International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
Article Received: 08 April 2023 Revised: 29 May 2023 Accepted: 16 June 2023

A Novel Chimp Optimized Linear Kernel Regression (COLKR) Model for Call Drop Prediction in Mobile Networks

Ashok G V[1], Dr. Vasanthi Kumari P.[2]
[1]Department of Computer Science and Engineering
Dayananda Sagar University
Bangalore, Karnataka, India
[2]Professor, Department of Computer Applications
Dayananda Sagar University
Bangalore, Karnataka, India

Abstract— Call failure can be caused by a variety of factors, including inadequate cellular infrastructure, undesirable system structuring, busy mobile phone towers, changing between towers, and many more. Outdated equipment and networks worsen call failure, and installing more towers to improve coverage might harm the regional ecosystems. In the existing studies, a variety of machine learning algorithms are implemented for call drop prediction in the mobile networks. But it facing problems in terms of high error rate, low prediction accuracy, system complexity, and more training time. Therefore, the proposed work intends to develop a new and sophisticated framework, named as, Chimp Optimized Linear Kernel Regression (COLKR) for predicting call drops in the mobile networks. For the analysis, the Call Detail Record (CDR) has been collected and used in this framework. By preprocessing the attributes, the normalized dataset is constructed using the median regression-based filtering technique. To extract the most significant features for training the classifier with minimum processing complexity, a sophisticated Chimp Optimization Algorithm (COA) is applied. Then, a new machine learning model known as the Linear Kernel Regression Model (LKRM) has been deployed to predict call drops with greater accuracy and less error. For the performance assessment of COLKR, several machine learning classifiers are compared with the proposed model using a variety of measures. By using the proposed COLKR mechanism, the call drop detection accuracy is improved to 99.4%, and the error rate is reduced to 0.098%, which determines the efficiency and superiority of the proposed system.

Keywords- Call dropout, chimp optimization algorithm, data preprocessing, linear kernel regression, mobile networks, machine learning, service provider.

I. INTRODUCTION

In present days, call dropouts in mobile networks have emerged as a new highly topical concern in the public discourse following deliberations about net neutrality [1], [2]. Call drops occur when a call is cut off before the users opt to end it, regardless of if they have phoned another user or someone else has called them on their mobile device. The expected level of quality of service (QoS) in a mobile wireless network is directly impacted by a dropped call [3]. Call drops are more problematic than blocked calls because there is no one waiting in line to re-establish the connection when a call drops. With fixed line networks, call drops are uncommon. Since the mobile wireless network is only an extension of the fixed line network, the majority of customers anticipate identical performance levels between mobile networks and fixed-line networks. The majority of people's life will revolve on the services offered by mobile wireless networks in the future [4], [5], hence the desire for better service quality will only grow. The user will incur additional costs as a result of having to call again in order to complete the conversation, in addition to the inconvenience that the dropped call would entail. If service providers take certain precautions, such as balancing traffic among the various frequency tiers, reducing interference and traffic, and expanding the service area, call drops are often avoidable [6]. Consequently, within the available spectrum, strengthening network infrastructure and implementing technological solutions to reduce call dropouts are the only ways to enhance the level of service in mobile networks. A considerable increase in network traffic is anticipated today with the growing use of cellular networks [7].

Call lengths on the phone are increasing, which is an alarming statistic. In networks, call drops—calls that are "dropped" (stopped) without any of the parties purposefully disconnecting the calls—are a typical occurrence. Call drop prevention [8] enhances service quality. Self-organized networks are a technique that makes complex networks simpler and guarantees better service. A major factor in it is Machine Learning (ML) [9], which is a subset of artificial intelligence that helps to comprehend actions and see patterns beyond human recognition, repairs the holes left by human limitations. ML allows technologies to improve the learning capacity without having to be explicitly programmed for performing particular tasks [10], [11]. It is highly regarded because it can learn from data, discover patterns, and make decisions without much or no human involvement in an era where vast volumes of data are readily available. Supervised learning (SL), unsupervised learning (UL), and reinforcement learning (RL) are the three categories into which machine learning [12]-[14] can be separated. In a supervised model, the algorithm learns from a labelled dataset that offers a better solution key the algorithm may use to gauge its efficiency on the dataset. In contrast, unsupervised models provide the computer unlabeled input to try and make sense of by extracting features and patterns on its own. In the reinforcement learning algorithm, the decision-making agent rely on both analyzing prior experience and investigating novel strategies that might offer a greater reward. There has been study on using ML modelling for data-centric networks. However, no comprehensive research has been conducted to determine whether ML would be a better choice than traditional techniques in anticipating and predicting a call drop in mobile networks. But, some of the works have been developed in the field for call drop and network failure prediction using machine learning [15], [16]. However, it faces issues such as a high error rate, complexity in system design, inability to handle huge datasets, and increased time requirements. As a result, the proposed study will employ an effective and sophisticated ML for call drop prediction. The main purpose of this research work is to develop a new ML model incorporated with the optimization model for predicting call drops in the mobile network with better performance.

The key objectives of this paper are as follows:

- To implement the system, a real time dataset, named as, CDR has been generated with the fields of String ID, A Party Calling Number, B Party Called number, Calling Time Stamp, Called Time Stamp, Event Code, Charge, Network Status, Call Result, Service Provider, Service Provider Code and Call Status.
- The median regression-based filtering technique is used to construct the normalized dataset by preprocessing the attributes.
- A sophisticated Chimp Optimization Algorithm (COA) was used to extract the most significant features for training the classifier with low processing complexity.
- A new machine learning model called as, Linear Kernel Regression Model (LKRM) has been deployed to reliably forecast call drops with a lower error rate and higher accuracy. To accurately predict the call drops with reduced error rate and increased accuracy, a new machine learning model, named as, Linear Kernel Regression Model (LKRM) has been deployed.
- A comprehensive analysis is performed to validate the performance and compare the outcomes of the proposed Chimp Optimized Linear Kernel Regression (COLKR) model.

