. For example, in one study, it utilized the integration of SVM and a network controller for the purpose of DDoS attacks detection. This makes a system capable of distinguishing benign flow entries from regular traffic by those malicious entries associated with DDoS attack traffic [15].

In brief, SVM is highly appreciated regarding stability and low false alarm rate for two-class problems in classification. The mechanism taken by SVM heavily minimizes the time in terms of attack detection and even attack classification. Lastly, with deployment at the level of the SDN controller, the complexity of SVM contributes minimally to the inefficiency of the entire system of SDN.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

The Role of Decision Trees (DT) in SDN-Based Networks

A decision tree is a strong predictive model that represents the relationship of different object properties with its values. The structure of a decision tree is similar to an actual tree; at every internal node, it represents an object, a branch is taken as each possible attribute value, and each leaf node represents some specific category. In the data mining technique, very often decision trees are employed for both analytical and prediction purposes. The primary usage of decision trees in networks is in packet classification [16].

Inductive inference is used by a lot of algorithms in a decision tree. For example, an algorithm known as C4.5 decision tree can be used in classifying packets with high volume flow and can therefore manage it efficiently. In module Real-time Detection Strategy, the choice of n-tuple feature focuses on whether it can have a powerful classification system that can therefore determine whether a certain n-tuple is considered to be an elephant flow or not. Comaneci and Dobre have employed C4.5 decision tree classifiers that serve as pre-trained models for the different types of traffic. They enhanced that with the features like the inter-packet arrival time, packet size, the number of packets, or flow tuples to make its accuracy high. Another classifier that has been developed is Least Cost Disruptive (LCD) that has been proposed for an optimization of the ASP Data Plane. This model takes into account trade-offs between good service provision, adaptation costs, and disruption to users. More recently, research also explores the application of Decision Trees in solving FTCP [17].

Harnessing Ensemble Learning in SDN-Based Networks

The supervised learning algorithms are developed to achieve stability in the model so that it can always give superior performance even when situations become uncertain. Ensemble learning further takes it one step forward by combining several weak supervised models to produce a more powerful and reliable model. The process starts with building a set of n individual learners, which is then combined with specific strategies [18]. Whereas ensemble methods are not as widely used as the traditional technique, they have proven surprisingly successful in many applications. For instance, whereas DT is used as a base learner, the bagging-based RF model has been surprisingly very efficient in quite a few occasions. Indeed, one of the studies, for example, revealed RF-based cross-validation had impressed the world with indoor

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

localisation accuracy at 98.3%, greater compared to other algorithms like KNN, SVM, NN, etc.

Furthermore, Lei et al. proposed a regression approach that was designed for the purpose of correctly modeling the latency distribution of a VNF, and which proved the versatility of ensemble learning in solving various problems [19]. It represents the fact that most approaches to ensemble learning are very much more accurate than a traditional approach, but quite naturally much more complex; this will be an important tradeoff in choosing a right kind of machine-learning strategy applicable to specific applications within a given SDN network [20].

Insights into Supervised Learning in SDN-Concept Networks

In our analysis, we examined the performance of different supervised methods in SDN and listed our findings in Table 3. Ensemble learning techniques, bagging, boosting, and AdaBoost, are usually stronger than traditional techniques like KNN, Neural Networks (NN), and SVM, which are dependent on a single classifier. Besides, although these contemporary techniques improve the performance of the system, they do not necessarily make it train faster compared to traditional techniques.

TABLE 3. Prediction Accuracy for Different Supervised Machine Learning Algorithms [21]

Application	Dataset	KNN	SVM	Naive-Bayes	Decision Tree	Others
Predicting Network Attack Patterns	Public dataset from the "Long Tail"	86.19%	90%	94%	88.52%	
Detecting DDoS Attacks	Real-time dataset	90%	94%	90%	94%	K-medoids: 88%, K-means: 86%

Exploring Unsupervised Learning in Software-Defined Networks

Unsupervised learning simply means digging deep in data without any prelabeled labels. In this way, one can dig deeper in patterns and relationships

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

that would never be found using any other approach. The results will thus help to direct the study to be undertaken next. Clustering probably is the most popular type of unsupervised learning; however, one specific type stands out- the K-means algorithm.

Clustering means you simply group your data points in some way, which is referred to as a "cluster", without having any information of how your data was created. In real-life implementation, for example, an unsupervised algorithm would help one detect anomalies on different layers of a network such that the problems would be detected earlier before they would become critical and where such problems actually originate from [22]. The standard K-means algorithm has not been seen to be widely applied to SDN. However, other variants are increasingly used in SDN. Among such methods, the hierarchical variant of K-means was up to now applied successfully in solving several controller placement challenges in WANs with proved better balance than optimized K-means. More related adaptations appear as heuristic ones combining K-means and Dijkstra's algorithm with approaches based on the cooperative game theory, oriented on specific challenges from the domain of SDN controller placement.

