IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 May 2024, accepted 27 May 2024, date of publication 19 June 2024, date of current version 9 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3416863

RESEARCH ARTICLE

Congestion Control Prediction Model for 5G Environment Based on Supervised and Unsupervised Machine Learning Approach

MOHAMMED B. M. KAMEL[1,2], IHAB AHMED NAJM[3], AND ALAA KHALAF HAMOUD[4]

[1]Department of Computer Science, University of Kufa, Najaf 54003, Iraq
[2]Department of Computer Algebra, Eötvos Lóránd University (ELTE), 1053 Budapest, Hungary
[3]Department of Mathematics, University of Tikrit, Tikrit 43000, Iraq
[4]Department of Cybersecurity, University of Basrah, Basrah 61004, Iraq

Corresponding author: Mohammed B. M. Kamel (mkamel@uokufa.edu.iq; mkamel@inf.elte.hu)

ABSTRACT With the emergence of 5G technology, congestion control has become a vital challenge to be addressed in order to have efficient communication. There are several congestion control models that have been proposed to control and predict the possible congestion in 5G technology. However, finding the optimal congestion control model is an important yet challenging task. In this paper, we examine the supervised and unsupervised machine learning approaches to the task of predicting the possible node that causes congestion in the 5G environment. Due to the huge variance in the domains of the data set columns, measuring the prediction's consistency was not an easy task. During our study, we tested twenty-six supervised and seven clustering algorithms. Finally, and based on the performance criteria, we have identified the best five algorithms out of the studied algorithms.

INDEX TERMS Machine learning, congestion control, 5G, supervised ML, unsupervised ML.

I. INTRODUCTION

Compared to previous network generations, 5G networks have higher speeds, lower latency, and improved coverage. These features and its superiority over previous generations resulted in its widespread adoption [1]. Due to the widespread adoption and joining of a high number of nodes in the network, many new challenges have been raised, especially in the area of congestion control [2]. The goal of a routing algorithm is to choose the best possible path and avoid any potential congestion; yet, it may result in additional costs during the routing process [3]. As it can result in severe delays and lower throughput, congestion during 5G routing decisions becomes critical.

Several studies have been made for implementing various congestion control approaches in the 5G environment [4], [5]. Among the features of 5G networks that reduce congestion is the ability to dynamically distribute available resources, including frequencies and bandwidth. Network slicing [6] and edge computing [7], which allow traffic-based optimization of 5G networks, may be utilized to achieve this. Implementing Quality of Service (QoS) techniques ensures that critical services remain unaffected by current traffic [8]. Important traffic, like emergency services, is assigned with a higher priority, and resources in the 5G network are allocated appropriately. Another congestion control mechanism is traffic offloading, which transfers the data traffic to Wi-Fi [9] or other networks. The offloading is done to decrease the load on 5G networks, thus minimizing congestion.

In addition to the discussed approaches, applying machine learning (ML) algorithms has shown positive results in controlling network congestion [10]. While unsupervised ML algorithms are trained with unlabeled data, supervised ML algorithms are trained with labeled data [11]. In order to control congestion, both supervised and unsupervised algorithms are trained to identify possible congestion nodes as well as the optimal congestion control window. Classification is an essential part of supervised ML, where data items are grouped

The associate editor coordinating the review of this manuscript and approving it for publication was Bilal Khawaja.

into classes based on the class labels information. On the other hand, clustering is an essential part of unsupervised ML, in which similar data items are categorized into clusters without the information of class labels. The adoption of supervised and unsupervised algorithms is dependent on several factors including data type and size, complexity, and accuracy.

Our paper highlights the importance of adopting and utilizing machine learning algorithms in the process of congestion control in the 5G environment and identifies the top algorithms in the process of congestion control prediction. The task of finding the optimal algorithm to be adopted for congestion control is challenging. We aimed to find the optimal algorithm that predicts the optimal node causing congestion during the congestion control process in 5G networks. In our study, we tested twenty-six supervised and seven unsupervised algorithms. Unsupervised machine learning algorithms have been used for classification. The approach of classification via clustering is utilized to improve the accuracy of congestion control prediction by clustering data to identify distinct groups of data which be used to enhance the classification process. Cronbach's alpha has been used to measure the consistency undimensionality or homogeneity of datasets. During the evaluation, the studied algorithms were evaluated based on performance criteria, including True Positive (TP) and False Positive (FP) rates, precision, recall, Receiver Operating Characteristic (ROC), and Area Under Curve (AUC). Fig. 1 shows the main steps of the congestion control prediction model.

The rest of this paper is structured as follows. In Section II, we studied the related works in the field of congestion control. Section III discusses machine learning and congestion control in detail. The model setting has been stated in Section IV. Our findings are explained and analyzed in Section V. We point out our observations in Section VI. Finally, Section VII includes our conclusions.

II. LITERATURE REVIEW

Many studies have handled the congestion control approach. Sangeetha et al. [12] proposed a model based on data loss and energy reduction since congestion appears in all WSNs. The sensor nodes' topology is adjusted regularly based on node degree and time interval to enhance the node's power consumption and interference and to provide a better and more effective energy congestion-aware technique for routing in WSN, which is called survival path routing (SPR). This protocol is used by IoT applications in high-traffic networks where all nodes try to send their packets simultaneously to destination nodes [13]. A new algorithm for congestion control for WSNs is developed by Singh et al. [14], where a simplified poisson process is used and the optimal rate is obtained by retransmitting with congestion control, while the old algorithm had a high complexity and high power usage.

Subsequently, many studies evaluated the performance of congestion control mechanisms over the 5G network [15] in terms of resource allocation [16], network selection, network scalability [17], and distributed telemetry [18]. To reduce network congestion, enhance the lifetime of the network and individual nodes, and reduce network divisions, Shelke et al. [19] proposed a routing algorithm that selects the best route by combining appropriate sleep scheduling mechanisms based on the opportunistic theory. Godoy et al. [20] analyzed and investigated the communication channel congestion in the environment based on configuration parameters of nodes such as the generation rate of the data packet, intervals of transmission time, and power level of transmitter output. Najm et al. [21] proposed a multi-criteria decision-making mechanism to improve congestion control in 4G networks.

Braham et al. [22] proposed an efficient and fair distributed algorithm for congestion control in tree-based communication WSNs to assign transmission rates for each node. The study lacked a performance comparison with previous traditional algorithms to see if it was optimal or not. Although the next scenario was poor and simple, applying machine learning algorithms, especially supervised ML, to improve congestion control in wireless or wired networks is considered a vital approach. Machine learning algorithms can be adopted in many fields [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33] to predict the required knowledge. Geurts et al. [34] proposed a model based on an automatic loss classifier based on a simulated database of random topologies of networks. Jagannathan and Almeroth [35] proposed a model called TopoSense for multi-cast congestion control. Many enhancements were required, such as the poor calculation of link capacity and the need for calculating interval size. Moreover, there was a need to minimize control traffic and burst traffic.

Following the trend, machine learning capabilities have been utilized with congestion control algorithms in 5G environments. Several attempts have been presented, for instance, in an open radio access network, a fast increase in data based on artificial intelligence, and an adaptive routing control approach to obtain effective congestion avoidance [36], [37], [38]. A controller is proposed by Sunny et al. [39] to ensure the efficient and fair work of WLAN that has multi-cochannel access and improvement of long-lived multi-TCP AP transfer. Next, many researchers adopted DT in their studies of network applications.

Katuwal et al. [40] proposed a model to solve the problem of multi-class classification based on the multi-classifier system. An efficient NN with oblique random forest DT is used to build the model. The model proved its efficiency based on the evaluation of 65 multi-class datasets compared with the evaluation of large or medium datasets. Gomez et al. [41] compared many ensemble algorithms of DT and proposed a new classifier based on its performance. The computed capacity for devices of a small network is not a limitation. A new model is proposed by Leng et al. [42] to solve the problem of congestion control flow table in a software-defined network (SDN) based on C4.5 DT. The flow entries are compared based on C4.5 DT to reduce the time and matching cost. Using the DT approach with an SDN flow

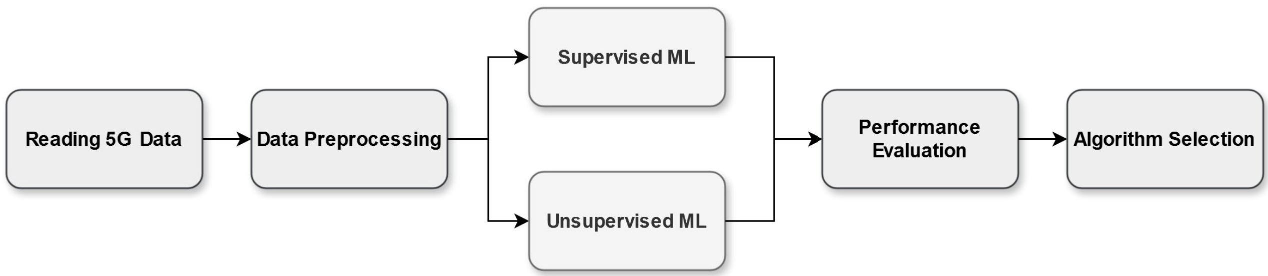


FIGURE 1. Main steps in ML based congestion prediction.

table was the first online machine learning model. Next, the clustering machine learning approach is used for localization and AP reselection. Liang et al. [43] proposed a model for WLAN that adopts a clustering algorithm and AP reflection.