The remainder of this work is divided into the following sections: Section II contains an extensive literature review that looks at the many types of machine learning algorithms used in mobile networks for call drop prediction. Section III provides a detailed and concise explanation of the proposed COLKR model. Section IV validates and assesses the performance and comparison outcomes using CDR data. Section V concludes the whole study by summarizing the findings, outcomes, and future scope.

II. RELATED WORKS

This section looks into different machine learning methods that help reduce call dropouts and enable reliable communication in wireless and mobile networks. For better comprehension, it also looks into the difficulties and limitations of the previous studies.

Ozturk *et al.* [17] developed a new architecture, named as, Control/Data Separation Architecture (CDSA) for minimizing call drop rate in cellular networks. Here, the low-cost mobility prediction model is implemented with the use of deep learning algorithm, termed as, Stacked Long Short-Term Memory (LSTM). The contribution of this paper was to reduce the cost function with ensure mobility management. In addition, a unique analytical model is put forth for the holistic handover (HO) cost evaluation that focuses on improving the parameters of transmission overhead, latency, call dropping, and dynamic resource usage. Many factors, including user profile, data flow, signal strength, and call flow rate, affect the likelihood of a call dropping. For instance, the probability of a call dropping is higher in the urban environment than in the rural environment because of greater network density, greater volume of data, and also more limited resource availability. Sun *et al.* [18] investigated about the issues and challenges in the deployment of machine learning techniques in wireless networks. Normally, a balance must be made between the user experience impacted by call dropping ratio and the signaling cost paid by handover. Mobility management is crucial for optimal service delivery in wireless networks. Recent developments in user mobility prediction, handover parameter optimization, and other areas have demonstrated the enormous benefits of machine learning. To further enhance network efficiency, it is standard procedure in wireless networking to group nodes or users into various clusters for internal collaboration or coordination. Abad, *et al*. [19] deployed a new algorithm, named as, Federated Edge Learning (FEL) for improving the performance of heterogeneous cellular networks with reduced call drop rate, and bandwidth utilization. In order to reduce the end-to-end latency, the authors used a resource allocation method and communication-efficient distributed learning techniques, including sparsification and periodic averaging.

Ahmad *et al*. [20] applied an eXtreme Gradient Boost (XGB) algorithm for churn prediction in the telecom sector. Here, the different types of data mining techniques such as preprocessing, feature transformation, selection, and classification for churn prediction. Organizations are working to create methods to predict probable customer churn because it has a direct impact on their revenues, particularly in the telecom industry. In order to reduce customer churn, it is crucial to identify the variables that contribute to this churn. This work's key contribution is the creation of a churn prediction model that helps telecom providers identify customers who are most likely to experience churn. Luo *et al*. [21] suggested an AI model for making an intelligent decision in the mobile networks to enable a reliable communication. Typically, the AI mimics an intelligent human behavior in machines, which has achieved tremendous success across a wide range of application fields. Moreover, the suggested work recommends that the AI has the capacity to understand the state of mobile networks for making wise decisions. As a result of the utilization of packet switching, the mobile network system is impacted by the functionality of individual nodes and protocols. In order to make wise and effective decisions, it is crucial to jointly understand user equipment, the wireless spectrum, and wireless service components. Moysen *et al*. [22] constructed a self-organized network management scheme with the use of machine learning technique. The authors mainly focused on reducing call drops by avoiding unwanted hand overs. Moreover, the interference among the cells has been minimized with the use of inter-cell coordination scheme. The suggested paradigm attempts to give end users a high-quality experience with no negative environmental impact. In order to reduce energy usage, Network Elements (NEs) are being designed with lower power requirements, and when not in use, unused capacity or nodes are being momentarily turned off.

Anand *et al*. [23] introduced a new resource allocation mechanism for reducing traffic rate in 5G networks. The authors intended to reduce the bandwidth consumption in the network by effectively allocating the resources. Asghar *et al.* [24] investigated about the different properties of cellular networks, which includes self-configuration, self-organization, and self-healing. Due to the limited ability of human experts to process vast amounts of network data simultaneously and draw conclusions about the existence of outages in the mobile network, the network monitoring and controlling operations are very essential. Also, it will increase the operational costs of operators. By using methods for the detection of service degradations and disruptions, the self-healing can lessen the burden on human specialists. Small cell-specific solutions must be created because they are anticipated to make up a sizable portion of the cellular network architecture in the future. This problem is increased by the sparse reporting that tiny cells are susceptible to because of the relatively few consumers that they typically have and the more densely packed mobile cellular network topology in terms of inter-node distances.

As a result, it is more challenging to detect service outages at small cells using the conventional techniques. The number of design and optimization factors is anticipated to increase dramatically with the development of new technologies. The likelihood of parameter misconfiguration might considerably increase when network control parameters and entities proliferate. The proposed approach makes use of level functions to continually monitor downlink signal parameters such channel quality, call drop rate, and handover timing advance in order to determine whether a cell drops below the accepted level established by expert systems. The authors have shown that the suggested method can react in close to real time by discovering outages within a few minutes of their occurrence, which is a significant improvement over the detection rate by human specialists, specifically in very complex systems. Unsupervised learning techniques are extremely well-liked in applications for interruption detection due to their special capacity to sort data into various groups without any prior pre-classification. Unsupervised learning is widely used to identify sleeping cells, or those that are experiencing an outage but are not producing any alerts. Because of the absence of alerts that come along with the interruption, it is not feasible to identify such cells manually instantly, making the detection of such cells a particularly beneficial application of unsupervised learning. Hadi *et al*. [25] presented a comprehensive survey to examine the different types of big data analytics models for both wired and wireless networks. One of the most aggravating situations is when a mobile subscriber is taken off guard by an unexpected call drop. Many of these incidents take place while a user is moving away from one technologically advanced location, such as leaving a base station, and is at the edge of that area's coverage. Pustokhina *et al*. [26] implemented a multi-objective

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
***Article Received**: 08 April 2023 **Revised:** 29 May 2023 **Accepted:** 16 June 2023*

rain optimization technique to effectively predict churn from the telecommunication sector. Here, the rain optimization technique is mainly used for tuning the parameters of classifier, which supports to obtain the better churn prediction accuracy. However, the suggested optimization has the drawbacks of low convergence rate, and high searching complexity. In paper [27], the churn prediction analysis is performed with the use hybrid optimization-based learning mechanism. Mishra *et al*. [28] deployed a self-optimization algorithm for avoiding call drops in the mobile communication networks. Here, the handover optimization is performed with increased robustness and reduced call drops.