There is a growing necessity to compare supervised learning methods with unsupervised ones, which would clarify the strength and weakness differences. There are studies that concluded the algorithm Naive Bayes and KNN has a capacity to classify normal or abnormal network traffic. This implies that K-means and K-medoids will always take less time for training, but at certain points in time, these will lose accuracy. Scientists have been studying how to intermingle supervised learning models such as SVM with some of the unsupervised algorithms, like K-means, to improve their ability in traffic classification [23].

These two learning methods are often combined in real-world deployments on large data platforms to leverage their strengths. For example, one system uses machine learning to classify traffic flows and use these classifications to inform high-level SDN policies that simplify complexity. The approach has been demonstrated with promising results, including high accuracy rates for normal traffic and good performance even when abnormal cases are encountered. The new generation of clustering algorithms that are proposed that combine K-means with decision-tree techniques improve user

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

experience about cloud services. Application management and data flows are optimized while meeting the expectations about the quality of service by the users. Finally, unsupervised learning thrives in analyzing large datasets where, instead of using some predetermined labels, the extracted information and rules are extracted directly from the data. Flexibility and intelligence make it such an important area of research within artificial intelligence and Software-Defined Networks [24].

The Role of Semi-Supervised Learning in Enhancing SDN Performance

Traditional machine learning can be broadly categorized into two types: supervised and unsupervised learning. Supervised learning is based on labeled datasets, where each data point is tagged with the correct output. On the other hand, unsupervised learning deals with unlabeled data, which tries to find patterns without predefined categories. However, in many real-world situations, acquiring labeled data is both difficult and costly, while unlabeled data is often plentiful. This challenge has led to very rapid development techniques of semi-supervised learning that effectively combine both labeled and unlabeled samples.

The approach of semi-supervised learning combines the best of both worlds. This approach requires a small amount of labeled data and utilizes a large amount of unlabeled data for training and classification purposes. This can be quite useful in applications that are similar to supervised learning applications, such as traffic classification and anomaly detection in networks [25].

Although semi-supervised learning is applied with synthetic data and controlled lab settings, nowadays, there is a high interest in real-world application of these techniques. As an example, some have applied semi-supervised learning to classify internet traffic well, which provides evidence that this technique may be applicable in practice. The data flows can then be handled by analyzing the QoS parameters, typically large in size, commonly referred to as "elephant" flows, to optimally make use of available resources.

Another area where semi-supervised learning benefits QoS classifiers is when it has to deal with traffic coming from unknown applications. This is an area of flexibility as networks are now able to adapt and respond to changing data types, thus bettering the performance of networks. Though it holds great promise, its practical application is not explored much; this calls for further

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

research. It will be best understood if one investigates the application of semi-supervised learning in real-world environments, so it can then be harnessed for the betterment of network performance and reliability [26].

Advancements in Reinforcement Learning for Software-Defined Networking

Reinforcement learning is a very interesting concept wherein an agent learns from interactions with its environment through a process called trial and error. This agent receives a reward after every interaction in terms of how effective those interactions were. The prime goal of a reinforcement learning system is to modify parameters in a way that optimizes these rewards. There are no clear instructions coming from the environment; feedback helps it evaluate how successful the actions are.

Our research work underlines the pivotal role that RL performs to augment resilience and scalability in the Software-Defined Networking environments. It outperforms in tasks related to path selection and optimization of routes. For example, the DROM algorithm revolves around delay reduction and high throughput boosting and shows a good number of improvements on performance as well as the stability of the networks. On another front, the SDCoR study stands out by providing adaptive optimal routing policies based on real-time sensing and learning from the Internet of Vehicles (IoV) compared to traditional IoV protocols [27].

Such factors as high jitter need to be dealt with to keep quality in place within networks. Techniques were developed to limit the number of alternative paths taken for each successive data frame. As an illustration, one research found that jitter remained under 40 ms if high-loss routes were avoided compared with more conventional techniques.

Recent advances in RL explore exciting collaborations with other technologies. Researchers have integrated random neural networks with RL to identify the optimal overlay path and still keep monitoring overheads minimal. Another innovation in use is the SRSA mechanism, which makes scaling decisions automatically using RL, a characteristic that's particularly beneficial to dynamic environments [28]. Deep Reinforcement Learning combines the strengths of deep learning and RL, thus significantly speeding up the learning process and increasing the efficiency of the algorithms of RL. DRL has already been gaining much attention in both theoretical research and practical

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

applications. AlphaGo is one example that Google DeepMind came up with and marked the historic moment in artificial intelligence.