A review of the communication technology of machine-to-machine was conducted by Hasan et al. [44] to list all challenges and solutions for diverse standards of developing organizations. Liu and Wu [45] utilized the random forest algorithm for congestion control prediction, where some variables were utilized to build the model, such as type of day, road quality, time period, and weather conditions.

Park et al. [46] proposed an approach utilizing Bayesian neural network and DT to predict the occurrence of incidents. Since the aim of the model is to reduce the potential incidents or any events that may cause these incidents, the model could not implemented in real systems since it needs realistic parameters for training and building a dataset. An improved route based on the support vector machine (SVM) with DT is used to estimate the link quality over WSN, where Shu et al. [47] used two estimation parameters: link quality and the strength of the received signal. SVM is used in the model due to its ability to handle binary classifications.

For network infrastructure and data centers, DT was adopted as an energy-saving solution [48]. Soltani and Mutka [49] proposed an approach utilizing DT for best path selection in the cognitive radio networks. In this model, the nodes can find better nodes to send data to after analyzing the tree and removing the choices that reduce node gain. DT is utilized to interpret the routing path of cognitive video over a dynamic radio network. The optimal path from the leaf node to the root is determined based on background induction to construct and receive the transmitted video. The DT algorithm is also used by Stimpfling et al. [50] to build a model to enhance data structure size and memory access. DT is considered a strength since it reduces the searching time. Moreover, DT is used by Singh et al. [51] to build a model for vehicular traffic noise prediction. Four machine learning algorithms are used for model implementation: DT, ANN, generalized linear model, and random forest. The random forest approach was found to be a better algorithm for prediction compared with other algorithms. Xia et al. [52] used DT with a proposed delegation schema (CP-ABE) to enhance the efficiency of decryption for VANETs. The purpose of DT is to improve the factors used for estimating vehicle decryption overhead.

Based on the results, DT was a better option than K-nearest and SVM for prediction because of its higher precision and accuracy. Many researchers have highlighted network protection by utilizing the DT notion. For example, researchers in [53], [54] presented an unknown detection threat approach in the network via recognition threat features. Following the trends, Mohamed et al. [55] developed a flexible scheme for reducing the quantity of data transmitted across the smart grid, but the intended scheme missed mentioning the outcome of paradigm updates. Pham and Yeo [56] presented an adaptive and protected scheme for cars to control both confidentiality and trust in the utilized recognition scheme. Next, Fadlullah et al. [57] highlighted and explored the survey requirements of propagation techniques related to deep learning utilizations concerning numerous traffic network control characteristics. The leading edge of peak network communications, which are compromised by algorithms and architectures in deep learning, also encourages the motivation to facilitate deep learning to compromise the network's challenges. Nevertheless, their viewpoints did not include the 5G environment.

Furthermore, Kong, Zang, and Ma [58] developed dual machine-learning approaches to address TCP congestion control issues in under-buffered connections over the wired environment. A supportive and adaptive loss prediction was assigned to obtain a superior tradeoff delay. In research issued by Taherkhani and Pierre [59] used the K-means algorithm to control congestion in VANET networks. It contains three sections for directing, detecting data congestion, and clustering communications. Next, the issue of prediction traffic status was settled by Chen et al. [60] by permitting DT and SVM to depend on enabling online data; however, the set value of both services was overridden.

Tariq et al. [61] presented a detection of botnet attacks by using the machine learning technique, regardless of the explanation of the carried packet. However, comprehensive calculations missed the stats plan. Wu et al. [62] implemented a developable machine learning method to predict or expose the limps of online video via feature extraction of monitored data in the network. The method defines characteristic features depending on diverse scale windows. The criterion of used information gain, however, demands audio pieces to be overridden, besides ignoring the cache amid the sending nodes and the video operator, constant section length, and the limited interaction of the user. Quality of experience prediction is based on consultative factors, namely the Video Quality Model (VQM) and Structural Similarity Index (SSIM) parameters. Nevertheless, a justification for why k = 9 assigned in a random forest was superior is shown by Abar et al. [63].

III. BACKGROUND

In this section, machine learning and congestion control have been discussed in detail.

A. MACHINE LEARNING

Machine learning is a very fast-growing path that can be classified into supervised, unsupervised, semi-supervised, and active ML. Supervised ML refers to classification, where the labeled examples of the training data set determine the learning in supervised ML. Unsupervised ML is a synonym for clustering, where the input classes in the training data set are not labeled. Unsupervised ML is essentially used to discover hidden patterns within data sets. Semi-supervised refers to using both labeled and unlabeled examples during model learning, labeled for learning the model and unlabeled for refining class boundaries. Active learning permits users to play an active role in the machine learning process [64]. The basic element in the machine learning model is the dataset. The size of the dataset, the features' domain, the number of features, and the data type of all features are essential elements in the machine learning model. The size of the data set affects the overall accuracy since the training data set will be larger and the learning model will have more data to learn from. Many reasons can improve the accuracy of the classification algorithm, such as data cleaning, adding missing values, increasing the data set, feature selection and transformation, bagging, and boosting.

1) SUPERVISED MACHINE LEARNING

There are many algorithms classified as supervised ML: Classification and Regression trees (CART) were introduced in 1984 by Breiman based on splitting the explicative variables' space into multidimensional rectangle form and a local predictor with each one of them. The tree structure development approach came from the recursive partitioning of the data set into two homogeneous data sets, which led to building a branching structure [65]. The regression model captures how one or more variables vary across more attribute domains, which can be used to predict the target variable [66].

The DT approach is a frequent approach that comes from the tree-based approach to data classification. The reliability, low cost, and ease of implementation are the reasons behind adopting such an approach [67]. The basic structure of the decision tree starts with one node and branches to many other nodes, forming the tree graph. DT algorithms (ID3, CART, J48, RepTree, Decision Stump, Hoeffding, and Random Forest) are used in many fields, such as market analysis, educational data mining, and estimating risks [68], [69], [70]. When the DT is used for attribute selection, the data within the data set is examined, and only the relevant data are chosen. Only the selected attributes appear in the graph after excluding the irrelevant data [64].

A Naäve Bayesian classifier (belief network) is a graphical model that represents the random variables set as knowledge, where each node represents the corresponding random variable and the edges represent the conditional dependency between variables. These conditional dependencies represent computational methods and statistical probabilistic theories. The Naäve Bayes algorithm is a simple algorithm built based on Bayesian theory, where it works basically on the conditional probability [67].

Artificial Neural Network (ANN), where the model is implemented like a human brain neuron network. The ANN and biological brain are similar in two keys: the connections between the neurons that determine the network function and the building blocks of the computational devices [71]. The multilayer perceptron algorithm is one of the ANN algorithms that work on training datasets by gathering information by minimizing the error and applying that information to the new dataset [67]. ANN is used for prediction, pattern recognition, optimization, and control, and many researchers have used ANN to solve many problems in many disciplines. ANN is used for many reasons, such as low energy consumption, adaptivity, learning ability, distributed computation and representation, massive parallelism, and fault tolerance [72]. Support Vector Machine (SVM) is an ML algorithm that learns from training data sets and assigns labels to dataset objects. SVM can be used in many disciplines, such as fraud detection, anomaly detection, image recognition, gene classification, and educational data mining [73].

2) UNSUPERVISED MACHINE LEARNING

The synonymous term for unsupervised machine learning is clustering, where the information of the class label is not presented. The clustering approach is defined as grouping similar data items into clusters. The clustering algorithms are provided with data items with no labels, and the task of these algorithms is to represent the data distribution suitably. The learning approach in unsupervised machine learning is based on observation, while the supervised machine learning approach is based on learning by examples. For large databases, efforts have been focused on exploring the most effective method for efficient cluster analysis. Clustering is used in exploring the different data types with different types of sizes, such as complex shapes, graph clustering, image clustering, and object clustering with a huge number of features [64], [74].

The process of cluster analysis is based on partitioning the similar objects in the data sets into subsets, where each subset represents a cluster. Based on the partitioning method, different clustering algorithms may result in different clusters for the same data set. This partitioning or data segmentation may lead to the discovery of unknown groups of clusters that can be noticed by humans. Clustering can be used for outlier detection. The outliers are the values far away from any cluster. Many approaches of clustering can be used to build clusters based on the method of selecting a group of objects.

Hierarchical methods based on a non-parametric clustering approach produce a dendrogram (a tree of clusters). These algorithms measure the dissimilarities among cluster sets for each iteration. The hierarchical methods can be classified based on the form of the hierarchical decomposition into divisive and agglomerative. The agglomerative or bottom-up approach forms the topmost group by grouping close objects. The other approach (divisive) or top-down approach starts with all objects that belong to the same cluster. The methods of hierarchical clustering can be continuity-based or distance-based. However, when the split or merge step is performed in hierarchical methods, it cannot be undone [75].