Based on the findings of this study, it has been determined that numerous machine learning and deep learning algorithms have been created in the literature for reducing call dropouts in mobile networks. Unfortunately, the majority of the methods have the following shortcomings:

- Inefficient outcomes as a result of variable dependencies.
- Dataset training procedure takes a long time.
- Overfitting problem.
- Reduced effectiveness.
- Very susceptible to noisy data.
- Not more suitable for handling huge-dimensional dataset, since the type of dataset has a significant impact on the classifier's accuracy.

Thus, the goal of the proposed work is to develop a new as well as sophisticated machine learning model for preventing call dropouts in a mobile network.

III. PROPOSED METHODOLOGY

This section gives a brief description of the proposed Chimp Optimized Linear Regression Model (COLKR) for predicting call drops in mobile networks. This work's original contribution is the implementation of a revolutionary machine learning technique for preventing call drops in the network while improving performance. For this purpose, a distinct Chimp Optimization Algorithm (COA) and Linear Kernel Regression Model (LKRM) techniques have been implemented in this work.

The novel contribution of this paper is to implement a cutting-edge optimization integrated regression model for avoiding call drops in the mobile networks. In the previous studies, the chimp optimization is used for solving some of the complex multi-objective engineering problems. But, in this work, we used this optimization technique to squeeze the dimensionality of the telecom dataset with optimized features. By using this algorithm, the efficacy of classifier is highly improved with proper training and testing results. When compared to the previous studies used in the field of call drop avoidance, the proposed COLKR technique provides the major benefits of reduced false predictions, accurate call drops identification, high system efficiency, and lower complexity.

Typically, the call dropouts have a significant influence on both voice and the web consumers as well as service providers' revenue. Unexpected call drops have an impact on both the work that must be done and the relationships that need to be upheld with various business customers. They also weaken trust in telecom service providers. There are numerous circumstances that can cause a call to be dropped in a scenario or use case. A subscriber began using telecom services provided by Service Provider Network (SPN). The work flow model of the proposed COLKR based call drop prediction framework is shown in Fig. 1, which comprises the following operations:

- Input telecom CDR data acquisition
- Median filtering based preprocessing
- COA based feature selection
- LKRM call drop prediction
- Performance analysis

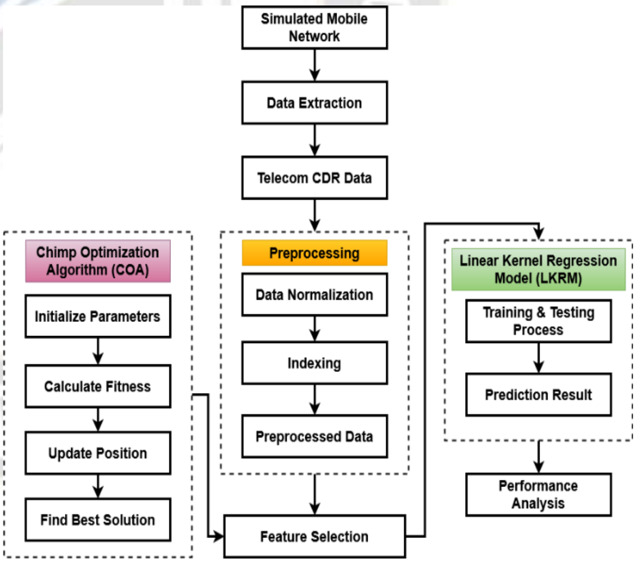

Figure 1. Workflow model

After getting the CDR data, the data cleaning and preprocessing operations are carried out to generate the normalized or noise free dataset for further processing. In the proposed framework, the input CDR is obtained with the call history and other information attributes, which is preprocessed before the identification of call drops. Since, the original dataset may comprise some irrelevant fields of information,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
***Article Received**: 08 April 2023 **Revised:** 29 May 2023 **Accepted:** 16 June 2023*

which may affect the performance and drop detection accuracy of the classifier. Hence, it is most essential to perform preprocessing before classification, in which the operations like normalization, indexing, and attribute balancing have been performed to generate the balanced dataset for an accurate call drop prediction. For this purpose, the median filtering technique is applied in the proposed framework, which preprocesses the given CDR Dataset with balanced attributes and fields of information. Consequently, the most important features are selected from the normalized dataset with the use of COA, which helps to improve the training of classification. Furthermore, the LKRM is applied to train the dataset for predicting call drops and network failures according to the features of CDR data. By using this operation, the reliable and effective communication has been established in the mobile networks with reduced call drops.

A. CDR Dataset Acquisition

Typically, acquiring real-time data from a particular organization or service provider can be difficult, we have devised a system to do so by leveraging cutting-edge tools and technologies. Using the coded methods and functions, the points of entry are created where the feature file can be initialized. The following fields are considered for creating CDR data; however, the framework may assist in the creation of new fields as needed. Moreover, the CDR Dataset includes the following fields:

- String ID
- Source calling number – A part
- Destination Called number – B Party
- Calling Time stamp
- Called Time stamp
- Event Code
- Charge
- Network Status
- Call Result
- Service Provider
- Service Provider Code
- Call Status

Then, the sample data pattern is provided as below:

String ID, A Party Calling Number, B Party Called number, Calling Time Stamp, Called Time Stamp, Event Code, Charge, Network Status, Call Result, Service Provider, Service Provider Code and Call Status

- 9143cb0d-a843-4dee-a7c5-fb2a3c41359f,9917618284,5084342972,2021-09-02T15:12:54.025+05:30,2021-09-02T15:15:53.524+05:30,0,6.97,NetworkConnected,0, Airtel,0,NoCallDrop
- b2bbdd5e-fdd2-45a1-863a-c227e70de01e,8131166797,3470914600,2021-08-19T23:19:42.950+05:30,2021-08-19T23:22:03.176+05:30,0,8.34,Network Connected,0, Airtel,0,NoCallDrop

B. Data Preprocessing

The input data values are first obtained from the CDR Dataset [29, 30], and then they are preprocessed, which may include noise removal or the replacement of missing data. The noiseless data aids in the detection of patterns linked with call drops. Because it explores the correlation between the variables in the dataset, the median filtering technique is used to eliminate undesirable or noisy data. This noise elimination approach improves the detection of call drops. The first step is to look over the data in the dataset and compute the average for the values that are missing.