DRL can be proven to adapt SDN for multimedia traffic control directly without necessarily having to employ complicated mathematical models for it to handle the multimedia streams. Deep Q-Learning techniques are mainly applied when handling different problems under a variety of network settings. For instance, an integrated DQL framework has been proposed to enhance the performance of eco-friendly heterogeneous wireless networks. Similarly, a dueling DQL method has been proposed to increase the throughput in blockchain systems by optimizing the trust features of blockchain nodes and controllers [29].

Insights into Machine Learning Techniques in SDN Concept Networks

In Section II, we have discussed promising applications of four key ML algorithms within the SDN environment. The ML algorithms discussed in the section are supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Results indicate that supervised learning algorithms are the most developed and effective algorithms in the SDN context since they create the best models from labeled existing datasets, which may, therefore, make the confident predictions and classifications.

On the other hand, unsupervised learning is different. It doesn't rely on labeled data; it works with data as it is. The best example of this would be clustering, where items are grouped together purely on their characteristics. This helps us find patterns and relationships within the data, even without predefined categories [30]

Network Applications in SDN Using Machine Learning Approaches

Section II tackles Machine Learning (ML) tools able to help SDN-based architectures. This approach is to sort ML algorithms systematically, which would help understand the different categories in depth and their effects on SDN. This post will aim to make clear what the benefits are in practice and how machine learning could potentially help SDN move onto the next level. We follow that up with some more concrete use cases within the SDN system, showing us how Machine Learning can push the envelope [31].

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

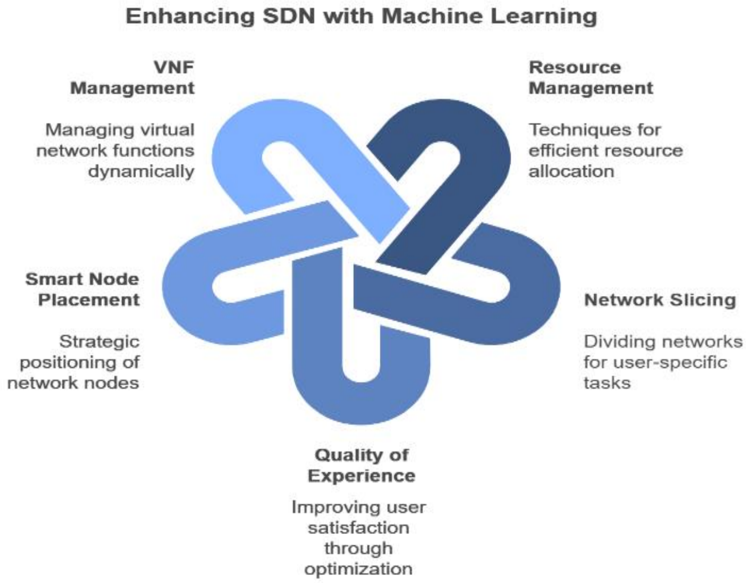

FIGURE 2. Enhancing SDN with Machine Learning

The methods of NIDS are generally divided into following two categories: 1.Misuse Detection. 2.Anomaly Detection However the misuse detection produces ideal results and unable to find new attacks, while anomaly detection is better in finding unknown attack but not fully mature among other technologies used by IDS [32].

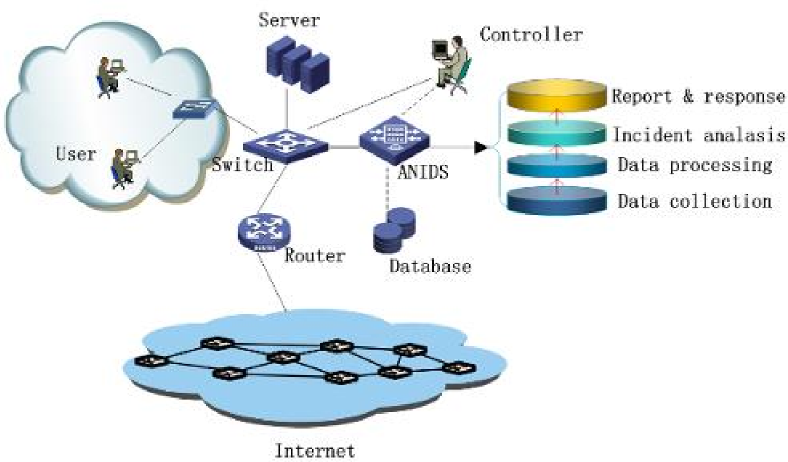

FIGURE 3. Network intrusion detection

Detection of abnormality within the network is achieved by recognizing pattern from well behaved networking activities as logged through log analysis. When the patterns show there are a notable number of deviations for this, then the more likely it is to be an anomaly that could pose potential threats. Detection operates by first learning the patterns for normal behavior

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

during a training phase, and second testing new data against the norms. When such deviation goes beyond a certain limit – it becomes an anomaly [33].