Partitioning methods divide each dataset into several groups, where each group contains at least one object. In these methods, the object must exactly belong to one group. Fuzzy partitioning is an example of these methods. Prohibitive computations and exhaustive enumeration are required to achieve optimal clustering in partitioning methods. For that, greedy methods, k-means, and k-medoids may be adopted to overcome this obstacle and lead to building optimal clusters. The partitioning methods work better with small to medium-sized data sets, where cluster building is based on finding clusters with a spherical shape [76].

Density-based methods differ from the other clustering methods by building clusters based on density, where the cluster is built as long as the objects are in the same neighborhood (or the same density). The density-based methods divide the objects into a hierarchy or multiple exclusive clusters. These methods are the optimal choice to find the outlier and noise. They are also optimal choices for discovering the arbitrary shape of clusters [77].

Grid-based methods form the grid structure of the cluster by quantizing the space of the objects into a limited number of cells. All the operations that can be performed on the clusters are performed on this grid structure. Fast processing is the main advantage of this approach. The short processing time results from the number of cells in the processed dimension. The grid-based approach is optimal for spatial datasets and can be used with other clustering approaches such as hierarchical and density-based methods [78].

B. CONGESTION CONTROL MECHANISM

Congestion control mechanisms are crucial to the transport layer protocols. The transport protocol can perform various functions, including message transmission, error detection, and message retrieval, by engaging throughout this layer [15], [79], [80]. Functionality is matched to network utilization. The number of terminals needed for these functions is determined by analyzing the terminal-area-used units. Data transit requires a transmitter-receiver link, where the transmitter connects to a certain endpoint. Congestion control mechanisms are divided into a slow-start algorithm and a congestion avoidance algorithm. The congestion control mechanism sends the initial message and awaits acknowledgment to monitor the congestion window and slow start threshold. The recipient sends an acknowledgment to the transmitter, identifying the congestion window and slow start threshold. As a result, congestion is controlled. The misplaced phase is retrieved if the recipient does not conduct acknowledgment. If the congestion window indicator is less than or equal to the slow start threshold, the slow start phase begins. Further related parametric settings and information can be found in [10].

IV. MODEL SETTING

The proposed congestion prediction model has been illustrated in Figure 2. The mmWave ns-3 module [81] and protocols were utilized to test network protocols, including TCP and SCTP, in the 5G environment [82]. This module for mmWave 5G cellular network simulation has many characteristics, including providing the ability to study the cwnd [83]. It supports multiple channel models, including 3GPP TR 38.901 for 0.5–100 GHz. It also provides adaptable PHY and MAC classes that support 3GPP NR frame structure and adaptable schedulers for dynamic TDD formats. Among its main features is the possibility of improving the RLC layer with packet re-segmentation. The model supports quick secondary cell handover, channel tracking, and dual LTE base station connectivity.

The utilized dataset in the proposed model, as shown in Table 1, is a 100-record dataset with five columns: sequence, congestion window size, throughput, queue size, and packet loss. This dataset is divided into two sets: 80 record dataset for training and 20 record dataset for testing the model. More enriched configurations are available in [10].

TABLE 1. Dataset structure.

N	Congestion window	Throughput	Queue size	Packet loss	Optimal
1	$cwnd_1$	th_1	qs_1	pl_1	N
. . .	. . .	. . .	. . .	. . .	. . .
100	$cwnd_{100}$	th_{100}	qs_{100}	pl_{100}	N

After the data visualization, the results show that there is no dirty, missing data, noise, or inconsistent data that needs to be handled or cleaned. Based on that, the only step in the data preprocessing is performing derived columns based on the columns. The derived column will be utilized as the goal column for the supervised ML. The associated parameters used to derive the goal column are utilized to determine the optimal node for prediction. A simple mechanism is followed in determining the optimal node, where the optimal node is the node with high throughput, congestion window size, and queue size with the lowest packet loss.

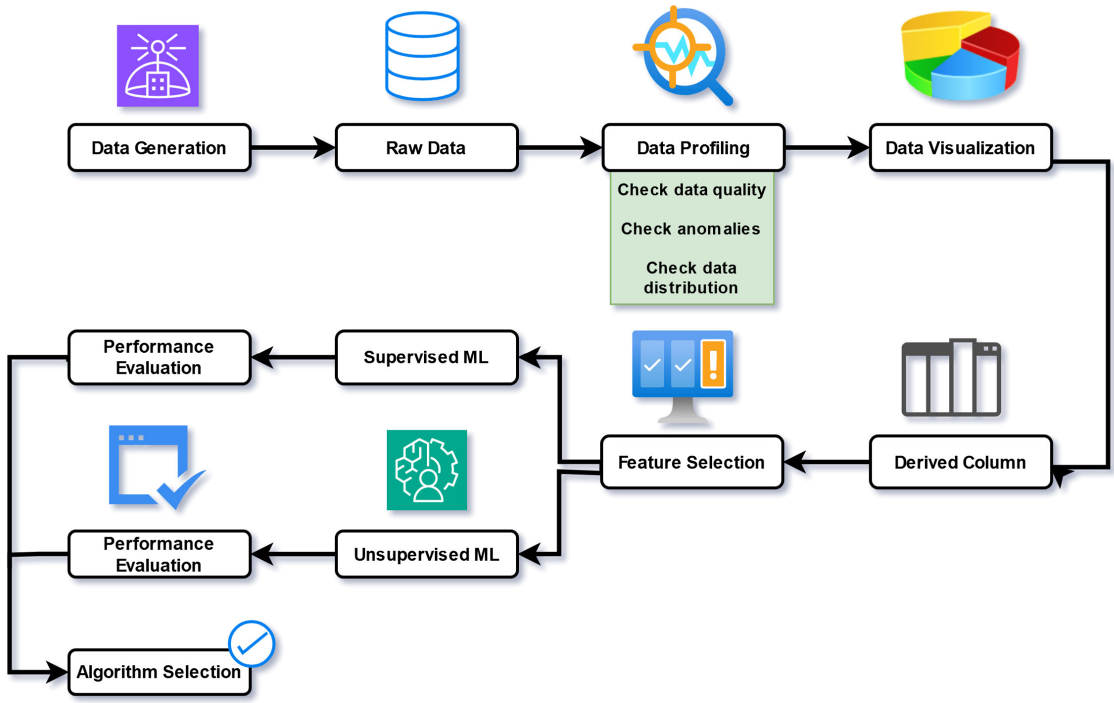


FIGURE 2. Congestion control prediction model.

The mechanism that has been followed to determine the optimal and non-optimal nodes is set by comparing the value of congestion window size to determine if it is greater than the mean value of congestion window sizes, if throughput is greater than the mean value of throughput, and if the queue size is greater than the mean value of queue sizes, as well as if the packet loss is less than the mean value of packet loss. The optimal (O) nodes in the dataset are less than the non-optimal (N) nodes based on the mentioned mechanism. The labeled nodes (O and N) have been utilized in implementing supervised ML in order to find the optimal algorithm for prediction and classification.

V. CONSISTENCY MEASUREMENT

Cronbach's alpha measures the internal consistency, undimensionality, or homogeneity of the dataset. The measure is the variance of the item appearance in the dataset, and its value falls between 0 and 1. Cronbach's alpha measures the inter-relatedness of the items in the same attribute domain of the dataset [84], [85]. The Cronbach's alpha can be calculated as follows:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum S_i^2}{S_T^2}\right)$$

where k represents the number of items, S_i^2 represents the variance of the ith item, and S_T^2 is the sum of all variances for all items. The results of implementing Cronbach's alpha formula on the dataset are listed in Table 2.

As shown in Table 2, the value of Cronbach's alpha is poor, which reflects the lack of consistency in the dataset. The reasons behind this are the variances in each column item, which increase both the sum of all variances (3.67503E+11) and the variances of total scores (3.69033E+11). These two high values result in Cronbach's alpha score being very weak due to huge variances in the same domain of each attribute, specifically in queue size, congestion window, packet loss, and throughput.

Table 3 lists the performance criteria for implementing the supervised machine learning algorithms. The performance criteria are: (True Positive (TP) rate, False Positive (FP) rate, Precision, and Recall). The TP rate represents the positive instances that are classified correctly, and the FP rate represents the false positive instances that are classified for a given class. Precision is the ratio of the positive predicted values, while recall represents sensitivity.

$$TP\ Rate = 100\ \frac{TP}{TP + FN}$$

where TP represents True Positive values and FN represents False Negative values.

$$FP\ Rate = 100\ \frac{FP}{FP + TN}$$

where FP represents False Positive values, TN represents True Negative values.

$$Precision = 100\ \frac{TP}{TP + FP}$$

where FP represents False Positive values.

$$Recall = 100\ \frac{TP}{TP + FN}$$

TABLE 2. Cronbach's alpha result.

Cronbach's alpha	Number of items	Sum of variances	Variance of total score	Number of rows
0.005182749	5	3.67503E+11	3.69033E+11	92

TABLE 3. Performance criteria of supervised algorithms.

