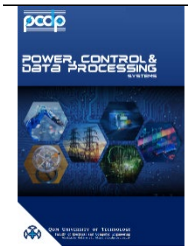
DOI: 10.30511/pcdp.2025.2072728.1047 - ISSN: 3060-8376 - Vol 3(1), 2026



POWER, CONTROL AND
DATA PROCESSING SYSTEMS
Available Online at: https://pcdp.qut.ac.ir/

ML-and VIKOR for Anomaly Detection and Cell Ranking in 5G/B5G

ARTICLE INFO

Article Type
Original Research

Authors
Reza Moammer Yami*,[1]
Soltanizadeh, Hadi[1]
Shahzadi, Ali[1]
Shirvani Moghadam, Shahriar[2]

[1] Department of Electrical and Computer Engineering, Semnan, Iran
r_moammer@yahoo.com
h_soltanizadeh@semnan.ac.ir
shahzadi@semnan.ac.ir
[2] Associate Professor of Communications/IEEE Senior Member,
sh_shirvani@sru.ac.ir

**** Correspondence***
Address: Telecommunications Infrastructure Company Tehran, Iran.
Phone: -
Fax: -
r_moammer@yahoo.com

Article History
Received: September 25, 2025
Accepted: November 28, 2025
ePublished: March 01, 2026

ABSTRACT

This study introduces a hybrid framework that integrates supervised machine learning (ML) algorithms with the VIKOR (VlseKriterijumska Optimizacija I Kompromisno Resenje) multi-criteria decision-making (MCDM) technique to advance anomaly detection and cell ranking in next-generation networks. The proposed model addresses critical challenges in heterogeneous network environments, including data imbalance and fault prioritization. Three ML algorithms—Naïve Bayes, Decision Tree, and Random Forest—were evaluated, with Random Forest achieving the highest accuracy (93.658%). However, the Decision Tree algorithm demonstrated optimal efficiency, balancing high accuracy (92.688%) with the fastest execution time (0.04 seconds), rendering it particularly suitable for real-time applications. The incorporation of VIKOR enhanced the framework by enabling fault prioritization based on severity and impact, improving detection of minority fault classes, and supporting multi-criteria resource management. This hybrid approach resulted in improved system accuracy, flexibility, and scalability, ultimately contributing to reduced operational response times and enhanced network reliability. The findings validate the efficacy of combining ML with MCDM for intelligent fault management and cell ranking in complex network ecosystems.

Keywords: Multi-criteria decision making, self-organizing networks, self-healing, faulty cell management, VIKOR, anomaly detection, 5G networks.

1 Introduction

The proliferation of big data analytics has revolutionized numerous industries, including mobile communications. For network operators, analyzing extensive network datasets has become imperative for efficient management, optimization, and configuration of next-generation systems. While historical data provides valuable performance insights, advanced modeling techniques can reveal critical patterns in self-configuration, network management, and self-healing operations. However, manual analysis of these substantial datasets proves both time-consuming and prone to errors, necessitating intelligent automated solutions

Recent advancements in artificial intelligence (AI) and machine learning (ML) have substantially transformed mobile network operations. These technologies have minimized errors in self-organizing networks (SONs), improved operational efficiency, and enhanced revenue streams while reducing human intervention. Standardized by 3GPP in Release 8 [1], SON has emerged as a fundamental component of 5G networks [2]. The transition toward 5G and beyond (B5G) has precipitated a dramatic increase in network-level metrics, driven by substantial infrastructure evolution. This growth has compelled researchers and operators to prioritize data-driven methodologies to achieve both technical and business objectives in modern SON environments.

Table 1 lists some network-level failures.

Table 1: Types of network failures

Failure Category	Specific Failure Types	Technical Specifications
Hardware Failures	Transmission Systems	Fiber cuts, Microwave link degradation >3dB, RF attenuation >15dB
	Base Station Components	BBU processing card failure, RRU power amplifier malfunction, AAS beamforming faults
	Power Infrastructure	Grid outages >5min, UPS battery failure, Rectifier module faults
	Thermal Management	HVAC failures, Temperature >85°C triggering protection
Software & Configuration	System Layer	Kernel panics, Memory leaks >80%, Driver conflicts
	Application Layer	Connection management state stuck, Resource deadlocks
	Configuration Errors	Incorrect Tx power settings, NRT configuration conflicts
Protocol Stack	Access Stratum	RACH preamble collisions >15%, Handover measurement delays >200ms
	Control Plane	RSRP < -120dBm within 100ms, Buffer congestion >90% for 5+ seconds
Environmental	Interference	Adjacent Channel Interference >-6dB, Thermal noise elevation
	Propagation	Building penetration loss >30dB (mmWave), Fast fading >20dB

To achieve the SON goals, intelligent rule-based techniques are needed to accurately detect failures and rank cells based on quality of service.