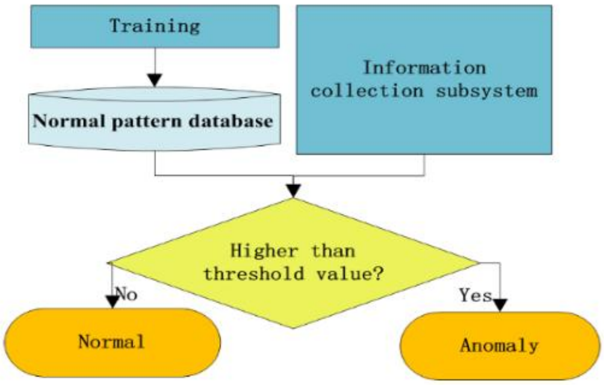

FIGURE 4. The flow of network anomaly detection.

FIGURE 4. The Flow of network anomaly

SDN: While a double-edged sword that brings with it vulnerabilities thanks to its open design, the separation of control and data planes means new security solutions can be added for even more enhanced security. The advent of recent SDN-based systems has enabled the rapid and accurate detection of intrusions, with most research focusing on low-overhead implementations using flow-based approaches an ML method [34].

Future ML in Future SDN

ML has evolved through three major stages. In the 1980s, connectionism was the trend, marked by the development of Perceptron's and Neural Networks (NN). The 1990s saw statistical learning methods like Support Vector Machines (SVM) take the lead. In the 21st century, Deep Neural Networks (DNN) emerged, bringing connectionism back to the forefront. The rapid growth of data and computing power has led to the maturation of many AI applications based on deep learning. Numerous ML methods are used, with the choice depending on data characteristics, training goals, and specific scenarios. In SDN-concept networks, understanding the unique aspects of each scenario is crucial for achieving optimal performance [35].

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

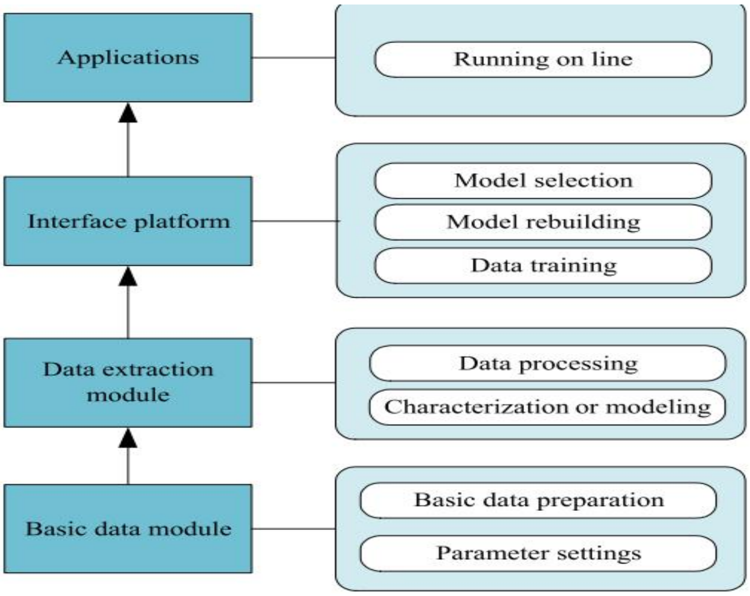

FIGURE 5. The abstract hierarchical structure of ML

An abstract hierarchical structure of ML is depicted in Figure 5. It starts with a basic data module (parameter settings and data preparation), followed by a data collection module (data processing and modeling), and ends with an interface tool and consumer-facing services. Appropriate methods are needed at every layer to implement these intelligent applications.

Challenges with Datasets

In Fig. 5, basic data is shown at the foundational layer. Access to high-quality datasets is a significant challenge in ML, as noted in [36] Challenges include data sources, labeling, class balance, and more. Ignoring data source characteristics during training can result in poor testing performance, unrepresentative of real systems.

Raw data often has issues such as missing attributes, unlabeled data, excessive or insufficient attributes, unseparated test and validation data, and category imbalances. Datasets determine the upper limit of ML results. They should represent the network architecture they correspond to, making high-quality data extraction a future research priority.

Problem Representation Challenge

Defining a problem involves understanding the issue you want to solve. Without sufficient analysis, finding a suitable method or model can be overwhelming. Early misdirection leads to invalid results, such as using a classification algorithm for a clustering problem.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

Overconfidence in powerful algorithms can be misleading. Sometimes, simpler methods like logistic regression outperform complex ones like SVM. Understanding the problem's nature is crucial.