The median value is calculated by organizing the data in increasing order and then calculating the mid value. The median value is used to replace missing and irrelevant variables. To make examining call drops easier, the data should be standardized within the range of 0 to 1 after missing values are removed. In the proposed framework, the data preprocessing is carried out with the use of a median filtering technique, where the median value is computed to replace the missing values and attributes. By using this technique, the balanced and attribute normalized dataset is generated, which is used for an accurate call drop prediction. For better prediction, the normalization method involves multiple data distributions and regression estimation. By using this algorithm, the data preprocessing and normalization processes are performed to generate the normalized CDR data for call drop prediction.

C. Chimp Optimization Algorithm (COA)

Following the normalization of the CDR data, a novel and intelligent Chime Optimization Algorithm (COA) has been employed to select the most important features for training and testing the classifier. A number of meta-heuristic models are used in existing studies to solve complicated optimization issues. Among other benefits, the COA offers a higher convergence rate, faster processing, lower complexity, and lower time consumption. As a result, the proposed study would employ this optimization technique to select the most relevant features from the CDR data. Typically, Chimps (also known as Chimpanzees) are one of only two African great ape species. They are the closest living ancestors of humans. Since they are descended from a single progenitor species (Hominoid) that lived approximately seven to eight million years ago, chimp and human DNA is so identical. The chimp colony is a society based on fission and fusion. This is a type

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
***Article Received**: 08 April 2023 **Revised:** 29 May 2023 **Accepted:** 16 June 2023*

of community in which the composition or size of the colonies fluctuates over time as members travel about the environment. Group composition is an evolving feature for chimps living in fission-fusion colonies. The major phases include in this optimization algorithm are as follows:

- Driving and pursuing prey
- Exploitation
- Attacking prey
- Exploration
- Social incentive

At first, the target is hunted for food during the exploration and exploitation stages, where the driving and hunting prey are mathematically simulated as shown in (1) and (2).

$$\mathrm{DP} = |\varrho \times \mathrm{CH_V(h)} - \psi \times \mathrm{Ch_V(h)}| \tag{1}$$

$$\mathrm{Ch_V(h+1)} = \mathrm{Ch_V(h)} - \beta \times \mathrm{DP} \tag{2}$$

Here, the coefficients vectors ϱ, ψ, β are computed by using the following models:

$$\varrho = 2\mathrm{q}\delta_1 - \mathrm{q} \tag{3}$$

$$\psi = \mathrm{CV} \tag{4}$$

$$\beta = 2\delta_2 \tag{5}$$

By using equ (3), the coefficient vector is estimated based on the parameters of δ_1 random number and Value of q is non-linearly reduced from 2.5 to 0 through iteration, where the value of δ_1 can be used in both exploitation and exploration operations. The effect of chimps' libido on their hunting behavior is represented by a chaotic vector that was derived using several chaotic maps. As shown in equ (1) to (5), the driving and chasing prey of chimp optimization is estimated according to the current iterations, coefficient vector value, position of prey, position of chimp, and coefficient vectors ϱ, ψ, β.

Moreover, the chaotic vector is computed using several chaotic maps to illustrate the effect of chimp desire in the hunting process. Two approaches are designed to mathematically model chimp attacking behavior: The chimps can investigate the location of the prey (by traveling, obstructing, and hunting) prior to encircle it. Typically, attacker chimps conduct the hunting process. Chimps serve as drivers, barriers, and chasers in the hunting process on sometimes. However, there is no information regarding the optimal placement in an abstract search space. During exploitation, the behavior of chimps is simulated for obtaining the best solutions as illustrated in (6).

$$\mathrm{Ch_V(\,h+1)} = \frac{\tau_1+\tau_2+\tau_3+\tau_4}{4} \tag{6}$$

By using this model, the four different positions such as attack, barrier, chaser and driver of chimps have been estimated. During exploration, they divide to seek prey and then aggregate to assault it. This chaotic behavior in the final stage aids chimps in alleviating the two challenges of trapping in local optima and slow convergence rate while solving high-dimensional problems. Then, the position updation is performed as shown in (7).

$$\mathrm{Ch}_V(h+1) = \{\mathrm{Ch}_V(h) - \tilde{n} \times \mathrm{DP} \text{ if } M > 0.5\mathrm{Ch}_V \text{ If } M > 0.5 \tag{7}$$

Moreover, the parameter and position updation have been performed to identify the best position of Chimp. At the end, the final best optimal solution is identified that helps to select the most important features from the CDR data.

D. Linear Kernel Regression Model (LKRM)

After choosing the features, the linear regression-based classification model is applied to accurately predict the call drops in the mobile network. It is a common statistical problem that involves determining a linear function that can model the relationships between predictors and response variables. When compared to the other machine learning algorithms, it is one of the most recent and novel classification model. Due to their enhanced performance rate and low time consumption, the Linear Kernel Regression Model (LKRM) is implemented in this work. The linear function has been computed using this method, which estimates the interdependence between the covariance and response variables. The ordinary least squares model is used to lessen the loss of factor for the provided set of training samples:

$$\{(p_i, q_i) \mid p_i \in \delta^N,\ q_i \in \delta^M,\ i = 1,\ 2,\ \cdots,\ X\}$$

as shown in (8):

$$\sum_i (q_i - \partial^T p_i)^2 \tag{8}$$

where p_i represents the training set, q_i is the vector, ∂^T and denotes the variance of estimate. The following model is used to illustrate the reduction in cost consumption $C(\partial)$ and squared error values through ridge regression as shown in (9):

$$C(\partial) = \sum_i \left\| q_i - \partial^T p_i \right\|^2 + \frac{\|\partial\|^2}{\tau} \tag{9}$$

where $/$ represents the regularization parameter, $\int$ indicates the regularization parameter, which is used to reduce the estimate variance by balancing the tradeoff among the bias and variance.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
***Article Received**: 08 April 2023 **Revised:** 29 May 2023 **Accepted:** 16 June 2023*

To find the best parameters for lowering the error values, cross validation is also carried out in this model. This method has allowed the kernel function to anticipate the multivariate label. After estimating the linear dependency, the estimated variance with the total cost function can be minimized using the above equation. Furthermore, the derivatives are obtained based on the following model:

$$\partial' = \left(PP^T + \frac{1}{\tau}\right)^{-1} PQ^T \tag{10}$$

Where, τ represents the regularization parameter, ∂' indicates the derivative of variance, PP^T and PQ^T are the random matrices.