Mobile network technology has undergone remarkable evolution since the introduction of 1G in 1983, with the industry now advancing toward sixth-generation (6G) systems. Each generational transition has introduced new technical requirements and innovations, consistently pushing the industry forward. Figure 1 illustrates the anticipated technologies expected to drive progress in 5G, B5G, and future networks.

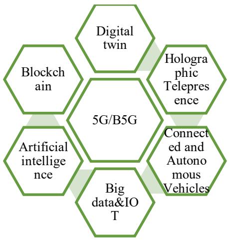


Figure 1: Expected technologies in generations 5G/B5G

Advancements in 6G Networks

Beyond current 5G/B5G specifications, 6G mobile networks are projected to introduce transformative capabilities across several domains:

- Terahertz (THz) band operation enabling ultra-high-speed communications
- Intelligent communication environments with adaptive signal transmission
- Pervasive AI integration for autonomous network operations
- Large-scale automation minimizing manual intervention
- Reconfigurable front-end systems for dynamic spectrum access
- Distributed environmental communications enhancing energy efficiency
- Space IoT utilizing CubeSats and UAVs for global coverage
- Cell-less massive MIMO ensuring seamless connectivity [4]

Challenges in Modern Mobile Networks

These technological advancements introduce substantial challenges in network management, configuration, and self-healing. Key issues include scalability constraints, management complexity, manual configuration processes, operational costs, and network inflexibility. Self-healing mechanisms in next-generation networks primarily focus on:

1. Anomaly detection and fault identification
2. Root cause analysis and diagnosis
3. Performance maintenance during component failures
4. Defective cell management and compensation [5]

Methodological Evolution and Limitations

Traditional network analysis has relied on rule-based systems, parametric algorithms, and statistical methods. However, these approaches demonstrate limited efficacy in heterogeneous network environments with complex structures. Consequently, research has shifted toward AI and ML techniques employing numerical approximation methods [6-7].

Machine learning algorithms offer significant advantages over classical statistical approaches, providing enhanced flexibility for network cell analysis. Within SON frameworks, three primary ML paradigms have gained prominence:

1. Supervised learning for fault prediction and classification
2. Unsupervised learning for pattern recognition in unlabeled data
3. Reinforcement learning for adaptive decision-making

These methodologies have enabled innovative applications categorized as:

- ML-based spectrum intelligence and radio resource management [9]
- ML-based transmission intelligence and signal processing [9]
- ML-based network intelligence and system optimization [9]

Research Objectives and Contributions

Despite these advancements, critical challenges persist in network scalability, management complexity, and configuration overhead. This research addresses these limitations by developing a hybrid framework integrating VIKOR multi-criteria decision-making with supervised machine learning algorithms. The specific objectives include:

1. Enhancing fault detection and diagnosis accuracy
2. Advancing self-healing capabilities through ML integration
3. Optimizing network performance using AI-driven decision frameworks

The paper is organized as follows: Section II examines self-healing activities in next-generation networks. Section III details the proposed methodology. Section IV presents experimental results and analysis. Section V concludes with implementation summary and research contributions.

2 Background Research

In the field of networked self-healing research, three primary approaches have been widely explored: rule-based, algorithm-based, and parametric-based methods. Rule-based approaches rely on empirical rules and are typically suited for small-scale networks. However, they often suffer from limited accuracy and increased complexity. Algorithm-based approaches, grounded in statistical theories, are also applicable to small networks.

While they offer reasonable accuracy, their flexibility is constrained when dealing with complex network structures. In contrast, parametric-based approaches leverage machine learning (ML), making them highly suitable for large-scale networks. These methods are characterized by their adaptability to complex structures, operational simplicity, and transparency. Table 2 summarizes the advantages and disadvantages of each approach.

Table 2. General methods for identifying defective cells

Role-based approaches Based on empirical rules	- Suitable for small network structures- - Lack of accuracy -Increase error - Very complex formulation
Algorithm-based approaches Based on statistical theories	- Suitable for small network structures - Low flexibility for complex structures - High operational complexity - Fairly accurate
Parametric -based approaches Based on machine learning	- Suitable for large network structures High -flexibility for complex structures - Operational simplicity and transparency - Based on data type and quality

The limitations of traditional statistical algorithms have driven the adoption of machine learning to enhance both accuracy and efficiency in self-healing mechanisms. Machine learning provides greater flexibility in analyzing network cells through three primary paradigms: unsupervised learning, supervised learning, and reinforcement learning. Due to the constraints of rule-based and parametric approaches, ML algorithms have shifted toward data-driven methods, enabling the automation of tasks that traditionally required significant human effort. These systems rely on algorithms and statistical models to analyze and infer patterns from data. However, the effectiveness of machine learning is heavily dependent on the quality of the input data. Without clean, consistent, and high-quality data, deriving meaningful quantitative insights becomes challenging.