Mathematical problem representation is intuitive and aids in solution finding, transforming practical problems into mathematical ones and aiding in feature engineering and model training, both time-consuming [37].

Challenges in Model Building and Optimization Once a problem is defined, a model is created to describe the objective world, abstracted from data. In data analysis, we often start with just the data and aim to derive rules from it, where the rule represents the model, we want.

Different models can produce varying results, even with the same algorithm. For example, with polynomial regression, we can create multiple hypothetical functions for a dataset and choose the best fit. However, building a model doesn't guarantee success. Factors like the number of features, sample size, and regularization parameters affect the final outcome. We need standards to determine that one model is superior. This involves comparing and selecting models based on criteria. Different strategies and perspectives can produce varied optimized results, leading to multiple models and solutions for the same problem [38].

Algorithm Selection Issues We have outlined several ML algorithms applicable in SDN-concept networks. ML involves building models implemented through algorithms. Challenges include understanding foundational models, recognizing practical problems suitable for the model, and applying the model to real-world situations. It's crucial to understand the nature of a method, its application scenarios, conditions, and limitations.

Choosing the right ML algorithm depends on the situation. Complex algorithms should be reserved for special cases. The ML algorithm cheat sheet helps in selecting suitable algorithms for specific problems [39-45].

Issues in ML Framework Innovation To facilitate the development and implementation of AI applications, many ML frameworks have emerged. As discussed in Section III-D, some research focuses on system frameworks that provide developers with shortcuts; some emphasize usability, while others focus on deployment or parameter optimization.

With the rise and evolution of SDN-concept architecture, ML-based algorithms and frameworks have seen continuous enhancement. Researchers

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

and engineers globally are encouraged to innovate, share, and even integrate large-scale new algorithms in this field. These algorithms are becoming versatile, capable of various combinations and extensions, and are used to construct general frameworks suitable for diverse applications. This means that multiple ML algorithms can be integrated into a more robust framework to analyze data more effectively and extract maximum value from it [46].

Discussion of Future ML Issues In this section, we summarize future issues needing more attention concerning the ML process. Though we outline five key problems, many practical aspects are involved. For instance, with datasets, all related processes, such as data collection, feature extraction, sample balancing, and anomaly handling, must be considered. A correct understanding of these issues will encourage researchers with ML or AI backgrounds to delve deeper into models or algorithms, enhancing the practicality and effectiveness of ML algorithms in the future [47].

The Future of SDN-Concept Networks with ML The increasing demand for various efficient and large-scale applications has set higher standards for future network architectures in terms of performance, flexibility, and controllability. However, innovation in networking has been relatively slow and inefficient. Traditional networks have struggled to support these applications, and without resolving these network issues, efficiency may not improve and could even decline [48].

SDN has become a prominent research topic worldwide. Initially, SDN specifically referred to networks using the OpenFlow South Bound Interface (SBI). Today, it refers to generalized networks supporting multiple SBIs (such as NETConf, OVSDB, BGPLS, PCEP), enabling flexible programming and intelligent analysis beyond traditional routing protocols [49, 50].

Conclusion

With the increase, in data and the progress of machine learning (ML) systems driving it forward ML has become a touted AI resource, for facilitating self-governing network operations and management by utilizing its ability to extract valuable information from datasets. Software Defined Networking (SDN) an aspect of networking has demonstrated its durability and flexibility especially as networks and large-scale datasets adapt to fulfill more intricate service requirements. Though certain surveys touch upon the topics and obstacles associated with implementing machine learning, in SD WAN

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

networks the development of solutions that facilitate autonomous network administration is currently, at a nascent phase.

This research also explores areas, for study in this field and highlights the key obstacles faced when utilizing machine learning efficiently within networks. What is often seen is that while machine learning has advanced considerably it encounters challenges in network applications due to data patterns and the scarcity of high-quality training data. This can impede machine learning programs from reaching their capabilities. We anticipate that this conversation will act as a roadmap, for enhancing Software Defined Networking (SDN) and integrating network solutions.

SDN, enhanced by ML methods, holds considerable promise for future network development and management, including intelligent routing, resource management, flow control, network security, and more. In the future, we aim to conduct deeper research into the key challenges identified in this paper, contributing to more robust and adaptable network infrastructures for the next generation.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Y. A. Khan, U. Khalil, H. Khan, A. Uddin, S. Ahmed, "Power flow control by unified power flow controller",sss Engineering, Technology & Applied Science Research., vol. 9, no. 2, pp. 3900-3904, Feb. 2019

[2] K. Benzekki, A. El Fergougui, and A. E. Elalaoui, Software-de ned networking (SDN): A survey, Secur. Commun. Netw., vol. 9, no. 18, pp. 58035833, 2016.