According to the Mercer condition, the kernel matrix is constructed using the formula in (11):

$$\{\varphi_U = G^T \varphi_{U_{c,d}} = p_r q_s = \varphi(p_r, q_s) \tag{11}$$

where φ_U is the kernel matrix that is replaced with the value of matrix G^T, and T indicates the number of observations. The kernel function maps the input samples to the feature space input. Following that, the kernel function of the model is estimated as shown in (12).

$$\varphi(M, V) = \exp\left(-Y\|M - V\|^2\right) \tag{12}$$

where **M** represents the mean value, **V** denotes the vector value, and Y is the kernel parameter. Moreover, the output function is estimated according to the following model:

$$\Psi(p) = [\varphi(p, p_1) \vdots \varphi(p, pn)] \left(\frac{1}{\tau} + \varphi_U\right)^{-1} Q^T \tag{13}$$

$$\varphi_U^* = \left(\frac{1}{\tau} + \varphi_U\right)^{-1} Q^T \tag{14}$$

As shown in (13) and (14), the parameter $\tau\tau$ is set with the kernel function φ_U. Finally, the output layer produced the label as φ_U^*, which is used to predict the solution for call drop avoidance according to the features of CDR data.

IV. RESULTS AND DISCUSSION

Several measurements are used in this part to validate the performance and results of the proposed COLKR-based call drop prediction framework. Furthermore, the performance is evaluated and compared to the use of the CDR Dataset.

The novel contribution of this paper is to implement a cutting-edge optimization integrated regression model for avoiding call drops in the mobile networks. In the previous studies, the chimp optimization is used for solving some of the complex multi-objective engineering problems. But, in this work, we used this optimization technique to squeeze the dimensionality of the telecom dataset with optimized features. By using this algorithm, the efficacy of classifier is highly improved with proper training and testing results. When compared to the previous studies used in the field of call drop avoidance, the proposed COLKR technique provides the major benefits of reduced false predictions, accurate call drops identification, high system efficiency, and lower complexity. The Call Detail Record (CDR) is a mobile call record dataset, which comprises the fields the following fields of information String ID, Source calling number – A part, Destination Called number – B Party, Calling Time stamp, Called Time stamp, Event Code, Charge, Network Status, Call Result, Service Provider, Service Provider Code, and Call Status. With this information, the call history data pattern has been created that is used to predict the status of call drops. By using this dataset, the call drop is identified by the proposed COLKR model.

To demonstrate the superiority of the proposed COLKR model, some of the most current cutting-edge models are compared using the following parameters:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\%$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\%$$

$$F_{\text{1-score}} = \frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \times 100\%$$

where TP - true positives, TN - true negatives, FP - false positives, and FN - false negatives. During performance analysis, the results of the proposed COLKR model is validated and assessed by using a variety of measures. Typically, the evaluation measures such as accuracy, sensitivity, specificity, precision and f1-score are the widely used parameters, which helps to determine that how effectively the classifier could predict the results according to the attributes of the given dataset. Moreover, the aforementioned parameters must be improved for assuring better system efficacy and detection performance. Hence, these parameters are used in the proposed study to test and validate the call drop detection performance of the proposed COLKR model.

Table I and Fig. 2 compare the accuracy of call drop prediction using CDR data of the standard and proposed

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
***Article Received**: 08 April 2023 **Revised:** 29 May 2023 **Accepted:** 16 June 2023*

techniques. The machine learning model's increased prediction ability can be measured using its level of accuracy. Based on the estimated results, it is concluded that the suggested COLKR model outperforms typical machine learning techniques in terms of accuracy (99%). As a result, the methodologies' error rates are validated and compared, as shown in Table II and Fig. 3. Since the increased error rate can degrade the classifier's overall prediction performance. To ensure better results, the classifier's error rate should be decreased by effective training and testing. When compared to the other models, the estimated results show that the COLKR model has a much lower error rate of 0.098. The classifier's training and testing operations are effectively performing due to effective data normalization and correct feature optimization. As a result, when compared to existing classifiers, the proposed COLKR produces better results.

Table I: Performance analysis based on accuracy using CDR data

Techniques	Accuracy (%)
Support Vector Machine (SVM) [31]	88.7
K-Nearest Neighbor (KNN) [32]	89.1
Decision Tree (DT) [33]	90
Logistic Regression (LR) [34]	95
Naïve Bayes (NB) [35]	91
Random Forest (RF) [36]	93
Ensemble Learning (EL) [37]	92.6
Extreme Gradient Boost (XGB) [38]	89.7
AdaBoost [39]	88.4
Proposed	99.4

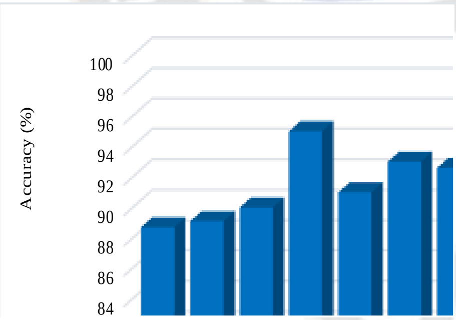

Fig 2. Accuracy analysis

Table II: Performance analysis based on error rate

Techniques	Error rate
Support Vector Machine (SVM)	0.36
K-Nearest Neighbor (KNN)	0.33
Logistic Regression (LR)	0.255
Naïve Bayes (NB)	0.319
Random Forest (RF)	0.238
Ensemble Learning (EL)	0.24
Extreme Gradient Boost (XGB)	0.39
AdaBoost	0.42
Proposed	0.098