Category	Algorithm	TP rate	FP rate	Precision	Recall	ROC	AUC
	Decision Stump	0.891	0.146	0.917	0.891	0.848	0.8476
DT	Hoeffding Tree	0.946	0.009	0.961	0.949	0.986	0.9864
	J48	0.924	0.334	0.919	0.924	0.853	0.853
	LMT	0.967	0.134	0.967	0.967	0.988	0.9833
	RandomForest	0.946	0.202	0.944	0.946	0.982	0.982
	RandomTree	0.957	0.136	0.957	0.957	0.91	0.9104
	RepTree	0.913	0.271	0.913	0.913	0.874	0.8744
BN	BayesNet	0.957	0.136	0.957	0.957	0.987	0.9873
	NaiveBayes	0.946	0.009	0.961	0.946	0.987	0.9873
	NaiveBayesUpdateable	0.946	0.009	0.961	0.946	0.987	0.9873
Rules	JRip	0.946	0.202	0.944	0.946	0.827	0.8267
	DecisionTable	0.935	0.268	0.932	0.935	0.966	0.9659
	OneR	0.859	0.409	0.863	0.859	0.725	0.7249
	PART	0.924	0.334	0.919	0.924	0.853	0.853
	ZeroR	0.859	0.859	0.737	0.859	0.413	0.4129
Misc	FLR	0.957	0.2	0.955	0.957	0.878	0.8783
	InputMappedClassifier	0.859	0.859	0.737	0.859	0.413	0.4129
Functions	Logistic	0.946	0.137	0.948	0.946	0.919	0.91916087
	MultilyerPreceprton	0.957	0.136	0.957	0.957	0.981	0.9815
	SGD	0.946	0.202	0.944	0.946	0.872	0.872
	SimpleLogistic	0.957	0.136	0.957	0.957	0.981	0.9815
	SMO	0.957	0.2	0.955	0.957	0.878	0.8783
	VotedPreceptron	0.859	0.859	0.737	0.859	0.488	0.48796087
Lazy	IBK	0.957	0.136	0.957	0.957	0.936	0.9357
	KStar	0.837	0.541	0.832	0.834	0.771	0.77144674
	LWL	0.913	0.143	0.927	0.913	0.9	0.89990109

Receiver Operating Characteristic (ROC) and Area Under Curve (AUC) are important criteria used for evaluating and measuring the performance of machine learning algorithms. ROC and AUC are suitable for visualizing the performance of different classifiers in supervised and unsupervised learning fields.

ROC takes a value between 0 and 1 that reflects the classifier's accuracy. As much as the ROC value reaches value 1, the model becomes more accurate [56], [86]. The AUC takes a value from 0 to 1, where 1 indicates the classifier performance as perfectly accurate and 0 indicates the classifier performance as perfectly inaccurate. A value falling between 0.7 and 0.8 is considered acceptable; an excellent value is a value between 0.8 and 0.9; and a value above 0.9 is considered an outstanding value. On the other hand, the value that lies under 0.5 indicates that the classifier performance is weak [71].

Table 3 lists a comparison among different categories of supervised learning, such as DT, where different algorithms such as Decision Stump, Hoeffding Tree, J48, LMT, RandomForest, RandomTree, and RepTree are examined. The field of Bayes Net (BN) is also examined, and different algorithms are used, such as BayesNet, NaiveBayes, and NaiveBayesUpdateable. In the regression category, Logistic, Stochastic Gradient Descent (SGD), and Sequential Minimal Optimization (SMO) have been examined. Additionally, other algorithms and classifier categories, such as Neural Network based algorithm (MultilayerPrecetron), K-Mean, KStar, and Locally Weighted Learning (LWL) have been examined. Based on the performance results, the top five supervised machine learning algorithms have been selected to visualize the best one among them. The comparison is implemented based on TP, FP, Precision, and Recall as a first step. The second step is to visualize the ROC and AUC of the selected top five algorithms.

Fig. 3 lists the performance criteria of the top five supervised algorithms (LMT, BaysNet, MultilyerPreceptron, SimpleLogistic, and IBK) as a chart. The LMT algorithm came in first with (96.7%) followed by the remaining four algorithms with (95.7%) in predicting the TP values. LMT also came in as the top algorithm in low predicting rate of FP values with (13.4%) followed by the remaining four algorithms with (13.6%). The LMT algorithm came in first place in precision with (96.7%) followed by the remaining algorithms with (95.7%) and also in first place in recall with (96.7%) followed by the remaining algorithms with (95.7%).

The value of ROC is considered outstanding if it exceeds the value of 0.9 [87], [88]. Based on that, all five algorithms are outstanding at predicting instances. LMT came in first in ROC value with (0.988), followed by the BaysNet algorithm with (0.987), MultilayerPrecepton and SimpleLogistic algorithms with (0.981), and the IBK algorithm with (0.936). The ROC of BaysNet with a value of (0.9873) came in first place, followed by the LMT algorithm with (0.9833), the MultilayerPreceptron and SimpleLogistic algorithms with (0.9815), and the IBK algorithm with (0.9357). The ROC

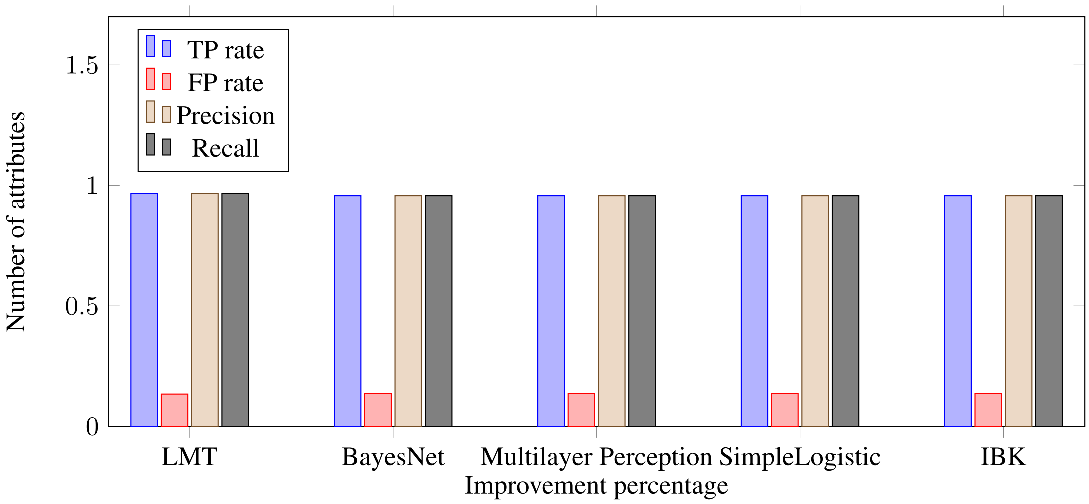


FIGURE 3. Performance of supervised machine learning algorithms.

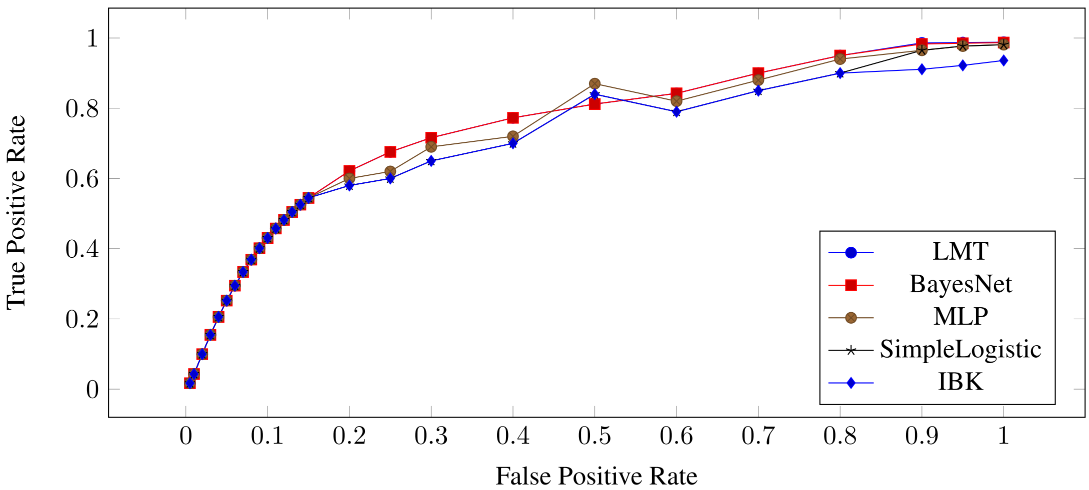


FIGURE 4. ROC of supervised machine learning algorithms.

and AUC have been shown in detail in Fig. 4 and Fig. 5, respectively.

Table 4 lists the performance criteria of unsupervised machine learning algorithms, specifically in the clustering approach. EM algorithm is the top accurate algorithm in predicting the TP values with (98.5%) followed by Farthest First, FilteredCluster, and SimpleKMean algorithms with scores of (92.4%), MakeDensityBased algorithm with (91.3%), HierarchicalCluster algorithm with (85.7%), and Canopy algorithm with (83.7%). Based on FP rate, EM came in first with a low prediction rate of FP values with (0.2%), followed by Simple K-Mean and Filtered Clusterer algorithms with (1.3%), MakeDensityBased algorithm with (1.4%), Farthest First algorithm with (7.7%), Canopy algorithm with (54.1%), and Hierarchical Clusterer algorithm with (87%). Based on the Precision criterion, EM came in first place with (98.6%), followed by SimpleKMean and Filtered Clusterer algorithms with (95.1%), MakeDensityBased algorithm with (94.6%), Farthest First algorithm with (94.1%), Canopy algorithm with (83.2%), and Hierarchical Clusterer algorithm with (75.2%). The ROC and AUC have been shown in detail in Fig. 7 and 8, repectively.