[3] Ahmad, S., Lavin, A., Purdy, S., & Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data.Neurocomputing,262, 134-147.

[4] Akhi, A. B., Kanon, E. J., Kabir, A., &Banu, A. (2019).Network Intrusion Classification Employing Machine Learning: A Survey. (Doctoral dissertation)Department of Computer Science and Engineering, United International University, Bangladesh.

[5] Alizadeh, H., Khoshrou, A., &Zuquete, A. (2015). Traffic classification and verification using unsupervised learning of Gaussian Mixture Models. In2015 IEEE international workshop on measurements & networking (M&N). 1-6. IEEE.

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X
Khan, I. Uddin, A. Ali, M. Husain, "An Optimal DPM Based Energy-Aware Task Scheduling for Performance Enhancement in Embedded MPSoC", Computers, Materials & Continua., vol. 74, no. 1, pp. 2097-2113, Sep. 2023
[6] S. Khan, I. Ullah, H. Khan, F. U. Rahman, M. U. Rahman, M. A. Saleem, A. Ullah, "Green synthesis of AgNPs from leaves extract of Salvia Sclarea, their characterization, antibacterial activity, and catalytic reduction ability", Zeitschrift für Physikalische Chemie., vol. 238, no. 5, pp. 931-947, May. 2024
[7] H. Khan, M. U. Hashmi, Z. Khan, R. Ahmad, "Offline Earliest Deadline first Scheduling based Technique for Optimization of Energy using STORM in Homogeneous Multi-core Systems", IJCSNS Int. J. Comput. Sci. Netw. Secur., vol. 18, no. 12, pp. 125-130, Dec. 2018
[8] Shehreyar Nawaz, Hamayun Khan, Wajiha Salman, Umer Shahid, Momin Latif Khokhar, M Zaid Iqbal, & Abdullah Hamid. (2024). A Survey on Latest Trends and Technologies of Computer Systems Network. Spectrum of Engineering Sciences, 2(4), 85–114
[9] Waleed, A. Ali, S. Tariq, G. Mustafa, H. Sarwar, S. Saif, I. Uddin, "An Efficient Artificial Intelligence (AI) and Internet of Things (IoT's) Based MEAN Stack Technology Applications", Bulletin of Business and Economics (BBE)., vol. 13, no. 2, pp. 200-206, July. 2024
[10] Noor, A. Ilyas, Z. Javaid, H. Khan, "Framing a Knowledge Domain Visualization on Green Human Resource Management: A Bibliometric Analysis from 2008-2022", Pakistan Journal of Humanities and Social Sciences., vol. 11, no. 4, pp. 4200-4212, Aug. 2023
[11] Muhammad Waleed Khawar, Hamayun Khan, Wajiha Salman, Samra Shaheen, Ariba Shakil, Fatima Iftikhar, & Khawaja Muhammad Ismail Faisal. (2024). Investigating the Most Effective AI/ML-Based Strategies for Predictive Network Maintenance to Minimize Downtime and Enhance Service Reliability. Spectrum of Engineering Sciences, 2(4), 115–132
[12] M. Gondal, Z. Hameed, M. U. Shah, H. Khan, "Cavitation phenomenon and its effects in Francis turbines and amassed adeptness of hydel power plant", In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), IEEE., pp. 1-9, Mar. 2019
[13] U. Hashmi, S. A. ZeeshanNajam, "Thermal-Aware Real-Time Task Schedulabilty test for Energy and Power System Optimization using

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

Homogeneous Cache Hierarchy of Multi-core Systems", Journal of Mechanics of Continua and Mathematical Sciences., vol. 14, no. 4, pp. 442-452, Mar. 2023

[14] Y. A. Khan, F. Khan, H. Khan, S. Ahmed, M. Ahmad, "Design and Analysis of Maximum Power Point Tracking (MPPT) Controller for PV System", Journal of Mechanics of Continua and Mathematical Sciences., vol. 14, no. 1, pp. 276-288, May. 2019

[15] Jawad Ahmad, Hamayun Khan, Wajiha Salman, Muzamal Amin, Zain Ali, & Shumail Shokat. (2024). A Survey on Enhanced Approaches for Cyber Security Challenges Based on Deep Fake Technology in Computing Networks. Spectrum of Engineering Sciences, 2(4), 133–149

[16] Khan, A. Yasmeen, S. Jan, U. Hashmi, "ENHANCED RESOURCE LEVELING INDYNAMIC POWER MANAGEMENT TECHNIQUEOF IMPROVEMENT IN PERFORMANCE FOR MULTI-CORE PROCESSORS" ,Journal of Mechanics of Continua and Mathematical Sciences., vol. 6, no. 14, pp 956-972, Sep. 2019