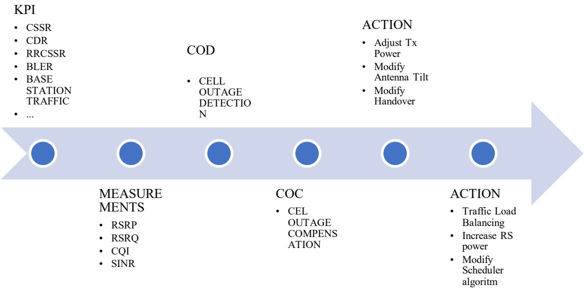


Figure 2. Structure of self-healing

Data analysis methods can be categorized into four main types, ordered by increasing complexity: descriptive, diagnostic, predictive, and prescriptive analysis. These methods can be effectively integrated with machine learning techniques to address the challenges of identifying and compensating for defective cells in networks. The application of each learning paradigm can be aligned with the framework illustrated in Figure 2, tailored to the specific conditions and objectives of the problem at hand.

In the context of self-healing, the discussion of compensation focuses on recovering, controlling, and optimizing the functionality of faulty cells within networks. Conversely, the detection of faulty cells centers on identifying these problematic cells within the network. Supervised learning is typically the primary approach for identifying defective cells. In next-generation networks, supervised learning is particularly valuable due to its predictive capabilities. Predicting the operational status of each cell at the network level is essential for reducing operational and capital expenditures (OPEX and CAPEX) while ensuring the delivery of desired services to subscribers. This predictive capability enables operators to make informed decisions in a timely manner, anticipate potential issues, and implement compensatory measures for affected cells.

One of the algorithms used in networks for fault detection in other target operations in the network is the decision tree [10]. In this method, the class of the sample is determined using an iterative process by moving through the tree levels and analyzing the existing features. One of the advantages of the decision tree algorithm is its resistance to noise and its low computational complexity in large volumes of data, and the main drawback of this method is the long processing time. Among the activities carried out with this algorithm, we can mention the research of Siva Kumar et al [11]. Using this algorithm, radio link failures in networks have been investigated. Another activity by Hangli and Zhuqiao [12] used the decision tree algorithm to select the appropriate network for the service according to the existing features. In the field of faulty cell detection, we can also mention the research conducted in [13], which was used to classify and predict faulty cells at the network level. Logistic regression [14]. is another algorithm that is often used for two-class learning problems. This method uses a linear model based on gradient descent to optimize and adjust the model parameters. For the logistic regression algorithm, the class probability is expressed based on the sigmoid function. Logistic regression works well for unbalanced and two-class data. Research conducted in [15-16] is the application of logistic regression in Internet networks. For example, Kulkarni et al. [16] used the logistic regression algorithm to identify failures and errors in the performance of 5G Internet network antennas. by simulating several network indicators using this algorithm, they estimated and identified failures and errors in the performance of network antennas. The random forest method [17] is another algorithm that is used in learning processes in networks. This algorithm is implemented by combining several main classifiers (decision trees) to create a more robust and accurate model. The main goal of this method is to solve the problem of data variance. Research that has been conducted using random forest in diagnostic processes in Internet networks can be mentioned in [18-19], in each of which the random forest method has been proposed as the central algorithm in the research. In the detection of defective cells, the implementation research [18] has been carried out using the random forest method to identify defective cells at the network level. Other learning methods include algorithms such as nearest neighbor and support vector machine, which have been used to identify and diagnose anomalies in network components such as [20-22]. For example, in the research [20] and [21], researchers used a combination of SVM and KNN algorithms to identify anomalies in the structure of cells. In each of these studies, the main goal has been to identify outliers in the network structure, and another method used in these studies is to change the index of neighboring cells, which to some extent indicates defects and defective cells. In neighboring cells, in other activities such as [23-24], the main goal of the research conducted is to identify structural patterns in network cells. As mentioned, in the process of identifying defective cells in networks, due to the nature of the data, supervisory methods are often used to determine the class of network cells. The reason for choosing such an approach to solve the problem is the nature of the data used. Using a reinforcement learning model together with deep learning-based approaches is often better than classical learning methods when the data has a large volume of high-dimensional samples. They are most useful in the discussion of defective cell compensation because in these algorithms the feature extraction and selection process is performed automatically. In this situation, the probability of error in selecting inappropriate features as well as computational processing is reduced. Among the researches that can be mentioned is the combination of reinforcement learning and deep learning to control the optimization of the network structure in the management of defective cells for the compensation process, the activities carried out in [26-27] can be mentioned. In general, the problem-solving approach in identifying and compensating defective cells in the network structure can be determined by focusing on the use of data types.