Fig. 6 represents the performance criteria of the top three unsupervised machine learning algorithms, namely: EM, Filtered Clusterer, and Simple K-Mean algorithms. Fig. 6 shows that the EM algorithm is the best algorithm for clustering with (98.5%), followed by both Simple K-Mean and Filtered Clusterer algorithms with (92.4%) in predicting the TP values. EM came in first place for the low predicting rate of FP values, as well as in precision and recall.

As discussed earlier, the values of ROC and AUC reflect the accuracy of predicting the TP rate and FP rate. As shown in Fig. 7 and Fig. 8, the ROC and AUC of both simple K-Mean and Filtered Clusterer algorithms are considered to be outstanding since they exceed the 0.9 value with 0.956 for ROC and 0.9557 for AUC, followed by the EM algorithm

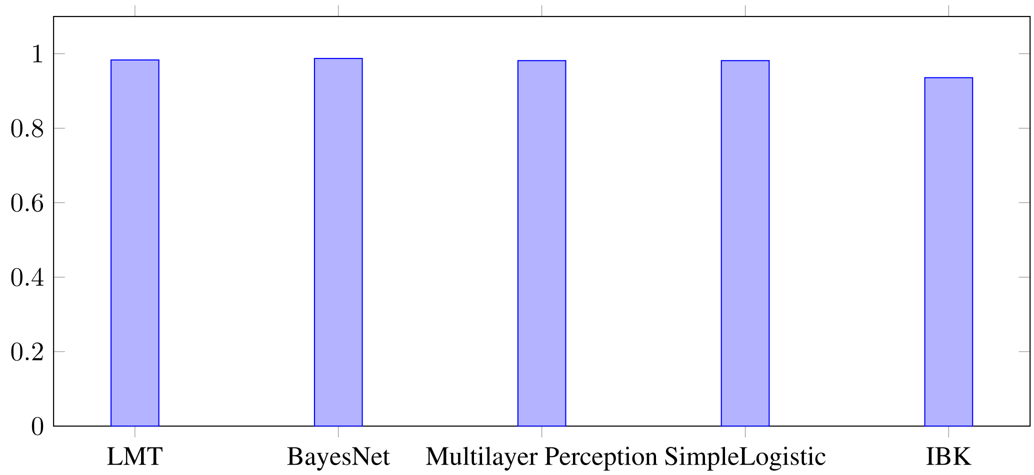


FIGURE 5. AUC of supervised machine learning algorithms.

TABLE 4. Performance criteria of unsupervised algorithms.

Category	Algorithm	TP rate	FP rate	Precision	Recall	ROC	AUC
Clustering	Canopy	0.837	0.541	0.832	0.837	0.771	0.88545
	EM	0.985	0.002	0.986	0.985	0.852	0.61151087
	Farthest First	0.924	0.077	0.941	0.924	0.924	0.9236
	FilteredCluster	0.924	0.013	0.951	0.924	0.956	0.9557
	HierarchicalCluster	0.857	0.87	0.752	0.857	0.725	0.52130761
	MakeDensityCluster	0.913	0.014	0.946	0.913	0.949	0.9494
	SimpleKMean	0.924	0.013	0.951	0.924	0.956	0.9557

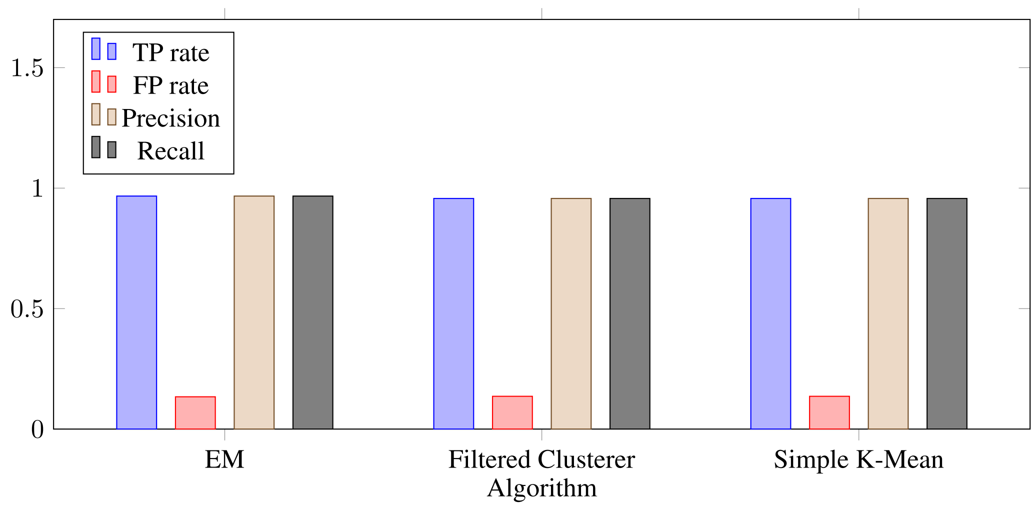


FIGURE 6. Performance of unsupervised machine learning algorithms.

with 0.852 for ROC and 0.611 for AUC. Since EM is considered to be the optimal algorithm for clustering based on the performance criteria discussed earlier, the AUC and ROC also support this concluded point.

VI. MODEL PERFORMANCE OBSERVATION

The basic concept of attribute selection in DT relies on the measures adopted for the selection method. DT adopts information gain (IG) as an attribute selection method, where the attribute with the highest IG value is chosen as the splitting node. IG is the average amount of information used to classify the instances as a class label and is calculated by using different information entropy equations. The attribute type, data set characteristics, size, and dimensionality affect the accuracy of the DT algorithms [89], [90], [91]. The data type of the predicted class makes the DT classifier prediction accuracy high or low.

The LMT algorithm is based on two classification approaches: tree induction and logistic regression. This algorithm uses logistic regression for the leaves of the tree produced. As a result, the accuracy of the small dataset with a low number of attributes and no missing values will be very high. Moreover, the concept of building a network for classification or feature selection in the Bayesian approach

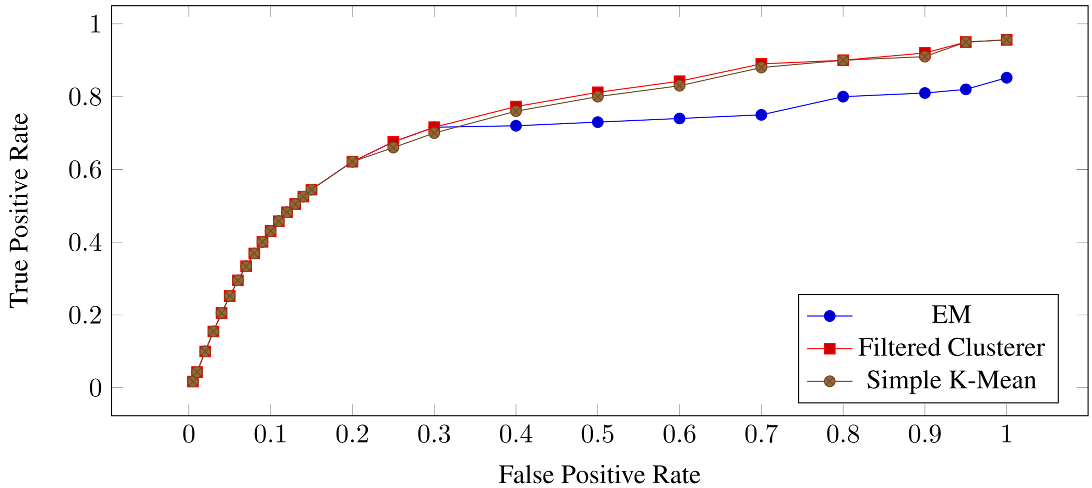


FIGURE 7. ROC of unsupervised machine learning algorithms.

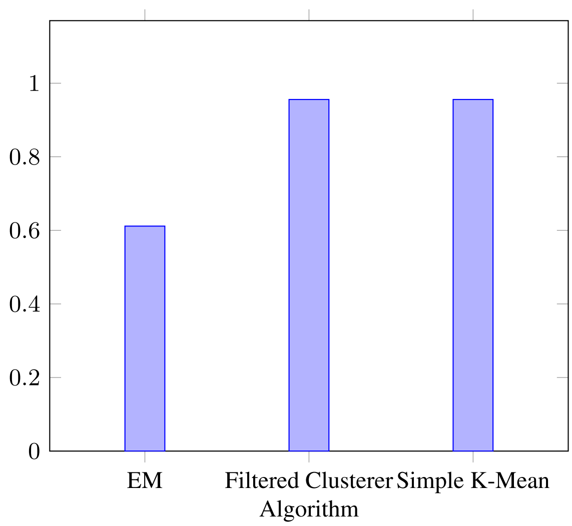


FIGURE 8. AUC of unsupervised machine learning algorithms.

relies on adopting probability to find the correlation among features. For the continuous features domain, the numeric attribute value is distributed, and then the distribution is represented later by its standard deviation and mean values. The probability after that can easily be calculated to find the correlation among attributes [92], [93]. Other algorithms, such as KNN and K star adopt probability and entropy approaches to measure the distance among attributes based on particular algorithms, such as IB1, 2, and 3, or even DT algorithms. Hence, the overall performance of the supervised algorithms is restricted by the type of the final class, whether it is a nominal or binary class, the type of the attributes, and whether there are missing values in the attributes, besides the previous characteristics of the dataset.