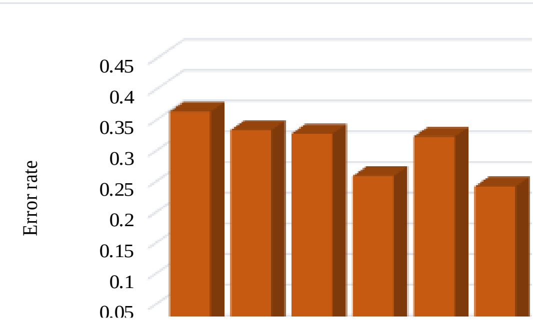

Figure 3. Error rate analysis

Table III: Overall performance analysis

Techniques	Sensitivity %)	Specificity %)	F1-score (%)
Support Vector Machine (SVM)	88.1	88.3	88
K-Nearest Neighbor (KNN)	88.9	88.5	88.2
Decision Tree (DT)	90.2	89.8	89.3
Logistic Regression (LR)	94.6	94.1	94.22
Naïve Bayes (NB)	90.7	91.1	90.3
Random Forest (RF)	92.9	92.5	91.9
Ensemble Learning (EL)	92.1	92.1	91.5
Extreme Gradient Boost (XGB)	88.96	90.1	89.76
AdaBoost	87.7	88.2	87.9
Proposed	99.1	99.3	99.2

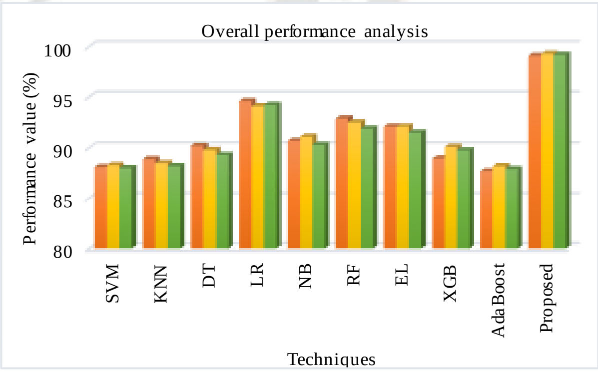

Figure 4. Overall performance analysis

Other metrics, such as sensitivity, specificity, and f1-score, are also validated and compared for the existing and proposed models, as shown in Table III and Fig. 4. These results also show that COLKR's performance is much better when compared to the other approaches. Because of the higher sensitivity, specificity, and f1-score values, the prediction rate has increased.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7147
***Article Received**: 08 April 2023 **Revised:** 29 May 2023 **Accepted:** 16 June 2023*

Typically, the Bit Error Rate (BER) is defined as the percentage of bits with errors dividing by the total amount of transmitted bits, as computed in the following model, as in (15):

$$\text{BER} = \frac{N_B}{\text{TC}_B} \times 100\% \tag{15}$$

where N_B indicates the number of error bits, and TC_B $TC_B$is the total amount of transmitted bits. This received bit is handled over a set length of time. This value represents the number of bits received with an error by the mobile station. Similarly, the Mean Squared Error (MSE) rate is defined as the total amount of error value in the statistical technique, which is computed as shown in (16).

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{n}\left(f_i - y_i\right)^2 \tag{16}$$

where N is the total number of data points, f_i f_i is the value that is given by the model, and y_i denotes the value that is actually observed at data point. The BER and MSE of existing and proposed machine learning algorithms are compared in Table IV and Fig. 5. When compared to the other strategies, the observed findings show that the suggested COLKR model efficiently minimizes the error rate. Because to the incorporation of COA, the suggested model's training and testing activities are effectively completed, resulting in a lower error rate.

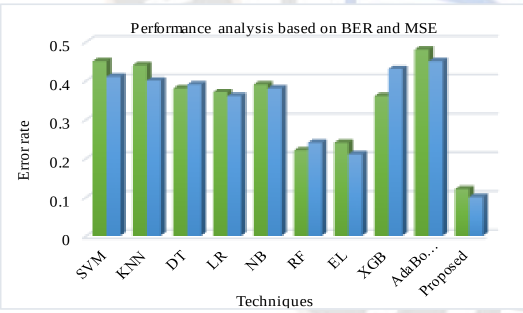

Fig. 5. Performance analysis based on BER and MSE.

Table IV: BER & MSE performance analysis

Techniques	BER	MSE
Support Vector Machine (SVM)	0.45	0.41
K-Nearest Neighbor (KNN)	0.44	0.40
Decision Tree (DT)	0.38	0.39
Logistic Regression (LR)	0.37	0.36
Naïve Bayes (NB)	0.39	0.38
Random Forest (RF)	0.22	0.24
Ensemble Learning (EL)	0.24	0.21
Extreme Gradient Boost (XGB)	0.36	0.43
AdaBoost	0.48	0.45
Proposed	0.12	0.1

The training and testing outcomes of the proposed COLKR classification model in terms of network failure and call connected rate are shown in Fig. 6 and Fig. 7. The training and testing results are validated in this analysis to demonstrate the classifier's performance. The training and testing performance results of COLKR in the proposed call drop prediction system have been greatly enhanced by employing the optimized feature set obtained with the assistance of COA.