[17] Abdul Rafay, Hamayun Khan, Wajiha Salman, Gulzar Yahya, & Uzair Malik. (2024). SD Network based on Machine Learning: An Overview of Applications and Solutions. Spectrum of Engineering Sciences, 2(4), 150–165

[18] Khan, Q. Bashir, M. U. Hashmi, "Scheduling based energy optimization technique in multiprocessor embedded systems", In 2018 International Conference on Engineering and Emerging Technologies (ICEET), IEEE., pp. 1-8, Sep. 2018

[19] Y. A. Khan, M. Ibrahim, M. Ali, H. Khan, E. Mustafa, "Cost Benefit Based Analytical Study of Automatic Meter Reading (AMR) and Blind Meter Reading (BMR) used by PESCO (WAPDA)", In 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), IEEE., pp. 1-7, Aug. 2020

[20] Naz, H. Khan, I. Ud Din, A. Ali, and M. Husain, “An Efficient Optimization System for Early Breast Cancer Diagnosis based on Internet of Medical Things and Deep Learning”, Eng. Technol. Appl. Sci. Res., vol. 14, no. 4, pp. 15957–15962, Aug. 2024

[21] Khan, I. Ullah, M. U. Rahman, H. Khan, A. B. Shah, R. H. Althomali, M. M. Rahman, "Inorganic-polymer composite electrolytes: basics, fabrications, challenges and future perspectives", Reviews in Inorganic Chemistry., vol. 44, no. 3, pp. 1-2, Jan. 2024

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

[22] Khan, M. U. Hashmi, Z. Khan, R. Ahmad, "Offline Earliest Deadline first Scheduling based Technique for Optimization of Energy using STORM in Homogeneous Multi-core Systems", IJCSNS Int. J. Comput. Sci. Netw. Secur., vol. 18, no. 12, pp. 125-130, Oct. 2018

[23] M. S. Abdul Razak, S. P. A. Gothandapani, N. Kamal, and K. Chellappan, "Presenting the Secure Collapsible Makerspace with Biometric Authentication", Eng. Technol. Appl. Sci. Res., vol. 14, no. 1, pp. 12880–12886, Feb. 2024.

[24] H. Khan, M. U. Hashmi, Z. Khan, R. Ahmad, A. Saleem, "Performance Evaluation for Secure DES-Algorithm Based Authentication & Counter Measures for Internet Mobile Host Protocol", IJCSNS Int. J. Comput. Sci. Netw. Secur., vol. 18, no. 12, pp. 181-185, July. 2018

[25] Amangele, P., Reed, M. J., Al-Naday, M., Thomos, N., & Nowak, M. (2019). Hierarchical Machine Learning for IoT Anomaly Detection in SDN. In2019 International Conference on Information Technologies (InfoTech). 1-4. IEEE.

[26] Anderson, J. P. (1980). Computer security threat monitoring and surveillance.Technical Report, Fort Washington, PA, James P. Anderson Co.

[27] Aung, Y. Y., & Min, M. M (2018). An analysis of K-means algorithm based network intrusion detection system. Advances in Science, Technology and Engineering Systems Journal,3(1), 496-501. Bauer, F. C., Muir, D. R., &Indiveri, G (2019). Real-Time Ultra-Low Power ECG Anomaly Detection

[28] Bhati, B. S., Rai, C. S., Balamurugan, B., & Al-Turjman, F. (2020). An intrusion detection scheme based on the ensemble of discriminant classifiers.Computers & Electrical Engineering,86, 106742.

[29] Bhattacharyya, D. K., &Kalita, J. K. (2013).Network anomaly detection: A machine learning perspective. CRCPress.

[30] Khan, A. Ali, S. Alshmrany, "Energy-Efficient Scheduling Based on Task Migration Policy Using DPM for Homogeneous MPSoCs", Computers, Materials & Continua., vol. 74, no. 1, pp. 965-981, Apr. 2023

[31] Shah, S. Ahmed, K. Saeed, M. Junaid, H. Khan, "Penetration testing active reconnaissance phase–optimized port scanning with nmap tool", In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), IEEE., pp. 1-6, Nov. 2019

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

[32] Y. A. Khan, "A GSM based Resource Allocation technique to control Autonomous Robotic Glove for Spinal Cord Implant paralysed Patients using Flex Sensors", Sukkur IBA Journal of Emerging Technologies., vol. 3, no. 2, pp. 13-23, Feb. 2020

[33] Y. A. Khan, "A high state of modular transistor on a 105 kW HVPS for X-rays tomography Applications", Sukkur IBA Journal of Emerging Technologies., vol. 2, no. 2, pp. 1-6, Jun. 2019