Regarding unsupervised learning, in addition to metrics or effects that increase the accuracy of the outputs, the selection of the algorithm is the most important challenge faced in order to obtain the best accuracy in results. Many properties should be found in a good clustering algorithm, such as performing well with massive data, analyzing single and mixtures of attribute types, and the ability of the algorithm to deal with deviations (outliers) to enhance the quality of the cluster. In addition, the results of the algorithm must be usable, interpretable, and easy to understand. Another feature of a good algorithm is its ability to operate with the lowest requirements for input parameters to avoid bias in the result, especially with higher dimensionalities and considerable data. Another feature to be considered is the sensitivity of the arrangement of inputs that can obtain different ideal results when presented to algorithms in different arrangements within the same data set. Finally, the optimal algorithm selection also depends on the kind of data set and the objective of the analysis [78], [94].

Since the clustering algorithm aims to be general, it is an important issue when selecting a clustering algorithm that makes the shape correspond to the resulting cluster. The clustering algorithms are biased toward determining the shapes and structures of the clusters, while it is not easy to determine the corresponding biased shape. The structure of the cluster may not be determined, especially with the datasets that hold categorical data types. The amount of dimension/attribute present in most datasets is huge. The majority of the existing clustering algorithms are unable to manage anything greater than a small number of dimensions, about eight to ten dimensions. Hence, the clustering of high-dimensional datasets is a challenge. An example of such high-dimensional datasets is the US census dataset. The presence of a huge number of attributes has proven to be the cause of dimensionality. This is associated with the following: (a) an increase in the number of attributes results in an increase in the number of resources needed to represent their growth; (b) for so many distance and distribution functions, the distance of a given point from the nearest and furthest neighbor is almost the same. As a result of the increase in time needed to process the data, both of the above-mentioned factors significantly affect the efficiency of a clustering algorithm.

Sequentially, the resulting clusters will have very poor quality [94].

VII. CONCLUSION

Many models and mechanisms are proposed to overcome congestion control problems and enhance the overall network performance. The proposed study handled the problem of congestion control in the 5G environment by examining supervised and unsupervised ML algorithms to find the optimal algorithm for predicting the optimal node.

In the field of supervised ML, twenty-six algorithms were tested: seven DT algorithms, three BN and lazy algorithms, five rules algorithms, and eight other algorithms. In the field of unsupervised ML, seven clustering algorithms are examined. Cronbach's alpha results showed that it is impossible to measure the consistency due to the huge variance in the data set columns' domains. This variance makes the prediction based on the changing data difficult. Many conditions determine optimal congestion window management, such as low packet loss, high congestion window, high queue size, and high throughput.

Since it is difficult to measure the performance of all supervised algorithms by charts, only the top five supervised algorithms were discussed based on their performance criteria, namely: LMT, BaysNet, MultilyerPreceptron, SimpleLogistic, and IBK. The LMT algorithm came in first with (96.7%) followed by the remaining four algorithms with (95.7%) in predicting the TP values. LMT also came in first place in the low predicting rate of FP values with (13.4%) followed by the remaining four algorithms with (13.6%). The LMT algorithm came in first place in precision with (96.7%) followed by the remaining algorithms with (95.7%) and also in first place in recall with (96.7%) followed by the remaining algorithms with (95.7%). The TP rate and FP rate are so close, due that, the ROC and AUC are measured for all algorithms to find the optimal one. LMT came in first in ROC value with (0.988) followed by the BaysNet algorithm with (0.987), MultilayerPreceprton and SimpleLogistic algorithms with (0.981), and IBK algorithm with (0.936). The ROC of BaysNet with a value (0.9873) came in first place, followed by the LMT algorithm with (0.9833), MultilayerPreceptron and SimpleLogistic algorithms with (0.9815), and the IBK algorithm with (0.9357).

In unsupervised ML, the performance criteria of the top three algorithms, namely: EM, Filtered Clusterer, and Simple K-Mean, were measured. The EM algorithm was the best algorithm for clustering with (98.5%) followed by both Simple K-Mean and Filtered Clusterer algorithms with (92.4%) in predicting the TP values. EM came in first place for the low predicting rate of FP values and first in precision and recall. The values of ROC and AUC reflect the accuracy of predicting the TP rate and FP rate. The ROC and AUC of both simple K-Mean and Filtered Clusterer algorithms were outstanding, with 0.956 for ROC and 0.9557 for AUC, followed by the EM algorithm with 0.852 for ROC and 0.611 for AUC.

Future research directions can include the implementation of optimal supervised and unsupervised ML algorithms in a real-world environment with stream data. Stacking can be examined to combine the best and optimal algorithms in the proposed model to predict the optimal node with higher accuracy and lower prediction time and resources.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] R. Dangi, P. Lalwani, G. Choudhary, I. You, and G. Pau, "Study and investigation on 5G technology: A systematic review," *Sensors*, vol. 22, no. 1, p. 26, Dec. 2021.

[2] N. Al-Falahy and O. Y. Alani, "Technologies for 5G networks: Challenges and opportunities," *IT Prof.*, vol. 19, no. 1, pp. 12–20, Jan. 2017.

[3] S. Malathy, P. Jayarajan, M. H. D. N. Hindia, V. Tilwari, K. Dimyati, K. A. Noordin, and I. S. Amiri, "Routing constraints in the device-to-device communication for beyond IoT 5G networks: A review," *Wireless Netw.*, vol. 27, no. 5, pp. 3207–3231, Jul. 2021.

[4] J. Lorincz, Z. Klarin, and J. Ožegović, "A comprehensive overview of TCP congestion control in 5G networks: Research challenges and future perspectives," *Sensors*, vol. 21, no. 13, p. 4510, Jun. 2021.

[5] B. Hindawi and A. S. Abbas, "Congestion control techniques in 5G mm wave networks: A review," in *Proc. 1st Babylon Int. Conf. Inf. Technol. Sci. (BICITS)*, Apr. 2021, pp. 305–310.

[6] G. Dandachi, A. De Domenico, D. T. Hoang, and D. Niyato, "An artificial intelligence framework for slice deployment and orchestration in 5G networks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 2, pp. 858–871, Jun. 2020.

[7] S. Douch, M. R. Abid, K. Zine-Dine, D. Bouzidi, and D. Benhaddou, "Edge computing technology enablers: A systematic lecture study," *IEEE Access*, vol. 10, pp. 69264–69302, 2022.

[8] Y. B. Zikria, S. W. Kim, M. K. Afzal, H. Wang, and M. H. Rehmani, "5G mobile services and scenarios: Challenges and solutions," *Sustainability*, vol. 10, no. 10, p. 3626, Oct. 2018.

[9] S. Han, "Congestion-aware WiFi offload algorithm for 5G heterogeneous wireless networks," *Comput. Commun.*, vol. 164, pp. 69–76, Dec. 2020.

[10] I. A. Najm, A. K. Hamoud, J. Lloret, and I. Bosch, "Machine learning prediction approach to enhance congestion control in 5G IoT environment," *Electronics*, vol. 8, no. 6, p. 607, May 2019.

[11] R. Sathya and A. Abraham, "Comparison of supervised and unsupervised learning algorithms for pattern classification," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 2, pp. 34–38, 2013.

[12] I. Khan, M. Zafar, M. Jan, J. Lloret, M. Basheri, and D. Singh, "Spectral and energy efficient low-overhead uplink and downlink channel estimation for 5G massive MIMO systems," *Entropy*, vol. 20, no. 2, p. 92, Jan. 2018.

[13] G. Sangeetha, M. Vijayalakshmi, S. Ganapathy, and A. Kannan, "A heuristic path search for congestion control in WSN," in *Proc. Int. Conf. Ind. Interact. Innov. Sci., Eng. Technol. (I3SET)*. Singapore: Springer, 2016, pp. 485–495.

[14] K. Singh, K. Singh, L. H. Son, and A. Aziz, "Congestion control in wireless sensor networks by hybrid multi-objective optimization algorithm," *Comput. Netw.*, vol. 138, pp. 90–107, Jun. 2018.

[15] S. F. Ahmed, M. S. B. Alam, S. Afrin, S. J. Rafa, S. B. Taher, M. Kabir, S. M. Muyeen, and A. H. Gandomi, "Toward a secure 5G-enabled Internet of Things: A survey on requirements, privacy, security, challenges, and opportunities," *IEEE Access*, vol. 12, pp. 13125–13145, 2024.

[16] S. Urooj, R. Arunachalam, M. A. Alawad, K. N. Tripathi, D. Sukumaran, and P. Ilango, "An effective model for network selection and resource allocation in 5G heterogeneous network using hybrid heuristic-assisted multi-objective function," *Expert Syst. Appl.*, vol. 248, Aug. 2024, Art. no. 123307.

[17] R. MacDavid, X. Chen, and J. Rexford, "Scalable real-time bandwidth fairness in switches," *IEEE/ACM Trans. Netw.*, vol. 32, no. 2, pp. 1423–1434, Apr. 2024.