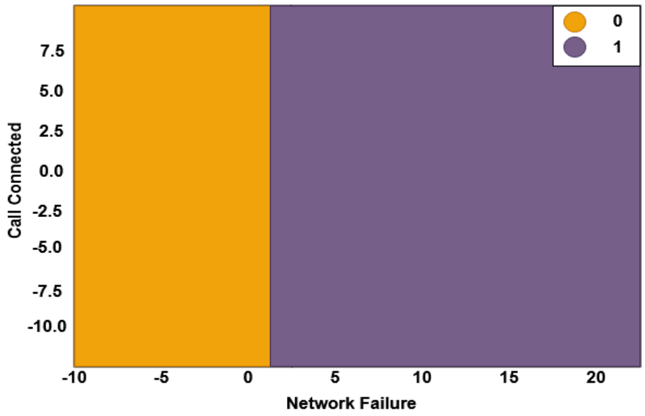

Figure. 6. Training results

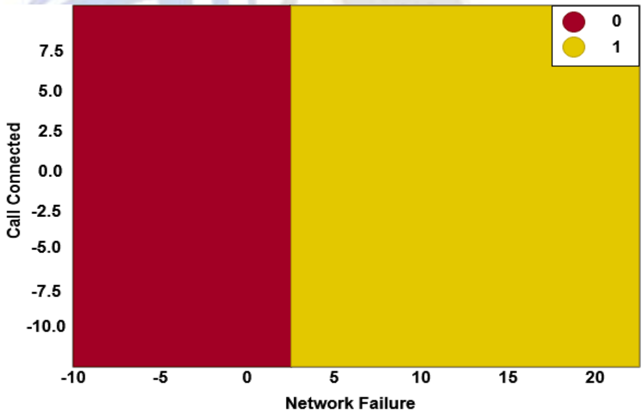

Fig. 7. Testing results.

V. CONCLUSION

Telecom companies generate vast amounts of data on a regular basis. Decision-makers emphasized that obtaining new consumers is more challenging than keeping existing ones. Despite the greatest efforts of authorities throughout the years, call dropouts remain an ongoing issue. In mobile networks, the call drop prediction is still exists one of the most important problem need to be resolved. For this purpose, this paper presents a new framework, named as, COLKR for an effective call drop prediction. The CDR data comprising the attributes of String ID, A Party Calling Number, B Party Called number, Calling Time Stamp, Called Time Stamp, Event Code, Charge, Network Status, Call Result, Service Provider, Service Provider Code and Call Status, has been taken as the input for processing. After dataset acquisition, the data preprocessing is performed with the use of median regression filtering approach for normalizing the attributes. Then, the new COA is implemented to choose the most relevant features for call drop prediction, which are used by the classifier for training & testing. Furthermore, the LKRM is employed to precisely predict the call drops with reduced error rate and increased accuracy. Moreover, the performance and results of the proposed COLKR technique has been validated and compared using a variety of measures such as error rate, accuracy, sensitivity, specificity, and etc. The obtained results indicate that the proposed COLKR provides an improved outcomes in terms of accuracy (99%), error (0.19), and other performance measures (99%).

In future, the current work can be enhanced by applying a light weight deep learning mechanism for call drop prediction and network failure detection. Also, we aim to implement an advanced big data analytics model for identifying the determinants of churns from the telecommunication industry. Moreover, the network performance can be effectively optimized for gaining an improved customer satisfaction.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Under the supervision of Prof. Dr. Vasanthi Kumari P., conducted the research, sharing valuable insights on Machine Learning algorithms, which helped improve the quality of this paper. We also thank Dr. A. Vadivel and Dr. Hema reviews in detailed and thoughtful feedback helped me this survey more depth and a broader scope. Three authors had approved the final version of the paper.

REFERENCES

[1] M. McClellan, C. Cervelló-Pastor, and S. Sallent, "Deep learning at the mobile edge: Opportunities for 5G networks," Applied Sciences, vol. 10, p. 4735, 2020.

[2] R. Su, D. Zhang, R. Venkatesan, Z. Gong, C. Li, F. Ding, et al., "Resource allocation for network slicing in 5G telecommunication networks: A survey of principles and models," IEEE Network, vol. 33, pp. 172-179, 2019.

[3] M.-F. Huang, M. Salemi, Y. Chen, J. Zhao, T. J. Xia, G. A. Wellbrock, et al., "First field trial of distributed fiber optical sensing and high-speed communication over an operational telecom network," Journal of Lightwave Technology, vol. 38, pp. 75-81, 2019.

[4] I. Tomkos, D. Klonidis, E. Pikasis, and S. Theodoridis, "Toward the 6G network era: Opportunities and challenges," IT Professional, vol. 22, pp. 34-38, 2020.

[5] K. Hirata, H. Yamamoto, S. Kamamura, T. Oka, Y. Uematsu, H. Maeda, et al., "System design for traveling maintenance in wide-area telecommunication networks," IEICE Transactions on Communications, vol. 103, pp. 363-374, 2020.

[6] A. Muradova and K. Khujamatov, "Results of calculations of parameters of reliability of restored devices of the multiservice communication network," in 2019 International Conference on Information Science and Communications Technologies (ICISCT), 2019, pp. 1-4.

[7] H. Abdulkareem, A. Tekanyi, A. Kassim, Z. Muhammad, U. Almustapha, and H. Adamu, "Analysis of a GSM network quality of service using call drop rate and call setup success rate as performance indicators," in Proceedings of: 2nd International Conference of the IEEE Nigeria, 2019, p. 300.

[8] M. Duraipandian, "Long term evolution-self organizing network for minimization of sudden call termination in mobile radio access networks," Journal of trends in Computer Science and Smart technology (TCSST), vol. 2, pp. 89-97, 2020.

[9] O. O. Erunkulu, E. N. Onwuka, O. Ugweje, and L. A. Ajao, "Prediction of call drops in GSM network using artificial neural network," Jurnal Teknologi dan Sistem Komputer, vol. 7, pp. 38-46, 2019.

[10] L. Xu, X. Zhao, Y. Yu, Y. Luan, L. Zhao, X. Cheng, et al., "A comprehensive operation and revenue analysis algorithm for LTE/5G wireless system based on telecom operator data," in 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 2019, pp. 1521-1524.

[11] A. Alhammadi, M. Roslee, M. Y. Alias, I. Shayea, S. Alraih, and K. S. Mohamed, "Auto tuning self-optimization algorithm for mobility management in LTE-A and 5G HetNets," IEEE Access, vol. 8, pp. 294-304, 2019.

[12] Z. Mammeri, "Reinforcement learning based routing in networks: Review and classification of approaches," Ieee Access, vol. 7, pp. 55916-55950, 2019.