Table 3. Previous classification of research conducted

ML technique	Objective	Algorithm
SL/UL	Anomaly detection	k-NN, LOF
SL	Diagnosis	FFNN
SL	Diagnosis	Naive Bayesian
UL	Pattern identification	Hidden Markov Model
RL	Anomaly detection	SVM and LOF
UL	Pattern identification	LOF
SL/UL	Anomaly detection	DBSCAN algorithm
UL	Pattern identification	DL, Autoencoder
SL	Pattern identification	LSTM
SL	Anomaly prediction	Prophet Algorithm
SL	Pattern identification	Decision tree
SL	Pattern identification	DT, SVM, Random Forest
SL	Pattern identification	SVM, NN
SL	Pattern identification	Hierarchical clustering

Despite significant advances in mobile network management and self-healing, several critical challenges remain unresolved. These challenges hinder the efficient operation, scalability, and reliability of modern networks, especially as they evolve towards 5G, B5G, and 6G. Key challenges include:

1. Scalability

As networks expand in size and complexity, traditional management and self-healing approaches struggle to scale effectively. This leads to inefficiencies, especially in large-scale and heterogeneous network environments.

2. Network management complexity

The increasing heterogeneity and dynamic nature of network infrastructures make it difficult to maintain optimal performance using conventional rule-based and parametric approaches. These approaches often do not adapt to rapidly changing network conditions.

3. Manual Configuration
Many existing networks still rely on manual configuration processes that are time-consuming, error-prone, and unstable for next-generation networks. Automating these processes is essential to reduce human intervention and improve efficiency.

4. High Costs
The financial burden associated with deploying, maintaining, and upgrading complex network infrastructures remains a significant barrier for network operators. Cost-effective solutions are needed to ensure the sustainability of next-generation networks.

5. Inflexibility
Traditional methods often lack the adaptability needed to respond to dynamic network conditions and user demands. This inflexibility can lead to suboptimal performance and reduced quality of service (QoS).

6. Data Imbalance
A common issue in fault diagnosis and self-healing is imbalance in the data set, where the number of defective cells is significantly lower than that of normal cells. This imbalance can lead to biased models that prioritize the majority class, resulting in poor detection rates for faulty cells. Developing robust algorithms capable of handling unbalanced data is crucial for accurate fault detection and diagnosis.

7. Lack of fault priority ranking
Current self-healing mechanisms often treat all faults equally, regardless of their severity or impact on network performance. This lack of prioritization can lead to inefficient resource allocation and delayed response to critical issues. Implementing fault prioritization strategies is essential to optimize self-healing processes and ensure timely resolution of high-impact faults.
The increasing complexity and heterogeneity of network structures exacerbate these challenges and make it difficult for traditional methods to effectively deal with them. This highlights the urgent need for innovative solutions, especially those that leverage artificial intelligence (AI) and machine learning (ML), to enhance network performance, reliability, and efficiency.
By integrating advanced ML techniques with multi-criteria decision-making frameworks, this research aims to fill these gaps and provide scalable, efficient, and more reliable self-healing mechanisms for next-generation networks. Specifically, the proposed approach focuses on:

- Developing robust algorithms to handle unbalanced data and improve fault detection accuracy.
- Implementing fault prioritization strategies to optimize resource allocation and response time.
- Automating network configuration and management processes to reduce human intervention and operational costs.

3 The proposed algorithm

This study aims to identify defective cells. To achieve this goal, we use multi-criteria decision making (MCDM) methods, which are divided into two types: multi-objective decision making (MODM) and multi-attribute decision making (MADM). The goal of these decision-making approaches is to select the best option or evaluate different decision factors. In this study, the relevant factors include network-level derived indicators that are crucial for determining the functional status of each cell. Each assignment method has its own unique characteristics. Some techniques focus on evaluating the quality of options, while others prioritize alternatives or decide on their implementation. In this study, the VIKOR multi-criteria technique was used to label the dataset. This technique works by assigning a weight to each network-level derived indicator to evaluate the quality of a cell. The score of each cell ranges from 0 to 1. In the VIKOR method, the best cell receives a high score.