[18] M.-R. Fida, A. H. Ahmed, T. Dreibholz, A. F. Ocampo, A. Elmokashfi, and F. I. Michelinakis, "Bottleneck identification in cloudified mobile networks based on distributed telemetry," *IEEE Trans. Mobile Comput.*, vol. 23, no. 5, pp. 5660–5676, May 2024.

[19] M. Shelke, A. Malhotra, and P. N. Mahalle, "Congestion-aware opportunistic routing protocol in wireless sensor networks," in *Proc. 1st Int. Conf. Smart Comput. Inform. (SCI)*, vol. 1. Singapore: Springer, 2016, pp. 63–72.
[20] P. D. Godoy, R. L. Cayssials, and C. G. García Garino, "Communication channel occupation and congestion in wireless sensor networks," *Comput. Electr. Eng.*, vol. 72, pp. 846–858, Nov. 2018.
[21] I. A. Najm, M. Ismail, J. Lloret, K. Z. Ghafoor, B. B. Zaidan, and A. A.-R.-T. Rahem, "Improvement of SCTP congestion control in the LTE-A network," *J. Netw. Comput. Appl.*, vol. 58, pp. 119–129, Dec. 2015.
[22] S. Brahma, M. Chatterjee, and K. Kwiat, "Congestion control and fairness in wireless sensor networks," in *Proc. 8th IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PERCOM Workshops)*, Mar. 2010, pp. 413–418.
[23] I. A. Najm, J. M. Dahr, A. K. Hamoud, A. S. Hashim, W. A. Awadh, M. B. M. Kamel, and A. M. Humadi, "OLAP mining with educational data mart to predict students' performance," *Informatica*, vol. 46, no. 5, pp. 11–19, Mar. 2022.
[24] J. M. Dahr, A. K. Hamoud, I. A. Najm, and M. I. Ahmed, "Implementing sales decision support system using data mart based on OLAP, KPI, and data mining approaches," *J. Eng. Sci. Technol.*, vol. 17, no. 1, pp. 275–293, 2022.
[25] M. Al-Asfoor and M. H. Abed, "Deep learning approach for COVID-19 diagnosis using X-ray images," in *Proc. Int. Conf. Inf. Technol. Appl. (ICITA)*. Singapore: Springer, 2021, pp. 161–170.
[26] H. K. Naji, H. K. Fatlawi, A. J. M. Karkar, N. Goga, A. Kiss, and A. T. Al-Rawi, "Prediction of COVID-19 patients recovery using ensemble machine learning and vital signs data collected by novel wearable device," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 1–10, 2022.
[27] H. K. Fatlawi and A. Kiss, "An adaptive classification model for predicting epileptic seizures using cloud computing service architecture," *Appl. Sci.*, vol. 12, no. 7, p. 3408, Mar. 2022.
[28] A. K. Hamoud, A. S. Alasady, W. A. Awadh, J. M. Dahr, M. B. M. Kamel, A. M. Humadi, and I. A. Najm, "A comparative study of supervised/unsupervised machine learning algorithms with feature selection approaches to predict student performance," *Int. J. Data Mining, Model. Manage.*, vol. 15, no. 4, pp. 393–409, 2023.
[29] H. K. Fatlawi and A. Kiss, "Handling delayed labeling of EEG data stream using semi-supervised label propagation," in *Proc. 15th Int. Conf. Electron., Comput. Artif. Intell. (ECAI)*, Jun. 2023, pp. 1–5.
[30] A. K. Hamoud, M. B. M. Kamel, A. S. Gaafar, A. S. Alasady, A. M. Humadi, W. A. Awadh, and J. M. Dahr, "A prediction model based machine learning algorithms with feature selection approaches over imbalanced dataset," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 28, no. 2, p. 1105, Nov. 2022.
[31] S. Al-yousif, A. Jaenul, W. Al-Dayyeni, A. Alamoodi, I. Najm, N. M. Tahir, A. A. A. Alrawi, Z. Cömert, N. A. Al-shareefi, and A. H. Saleh, "A systematic review of automated pre-processing, feature extraction and classification of cardiotocography," *PeerJ Comput. Sci.*, vol. 7, pp. 1–37, Apr. 2021.
[32] I. A. Najm, M. Ismail, T. Rahem, and A. Al-Razak, "Wireless implementation selection in higher institution learning environment," *J. Theor. Appl. Inf. Technol.*, vol. 67, no. 2, pp. 477–484, 2014.
[33] H. K. Fatlawi and A. Kiss, "An elastic self-adjusting technique for rare-class synthetic oversampling based on cluster distortion minimization in data stream," *Sensors*, vol. 23, no. 4, p. 2061, Feb. 2023.
[34] P. Geurts, I. El Khayat, and G. Leduc, "A machine learning approach to improve congestion control over wireless computer networks," in *Proc. 4th IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2004, pp. 383–386.
[35] S. Jagannathan and K. C. Almeroth, "Using tree topology for multicast congestion control," in *Proc. Int. Conf. Parallel Process.*, Sep. 2001, pp. 313–320.
[36] X. Zhang, J. Zuo, Z. Huang, Z. Zhou, X. Chen, and C. Joe-Wong, "Learning with side information: Elastic multi-resource control for the open RAN," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 2, pp. 295–309, Feb. 2024.
[37] V. Murgai, V. Kanakaraj, and I. Kommineni, "AI in the wireless 5G core (5GC)," in *AI in Wireless for Beyond 5G Networks* (5GC). Boca Raton, FL, USA: CRC Press, 2023, pp. 147–154.
[38] Y. Watanabe, Y. Kawamoto, and N. Kato, "A novel routing control method using federated learning in large-scale wireless mesh networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9291–9300, Dec. 2023.
[39] A. Sunny, S. Panchal, N. Vidhani, S. Krishnasamy, S. V. R. Anand, M. Hegde, J. Kuri, and A. Kumar, "A generic controller for managing TCP transfers in IEEE 802.11 infrastructure WLANs," *J. Netw. Comput. Appl.*, vol. 93, pp. 13–26, Sep. 2017.
[40] R. Katuwal, P. N. Suganthan, and L. Zhang, "An ensemble of decision trees with random vector functional link networks for multi-class classification," *Appl. Soft Comput.*, vol. 70, pp. 1146–1153, Sep. 2018.
[41] S. E. Gómez, B. C. Martínez, A. J. Sánchez-Esguevillas, and L. Hernández Callejo, "Ensemble network traffic classification: Algorithm comparison and novel ensemble scheme proposal," *Comput. Netw.*, vol. 127, pp. 68–80, Nov. 2017.
[42] B. Leng, L. Huang, C. Qiao, and H. Xu, "A decision-tree-based on-line flow table compressing method in software defined networks," in *Proc. IEEE/ACM 24th Int. Symp. Quality Service (IWQoS)*, Jun. 2016, pp. 1–2.
[43] D. Liang, Z. Zhang, and M. Peng, "Access point reselection and adaptive cluster splitting-based indoor localization in wireless local area networks," *IEEE Internet Things J.*, vol. 2, no. 6, pp. 573–585, Dec. 2015.
[44] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, Jun. 2013.
[45] Y. Liu and H. Wu, "Prediction of road traffic congestion based on random forest," in *Proc. 10th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 2, Dec. 2017, pp. 361–364.
[46] H. Park, A. Haghani, S. Samuel, and M. A. Knodler, "Real-time prediction and avoidance of secondary crashes under unexpected traffic congestion," *Accident Anal. Prevention*, vol. 112, pp. 39–49, Mar. 2018.
[47] J. Shu, S. Liu, L. Liu, L. Zhan, and G. Hu, "Research on link quality estimation mechanism for wireless sensor networks based on support vector machine," *Chin. J. Electron.*, vol. 26, no. 2, pp. 377–384, Mar. 2017.
[48] A. C. Riekstin, G. C. Januário, B. B. Rodrigues, V. T. Nascimento, T. C. M. B. Carvalho, and C. Meirosu, "Orchestration of energy efficiency capabilities in networks," *J. Netw. Comput. Appl.*, vol. 59, pp. 74–87, Jan. 2016.
[49] S. Soltani and M. W. Mutka, "Decision tree modeling for video routing in cognitive radio mesh networks," in *Proc. IEEE 14th Int. Symp. World Wireless, Mobile Multimedia Networks (WoWMoM)*, Jun. 2013, pp. 1–9.
[50] T. Stimpfling, N. Bélanger, O. Cherkaoui, A. Béliveau, L. Béliveau, and Y. Savaria, "Extensions to decision-tree based packet classification algorithms to address new classification paradigms," *Comput. Netw.*, vol. 122, pp. 83–95, Jul. 2017.
[51] D. Singh, S. P. Nigam, V. P. Agrawal, and M. Kumar, "Vehicular traffic noise prediction using soft computing approach," *J. Environ. Manage.*, vol. 183, pp. 59–66, Dec. 2016.
[52] Y. Xia, W. Chen, X. Liu, L. Zhang, X. Li, and Y. Xiang, "Adaptive multimedia data forwarding for privacy preservation in vehicular ad-hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2629–2641, Oct. 2017.
[53] E. Adi, Z. Baig, and P. Hingston, "Stealthy denial of service (DoS) attack modelling and detection for HTTP/2 services," *J. Netw. Comput. Appl.*, vol. 91, pp. 1–13, Aug. 2017.
[54] B. Tierney, E. Kissel, M. Swany, and E. Pouyoul, "Efficient data transfer protocols for big data," in *Proc. IEEE 8th Int. Conf. E-Sci.*, Oct. 2012, pp. 1–9.
[55] M. F. Mohamed, A. E.-R. Shabayek, M. El-Gayyar, and H. Nassar, "An adaptive framework for real-time data reduction in AMI," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 31, no. 3, pp. 392–402, Jul. 2019.
[56] T. N. D. Pham and C. K. Yeo, "Adaptive trust and privacy management framework for vehicular networks," *Veh. Commun.*, vol. 13, pp. 1–12, Jul. 2018.
[57] Z. Md. Fadlullah, F. Tang, B. Mao, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, "State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2432–2455, 4th Quart., 2017.
[58] Y. Kong, H. Zang, and X. Ma, "Improving TCP congestion control with machine intelligence," in *Proc. Workshop Netw. Meets AI ML NetAI*, 2018, pp. 60–66.
[59] N. Taherkhani and S. Pierre, "Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3275–3285, Nov. 2016.