[13] O. A. Wahab, A. Mourad, H. Otrok, and T. Taleb, "Federated machine learning: Survey, multi-level classification, desirable criteria and future directions in communication and networking systems," IEEE Communications Surveys & Tutorials, vol. 23, pp. 1342-1397, 2021.

[14] G. Manogaran, M. Alazab, V. Saravanan, B. S. Rawal, P. M. Shakeel, R. Sundarasekar, et al., "Machine learning assisted information management scheme in service concentrated IoT," IEEE transactions on industrial informatics, vol. 17, pp. 2871-2879, 2020.

[15] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, "A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector," IEEE access, vol. 7, pp. 60134-60149, 2019.

[16] F. Tang, B. Mao, Y. Kawamoto, and N. Kato, "Survey on machine learning for intelligent end-to-end communication toward 6G: From network access, routing to traffic control and streaming adaption," IEEE Communications Surveys & Tutorials, vol. 23, pp. 1578-1598, 2021.

[17] M. Ozturk, M. Gogate, O. Onireti, A. Adeel, A. Hussain, and M. A. Imran, "A novel deep learning driven, low-cost mobility prediction approach for 5G cellular networks: The case of the

Control/Data Separation Architecture (CDSA)," Neurocomputing, vol. 358, pp. 479-489, 2019/09/17/ 2019.
[18] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," IEEE Communications Surveys & Tutorials, vol. 21, pp. 3072-3108, 2019.
[19] M. S. H. Abad, E. Ozfatura, D. Gunduz, and O. Ercetin, "Hierarchical federated learning across heterogeneous cellular networks," in ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 8866-8870.
[20] A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," Journal of Big Data, vol. 6, pp. 1-24, 2019.
[21] G. Luo, Q. Yuan, J. Li, S. Wang, and F. Yang, "Artificial intelligence powered mobile networks: From cognition to decision," IEEE Network, vol. 36, pp. 136-144, 2022.
[22] J. Moysen and L. Giupponi, "From 4G to 5G: Self-organized network management meets machine learning," Computer Communications, vol. 129, pp. 248-268, 2018.
[23] A. Anand and G. de Veciana, "Resource allocation and HARQ optimization for URLLC traffic in 5G wireless networks," IEEE Journal on Selected Areas in Communications, vol. 36, pp. 2411-2421, 2018.
[24] A. Asghar, H. Farooq, and A. Imran, "Self-healing in emerging cellular networks: Review, challenges, and research directions," IEEE Communications Surveys & Tutorials, vol. 20, pp. 1682-1709, 2018.
[25] M. S. Hadi, A. Q. Lawey, T. E. El-Gorashi, and J. M. Elmirghani, "Big data analytics for wireless and wired network design: A survey," Computer Networks, vol. 132, pp. 180-199, 2018.
[26] I. V. Pustokhina, D. A. Pustokhin, P. T. Nguyen, M. Elhoseny, and K. Shankar, "Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector," Complex & Intelligent Systems, 2021/04/05 2021.
[27] G. RM, "Prediction of customer plan using churn analysis for telecom industry," Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), vol. 13, pp. 926-929, 2020.
[28] D. Mishra and A. Mishra, "Self-optimization in LTE: An approach to reduce call drops in mobile network," in Futuristic Trends in Network and Communication Technologies: First International Conference, FTNCT 2018, Solan, India, February 9–10, 2018, Revised Selected Papers 1, 2019, pp. 382-395.
[29] G. Ashok and P. V. Kumari, "To Analyse and optimise the issues in call drops and enhance the real time (CDR) processing," EasyChair 2516-2314, 2020.
[30] G. Ashok and V. Kumari, "An investigation of various machine learning techniques for mobile call data analysis for reducing call drop," Materials Today: Proceedings, vol. 51, pp. 2476-2478, 2022.
[31] G. Kaur, R. K. Goyal, and R. Mehta, "An efficient handover mechanism for 5G networks using hybridization of LSTM and SVM," Multimedia Tools and Applications, vol. 81, pp. 37057-37085, 2022.
[32] S. Verma, O. Singh, S. Kumar, and S. Mishra, "Fuzzy—KNN-Assisted Vehicular Localization for Bluetooth and Wi-Fi Scenario," in Emerging Technologies for Computing, Communication and Smart Cities: Proceedings of ETCCS 2021, ed: Springer, 2022, pp. 451-465.
[33] C. Gupta, I. Johri, K. Srinivasan, Y.-C. Hu, S. M. Qaisar, and K.-Y. Huang, "A systematic review on machine learning and deep learning models for electronic information security in mobile networks," Sensors, vol. 22, p. 2017, 2022.
[34] K. Mandal and G. Gong, "PrivFL: Practical privacy-preserving federated regressions on high-dimensional data over mobile networks," in Proceedings of the 2019 ACM SIGSAC Conference on Cloud Computing Security Workshop, 2019, pp. 57-68.
[35] D. Clemente, G. Soares, D. Fernandes, R. Cortesao, P. Sebastiao, and L. S. Ferreira, "Traffic forecast in mobile networks: Classification system using machine learning," in 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), 2019, pp. 1-5.
[36] E. Alimpertis, A. Markopoulou, C. Butts, and K. Psounis, "City-wide signal strength maps: Prediction with random forests," in The World Wide Web Conference, 2019, pp. 2536-2542.
[37] K. Abbas, T. A. Khan, M. Afaq, and W.-C. Song, "Ensemble learning-based network data analytics for network slice orchestration and management: An intent-based networking mechanism," in NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium, 2022, pp. 1-5.
[38] X. Liu and T. Wang, "Application of XGBOOST model on potential 5G mobile users forecast," in Signal and Information Processing, Networking and Computers: Proceedings of the 8th International Conference on Signal and Information Processing, Networking and Computers (ICSINC), 2022, pp. 1492-1500.
[39] F. Wang, D. Jiang, H. Wen, and H. Song, "Adaboost-based security level classification of mobile intelligent terminals," The Journal of Supercomputing, vol. 75, pp. 7460-7478, 2019.