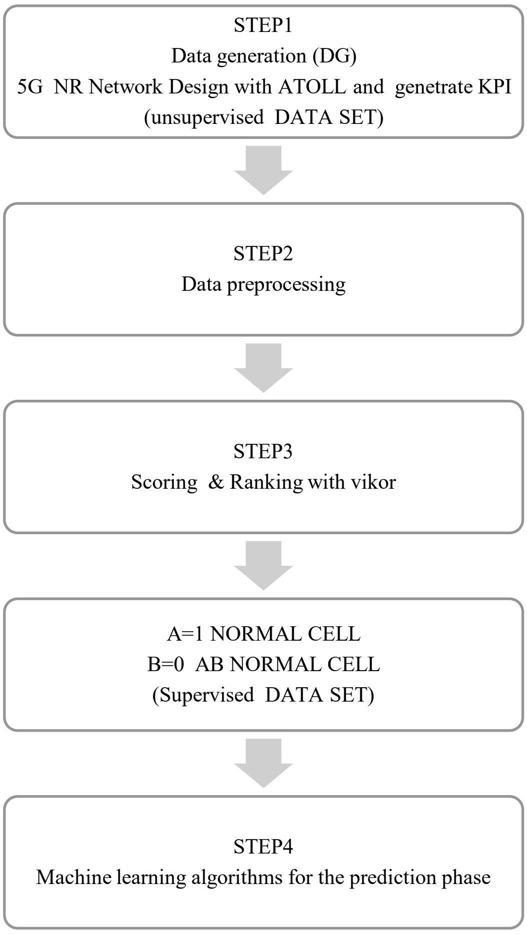


Figure3: Proposed process of ranking, recognition and classification

A. Data set

The data used in the research were 67120 samples. According to the existing protocols in the network, there are 3399 defective cells according to the received indicators used as input features and also the approach adopted using the VIKOR method algorithm. It is clear that the imbalance in the data class for damaged cells is very high, and this problem was expressed as one of the 3 main challenges in solving the problem. To train the model, we considered 80% of the data as training and 20% as testing. The number of features used was equal to 12, which is shown in Table 4. Also, the coordinates of the data are stated in Table 4.

Table 4. Features used in the data

Features(Kpi)s	Description
A:ERAB success Rate	This indicator indicates the percentage of success in establishing or maintaining an ERAB connection between the User Equipment (UE) and the Core Network. In other words, ERAB Success Rate tells us what percentage of user requests to establish a Data Bearer connection were successfully established or continued without problems.
B:S1Signal_ERAB_Setup_SR	S1Signal refers to the signaling between eNodeB (Evolved Node B) and MME (Mobility Management Entity) in an LTE network. S1 signaling is responsible for initiating and managing data communications in the network.
C:RAN_avail_Rate	RAN (Radio Access Network): The radio access network is responsible for wireless communication between user devices (UEs) and the core network. It consists of various equipment such as base stations (eNodeB in LTE or gNodeB in 5G).
D:HandOver	Handover (or Handover) in mobile networks is the process of transferring a user's communications from one cell or base station to another. This process occurs when the user moves or changes location (such as moving inside a car or when moving between cell towers) in order to maintain uninterrupted connectivity.
E:InterF_HOOut_SR	Inter-Frequency Handover (InterF-HO): This refers to the transfer of a user connection from one radio frequency to another in the network. This type of handover occurs when a user needs to move from one frequency to another, for example due to network traffic or better signal conditions.
F:IntraF_HOOut_SR	Intra-Frequency Handover (IntraF-HO): It means transferring a user connection from one cell to another cell on the same radio frequency. This type of handover usually occurs when the user is moving and needs to change cells on the same frequency to maintain the quality of the connection.
G:Call_Drop_Rate	Call Drop Rate, in simple terms, indicates what percentage of users' voice calls are unintentionally dropped due to network problems. This indicator is an important metric for evaluating network quality and performance.
H:CSFB_Rate	Circuit-Switched Fallback Rate (CSFB Rate) is a key performance indicator (KPI) in mobile networks that refers to the success or quality of the process of transferring voice calls from an LTE (4G) network to older networks (such as 3G or 2G). This process is used due to the limitations of the LTE network in providing traditional voice calls.
I:Call Setup Success Rate (CSSR)	Call Setup Success Rate (CSSR) is a key performance indicator (KPI) in telecommunications networks that indicates the percentage of calls that are successfully established. This indicator evaluates the quality and reliability of the network in the process of initiating voice calls.
G:Average_CQI	CQI (Channel Quality Indicator): A measure that the user device uses to report the quality of the radio channel to the base station. This value ranges from 0 to 15: A value of 0 indicates very poor channel quality. A value of 15 indicates the best channel quality.
I:Radio Resource Control (RRC)	Radio Resource Control (RRC) is a control layer protocol in mobile networks (such as LTE and 5G) that manages radio communications between the user equipment (UE) and the network (eNodeB or gNodeB). It operates in the Layer 3 Control Plane (L3 Control Plane) and is critical for managing and maintaining user communications.
J:Average_UL_Packet_Loss	Average UL Packet Loss (UL) is an important metric in telecommunications networks that measures the amount of data loss from a user equipment (UE) to the network over a specified period of time. It is commonly used to evaluate the quality of data transmission from the user equipment to the network for data services (such as mobile internet).