[60] Y.-Y. Chen, Y. Lv, Z. Li, and F.-Y. Wang, "Long short-term memory model for traffic congestion prediction with online open data," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 132–137.

[61] F. Tariq and S. Baig, "Multiclass machine learning based botnet detection in software defined networks," *Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 3, p. 150, 2019.

[62] T. Wu, S. Petrangeli, R. Huysegems, T. Bostoen, and F. De Turck, "Network-based video freeze detection and prediction in HTTP adaptive streaming," *Comput. Commun.*, vol. 99, pp. 37–47, Feb. 2017.

[63] T. Abar, A. Ben Letaifa, and S. El Asmi, "Machine learning based QoE prediction in SDN networks," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2017, pp. 1395–1400.

[64] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2022.

[65] C. Crisci, B. Ghattas, and G. Perera, "A review of supervised machine learning algorithms and their applications to ecological data," *Ecol. Model.*, vol. 240, pp. 113–122, Aug. 2012.

[66] N. Ye, *Data Mining: Theories, Algorithms, and Examples*. Boca Raton, FL, USA: CRC Press, 2013.

[67] B. Çığşar and D. Ünal, "Comparison of data mining classification algorithms determining the default risk," *Sci. Program.*, vol. 2019, pp. 1–8, Feb. 2019.

[68] A. Hamoud, "Selection of best decision tree algorithm for prediction and classification of students' action," *Amer. Int. J. Res. Sci., Technol., Eng. Math.*, vol. 16, no. 1, pp. 26–32, 2016.

[69] A. Hamoud, "Applying association rules and decision tree algorithms with tumor diagnosis data," *Int. Res. J. Eng. Technol.*, vol. 3, no. 8, pp. 27–31, 2017.

[70] A. K. Hamoud, A. S. Hashim, and W. A. Awadh, "Predicting student performance in higher education institutions using decision tree analysis," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 5, no. 2, pp. 26–31, 2018.

[71] A. K. Hamoud and A. M. Humadi, "Student's success prediction model based on artificial neural networks (ANN) and a combination of feature selection methods," *J. Southwest Jiaotong Univ.*, vol. 54, no. 3, pp. 1–19, Jun. 2019.

[72] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, Mar. 1996.

[73] W. S. Noble, "What is a support vector machine?" *Nature Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006.

[74] G. Fung, "A comprehensive overview of basic clustering algorithms," Dept. Comput. Sci., Univ. Wisconsin, Madison, WI, USA, Tech. Rep., 2001. [Online]. Available: https://pages.cs.wisc.edu/~gfung/

[75] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, no. 3, pp. 241–254, Sep. 1967.

[76] Z. Zhang, J. Zhang, and H. Xue, "Improved K-means clustering algorithm," in *Proc. Congr. Image Signal Process.*, May 2008, pp. 169–172.

[77] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, 1996, vol. 96, no. 34, pp. 226–231.

[78] P. Rai and S. Singh, "A survey of clustering techniques," *Int. J. Comput. Appl.*, vol. 7, no. 12, pp. 1–5, Oct. 2010.

[79] V. R. Gannapathy, R. Nordin, N. F. Abdullah, and A. Abu-Samah, "A smart handover strategy for 5G mmWave dual connectivity networks," *IEEE Access*, vol. 11, pp. 134739–134759, 2023.

[80] V. K. Quy, A. Chehri, N. M. Quy, N. D. Han, and N. T. Ban, "Innovative trends in the 6G era: A comprehensive survey of architecture, applications, technologies, and challenges," *IEEE Access*, vol. 11, pp. 39824–39844, 2023.

[81] M. Mezzavilla, M. Zhang, M. Polese, R. Ford, S. Dutta, S. Rangan, and M. Zorzi, "End-to-end simulation of 5G mmWave networks," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2237–2263, 3rd Quart., 2018.

[82] M. Rebato, M. Polese, and M. Zorzi, "Multi-sector and multi-panel performance in 5G mmWave cellular networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[83] A. A. Oliveira, D. Batista, and R. Hirata, "Exploring the ns-3 mmWave module," Dept. Comput. Sci., Univ. São Paulo, São Paulo, Brazil, Tech. Rep., 2019, p. 23. [Online]. Available: http://vision.ime.usp.br/~arturao/

[84] J. M. Bland and D. G. Altman, "Statistics notes: Cronbach's alpha," *BMJ*, vol. 314, no. 7080, p. 572, 7080.

[85] M. Tavakol and R. Dennick, "Making sense of Cronbach's alpha," *Int. J. Med. Educ.*, vol. 2, pp. 53–55, Jun. 2011.

[86] A. Khalaf, A. Majeed, W. Akeel, and A. Salah, "Students' success prediction based on Bayes algorithms," *Int. J. Comput. Appl.*, vol. 178, no. 7, pp. 6–12, Nov. 2017.

[87] J. N. Mandrekar, "Receiver operating characteristic curve in diagnostic test assessment," *J. Thoracic Oncol.*, vol. 5, no. 9, pp. 1315–1316, Sep. 2010.

[88] D. W. Hosmer Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 398. Hoboken, NJ, USA: Wiley, 2013.

[89] H. Uğuz, "A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm," *Knowl.-Based Syst.*, vol. 24, no. 7, pp. 1024–1032, Oct. 2011.

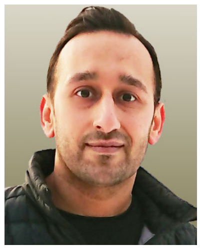
[90] C. Lee and G. G. Lee, "Information gain and divergence-based feature selection for machine learning-based text categorization," *Inf. Process. Manage.*, vol. 42, no. 1, pp. 155–165, Jan. 2006.

[91] B. Suri, Mani, and M. Kumar, "Performance evaluation of data mining techniques," in *Proc. Inf. Commun. Technol. for Sustain. Develop. (ICT4SD)*, vol. 1. Singapore: Springer, 2016, pp. 375–383.

[92] G. H. John and P. Langley, "Estimating continuous distributions in Bayesian classifiers," 2013, *arXiv:1302.4964*.

[93] Q. Wang, G. M. Garrity, J. M. Tiedje, and J. R. Cole, "Naiïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy," *Appl. Environ. Microbiol.*, vol. 73, no. 16, pp. 5261–5267, Aug. 2007.

[94] O. J. Oyelade, O. O. Oladipupo, and I. C. Obagbuwa, "Application of *k* means clustering algorithm for prediction of students academic performance," 2010, *arXiv:1002.2425*.



MOHAMMED B. M. KAMEL received the Ph.D. degree in computer science from Eötvös Loránd University and Furtwangen University. He is currently a Senior Researcher and a Certified Cybersecurity Consultant. He was a part of several projects as a cybersecurity researcher that have been accomplished. His research interest includes designing distributed secure protocols. He was a Gold and Bronze Award Winner from EIT, a body of the European Union.

IHAB AHMED NAJM received the Diploma degree from the Technical University of Berlin, Berlin, Germany, in 2012, the B.S. and M.S. degrees in information technology from the Utara University of Kedah Malaysia, in 2012, and the Ph.D. degree from the Department of Electrical, Electronic and Systems Engineering, National University of Malaysia, in 2017. From 2004 to 2007, he was a Laboratory Instructor with the Department of Computer Science, University of Tikrit. He is the author of more than 20 articles. His research interests include broadcasting, postal communication, tele-education, teleworking, multiservice convergence networks, cellular networks and the IoT, data science, machine learning and deep learning, internet protocol, short-range communication, VoIP, teleport and tele-health, digital communications networks, area of expertise smart mobile wireless communication, and future internet protocols.

ALAA KHALAF HAMOUD received the B.Sc. and M.Sc. degrees (Hons.) from the Department of Computer Science, University of Basrah, Iraq, in 2008 and 2014, respectively. He is currently an Assistant Professor with the Department of Cybersecurity, University of Basrah. He participated in (seven months) IT administration course at the Technical University of Berlin, Germany. His scientific research interests include data mining and data warehousing.