[34] Khan, S. Ahmad, N. Saleem, M. U. Hashmi, Q. Bashir, "Scheduling Based Dynamic Power Management Technique for offline Optimization of Energy in Multi Core Processors", Int. J. Sci. Eng. Res., vol. 9, no. 12, pp. 6-10, Dec. 2018

[35] Ibrahim Akmal, Hamayun Khan, Ayesha Khushnood, Fatima Zulfiqar, & Eman Shahbaz. (2024). An Efficient Artificial Intelligence (AI) and Blockchain-Based Security Strategies for Enhancing the Protection of Low-Power IoT Devices in 5G Networks. Spectrum of Engineering Sciences, 2(3)

[36] Y. A. Khan, "Enhancing Energy Efficiency in Temperature Controlled Dynamic Scheduling Technique for Multi Processing System on Chip", Sukkur IBA Journal of Emerging Technologies., vol. 2, no. 2, pp. 46-53, Jan. 2019

[37] Khan, K. Janjua, A. Sikandar, M. W. Qazi, Z. Hameed, "An Efficient Scheduling based cloud computing technique using virtual Machine Resource Allocation for efficient resource utilization of Servers", In 2020 International Conference on Engineering and Emerging Technologies (ICEET), IEEE., pp. 1-7, Apr. 2020

[38] Hassan, H. Khan, I. Uddin, A. Sajid, "Optimal Emerging trends of Deep Learning Technique for Detection based on Convolutional Neural Network", Bulletin of Business and Economics (BBE)., vol. 12, no. 4, pp. 264-273, Nov. 2023

[39] Mian Muhammad Abdullah, Hamayun Khan, Sabestean, Muhammad Farhan, Farhan Khadim, & Hafiz-ul-Asad. (2024). An Advance Machine Learning (ML) Approaches for Anomaly Detection based on Network Traffic. Spectrum of Engineering Sciences, 2(3)

[40] Sarwar, H. Khan, I. Uddin, R. Waleed, S. Tariq, "An Efficient E-Commerce Web Platform Based on Deep Integration of MEAN Stack Technologies", Bulletin of Business and Economics (BBE)., vol. 12, no. 4, pp. 447-453, Jun. 2023

[41] Haripriya, L.A., Jabbar, M., &Seetharamulu, B. (2018). A Novel Intrusion Detection System Using Artificial Neural Networks and Feature Subset

Spectrum of Engineering Sciences
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

Selection. International Journal of Engineering and Technology, 7(4), 181. http://doi.org/10.14419/ijet.v7i4.6.20458

[42] Aaraiz Naveed, Hamayun Khan, Zaeem Imtiaz, Wahaj Hassan, & Uneeb Fareed. (2024). Application and Ethical Aspects of Machine Learning Techniques in Networking: A Review. Spectrum of Engineering Sciences, 2(3)

[43] Hafiz M. Saqlain Khan, Hamayun Khan, Ch. Muhammad Akhtar Hayat, Hassan Tayyab, & Kashif Ali. (2024). An Enhanced Cost Effective and Scalable Network Architecture for Data Centers. Spectrum of Engineering Sciences, 2(3).

[44] M. Shahzar Nasir, Hamayun Khan, Ahmad Qureshi, Arslan Rafiq, & Taha Rasheed. (2024). Ethical Aspects In Cyber Security Maintaining Data Integrity and Protection: A Review. Spectrum of Engineering Sciences, 2(3)

[45] Hoang, D. H., & Nguyen, H. D. (2018). A PCA-based method for IoT network traffic anomaly detection. In2018 20th International Conference on Advanced Communication Technology (ICACT), 381-386. IEEE.

[46] Hodo, E., Bellekens, X., Hamilton, A., Dubouilh, P. L., Iorkyase, E., Tachtatzis, C., & Atkinson,
R. (2016). Threat analysis of IoT networks using artificial neural network intrusiondetection system. In2016 International Symposium on Networks, Computers and Communications (ISNCC), 1-6. IEEE.

[47] Blanco, R., Malagón, P., Briongos, S., & Moya, J. M. (2019). Anomaly Detection Using Gaussian Mixture Probability Model to Implement Intrusion Detection System. InInternational Conference on Hybrid Artificial Intelligence Systems, 648-659. Springer, Cham.

[48] Muhammad Abdullah, Hamayun Khan, Ayesha Shafqat, Muhammad Daniyal, Muhammad Bilal, & Muhammad Anas. (2024). Internet of Things (IoT's) in Agriculture: Unexplored Opportunities in Cross – Platform . Spectrum of Engineering Sciences, 2(3)

[49] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey.ACM Computing Surveys (CSUR),41(3), 1-58.