B: Data preparation

Data preparation is a crucial step in most learning operations, serving various objectives across different learning domains. It is often viewed as a preprocessing phase that takes place before the main training process begins. Data preprocessing consists of tasks such as feature equalization and mapping values to specific time intervals. Additionally, a common practice for collected data is to label the samples and assign them to specific classes, which is essential for supervised learning models.

C: data standardization

The primary goal of preprocessing is to eliminate noise from the data structure. For numerical data, data cleaning typically involves processes such as normalization and standardization. These operations help to remove the misleading effects of features that have values significantly larger or smaller than those of other features, ensuring a more accurate analysis.

D: Data labeling

To effectively identify defective cells, the datasets used must have binary labels. Decision making involves selecting the best option or assigning weights to different decision factors. Each method has a specific role. Some measure criteria, others rank the options, and some evaluate the criteria. For labeling the dataset, a multi-criteria technique such as VIKOR is used. The VIKOR method is used to assign weights to each criterion to evaluate the quality of the cell, while the options are ranked based on their similarity to the ideal. The cell labeling process consists of three general steps: 1) cell scoring, 2) cell ranking, and 3) determining a threshold value to separate healthy cells from defective cells and finally labeling the desired data sample. In this case, the threshold value was determined to be 0.3, which is obtained by considering the weight of each indicator and the positive or negative coefficient of that indicator on the network performance.

3.1 VIKOR Calculation Methodology

3.1.1 Decision Matrix Formation

Construct initial matrix with dimensions 67,120 × 12:

- **Rows**: Each cell (67,120 cells)
- **Columns**: Performance indicators (12 indicators)
- **Columns**: Performance indicators (12 indicators)

(1) $D = \left[X_{ij}\right]\ where\ i = \ 1\ to\ 67120\ ,\ j = 1\ to\ 12$

3.1.2 Decision Matrix Normalization

Using linear normalization due to uniform measurement scale:

(2) $R_{ij} = \frac{X_{ij}}{(max\ X_j)}\ for\ positive\ indicator$

(3) $R_{ij} = \frac{min\ (X_j)}{X_{ij}}\ for\ negative\ indictor$

3.1.3 Weight Assignment

Equal weights for all indicators:

(4) $Wj\ = \frac{1}{12} \approx\ 0.0833\ \ for\,j\ =\ 1\ to\ 12$

3.1.4 Ideal and Anti-Ideal Values

Ideal values:

(5) $for\ positive\ indicators\ f_j{*} =\ max(R_{ij})$

(6) $for\ negative\ indicators\ (J,G) f_j{*} = min(R_{ij})$

Anti-ideal values:

(7) $for\ positive\ indicators\ f_j^{*} =\ min(R_{ij})$

(8) $for\ negative\ indicators\ (J,G)\ f_j^{*}\ =\ max(R_{ij})$

3.1.5 Calculation of S_i and R_i

(9) $S_i = \sum\left[Wj * \frac{f_{j}{*}-Rij}{f_{j}{*}-f_{j^-}}\right], for\ j = 1\ to\ 12$

(10) $R_i = max\left[(Wj * (f_j{*} - f_j{-})\right],\ for\ j\ 1\ to\ 12$

3.1.6 VIKOR Index (Q_i) Calculation

(11) $Q_i = V^{*}\left[\frac{(S_i-s^{*})}{s^{-}-S^{+}}\right] + (1-V)^{*}\left[\frac{(R_i-R^{*})}{R^{-}-R^{+}}\right]$

(12) $S^{*} =\ min(\,S_i),S^{-} = max(\,S_i)$
$R^{*} = min(\,Ri),R^{-} = max(\,Ri)$

4 Results

In this section, the results obtained from the proposed algorithm are presented. First, each of the supervised algorithms used in the research are introduced. Then, in the next section, the results obtained from combining each algorithm with VIKOR are shown and then; after reviewing the results, the output obtained from the implemented algorithm is compared with the baseline methods.

Naive Bayes Classification & VIKOR Result

Correctly Classified Instances	91.143 %
Incorrectly Classified Instances	4.857 %
Kappa statistic	0.1581
Mean absolute error	0.0485
Root mean squared error	0.2193
Relative absolute error	51.4057 %
Root relative squared error	100.985 %
Time taken to build model	0.18 seconds

Table 5. Naive Bayes Classification Result

Precision	Recall	F-Measure	Cell class
0.913	0.964	0.938	A=0
0.866	0.887	0.872	B=1

Decision Tree Classification& VIKOR Result

Correctly Classified Instances	92.6878 %
Incorrectly Classified Instances	11.312 %
Kappa statistic	0.3023
Mean absolute error	0.1696
Root mean squared error	0.3252
Relative absolute error	83.239 %
Root relative squared error	102.589 %
Time taken to build model:	0.04 seconds

Table 6. Decision Tree Classification Result

Precision	Recall	F-Measure	Cell class
0.904	0.914	0.900	A=0
0.923	0.985	0.953	B=1

Random Forest Classification & VIKOR Result

Correctly Classified Instances	93.658 %
Incorrectly Classified Instances	11.312 %
Kappa statistic	0
Mean absolute error	0.1683
Root mean squared error	0.2179
Relative absolute error	82.6153 %
Root relative squared error	88.022 %
Time taken to build model:	0.08 seconds

Table 7. Random Forest Classification Result

Precision	Recall	F-Measure	Cell class
0.936	0.949	0.944	A=0
0.897	0.900	0.899	B=1

Table 8. Comprehensive evaluation of algorithms based on comparison of overall performance indicators

Algorithm	Accuracy	Kappa	MAE	RMSE	Runtime(s)
Naive Bayes	91.143%	0.1581	0.0485	0.2193	0.18
Decision Tree	92.6878%	0.3023	0.1696	0.3252	0.04
Random Forest	93.658%	0	0.1683	0.2179	0.08

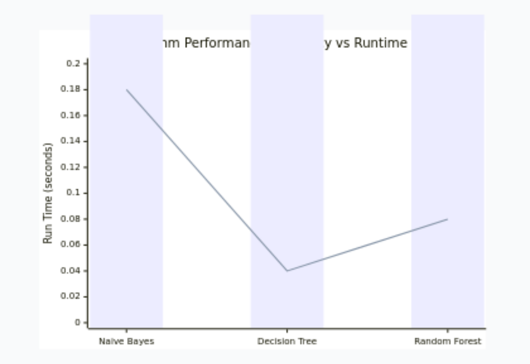


Figure 4. Results based on response time

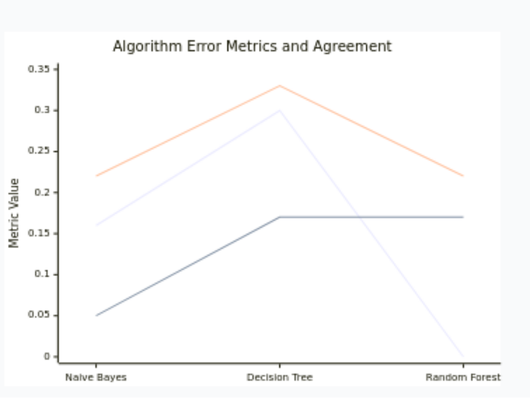


Figure 5. Evaluation based on error criteria

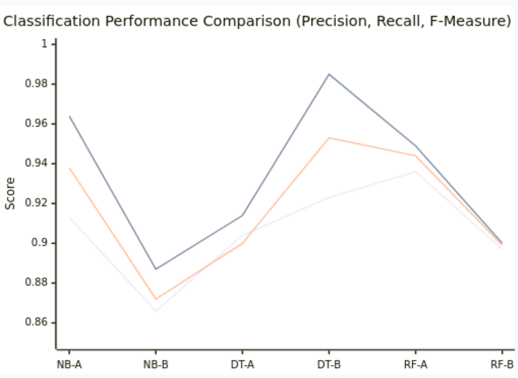


Figure:6: Evaluation based on performance criteria

5 Conclusion

This research proposed a hybrid approach that integrates supervised machine learning algorithms with the VIKOR multi-criteria decision-making technique to address key challenges in self-healing and fault management in next-generation mobile networks. The results demonstrate significant improvements in fault detection, diagnosis, and compensation, particularly in handling the complexities and heterogeneities of modern network infrastructures. Below are a comprehensive evaluation of the findings and the advantages of using VIKOR to overcome the challenges identified in this study:

1. Comprehensive Evaluation of Results

The proposed hybrid approach was evaluated using three supervised learning algorithms: Naive Bayes, Decision Tree, and Random Forest. The results highlighted the strengths and weaknesses of each algorithm:

Naive Bayes: Achieved 91.143% accuracy with low prediction errors (MAE: 0.0485, RMSE: 0.2193), making it suitable for real-time applications. However, its low Kappa score (0.1581) indicated limitations in detecting minority classes.

Decision Tree: Achieved 92.6878% accuracy with a better Kappa score (0.3023) and the fastest run time (0.04 seconds), making it ideal for real-time fault detection.

Random Forest: Achieved the highest accuracy (93.658%) and the lowest RMSE (0.279), but its Kappa score of 0 revealed challenges in handling imbalanced data.

2. Advantages of Using VIKOR

The integration of VIKOR with machine learning algorithms provided several key benefits:

Handling Data Imbalance: VIKOR's multi-criteria ranking mechanism ensured that minority classes (faulty cells) were given appropriate attention, improving the accuracy of fault detection in imbalanced datasets.

Fault Prioritization: By ranking faults based on their severity and impact, VIKOR enabled operators to prioritize critical issues, optimizing resource allocation and reducing downtime.

Multi-Criteria Decision-Making: VIKOR's ability to balance conflicting criteria (e.g., fault severity, performance metrics, and operational costs) led to more comprehensive and realistic evaluations of network health.

Improved Scalability and Flexibility: The combination of VIKOR with machine learning algorithms enhanced the adaptability of the solution to large-scale and heterogeneous networks.

3. Overcoming Key Challenges

The proposed approach effectively addressed several challenges in self-healing and fault management:

Data Imbalance: By prioritizing minority classes, VIKOR mitigated the bias toward majority classes, improving fault detection accuracy.

Lack of Fault Prioritization: VIKOR's ranking mechanism ensured that high-impact faults were addressed promptly, enhancing network reliability.

Complexity and Heterogeneity: The hybrid approach adapted to the dynamic nature of modern networks, outperforming traditional rule-based and parametric methods.

The integration of VIKOR with supervised learning algorithms represents a significant advancement in self-healing and fault management for next-generation networks. By addressing key challenges such as data imbalance, fault prioritization, and multi-criteria decision-making, this hybrid approach offers a scalable, flexible, and efficient solution for modern network infrastructures. Future work could explore the application of this framework in other domains, such as IoT and edge computing, to further validate its effectiveness and adaptability. In general, it can be said:

Phase 1 - Data Quality Enhancement: The VIKOR method transforms unsupervised network data into structured, labeled datasets by applying multi-criteria decision rules based on key performance indicators. This resolves fundamental scalability challenges in heterogeneous networks.

Phase 2 - Intelligent Classification: Supervised machine learning algorithms then leverage this enhanced dataset to achieve precise fault detection and classification, effectively managing network complexity through automated learning mechanisms.

The integration of rule-based VIOR prioritization with ML adaptive learning creates a synergistic framework that overcomes traditional limitations in handling large-scale, complex network environments. This combined approach ensures both scalability and accuracy in fault management.

Disclosure of Potential Conflicts of Interest

The Authors declare that there is no conflict of interest

Reference

[1] [1]Moysen, J., & Giupponi, L. (2018). From 4G to 5G: Self-organized network management meets machine learning. *Computer Communications*, *129*, 248-268. https://doi.org/10.1016/j.comcom.2018.07.015

[2] J. Rodriguez, Fundamentals of 5G Mobile Networks. John Wiley & Sons, 2015.

[3] Kim, H. (2022). Artificial intelligence for 6G (pp. 978-3030950408). Springer. https://doi.org/10.1002/9781118867464

[4] Akyildiz, I. F., Kak, A., & Nie, S. (2020). 6G and beyond: The future of wireless communications systems. *IEEE access*, *8*, 133995-134030. https://doi.org/10.1109/access.2020.3010896

[5] Klaine, P. V., Imran, M. A., Onireti, O., & Souza, R. D. (2017). A survey of machine learning techniques applied to self-organizing cellular networks. *IEEE Communications Surveys & Tutorials*, *19*(4), 2392-2431. https://doi.org/10.1109/comst.2017.2727878

[6] Sridharan, S. (2020). Machine learning (ML) in a 5G standalone (SA) self organizing network (SON). *arXiv preprint arXiv:2011.12288*. https://doi.org/10.14445/22312803/ijctt-v68i11p105

[7] Fourati, H., Maaloul, R., Chaari, L., & Jmaiel, M. (2021). Comprehensive survey on self-organizing cellular network approaches applied to 5G networks. *Computer Networks*, *199*, 108435.

[8] Papidas, A. G., & Polyzos, G. C. (2022). Self-organizing networks for 5g and beyond: A view from the top. *Future Internet*, *14*(3), 95. https://doi.org/10.1016/j.comnet.2021.108435

[9] Luo, F. L. (Ed.). (2020). Machine learning for future wireless communications. https://doi.org/10.1002/9781119562306

[10] Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, *2*(01), 20-28. https://doi.org/10.38094/jastt20165

[11] Li, W. H., & Qi, Z. (2018). Network selection algorithm based on decision tree in heterogeneous wireless networks. In *MATEC Web of Conferences* (Vol. 189, p. 04010). EDP Sciences. https://doi.org/10.1051/matecconf/201818904010

[12] Nethraa Sivakumar, Pooja Srinivasan, Nikhil Viswanath, and Venkateswaran N.(2022) "Decision tree-based radio link failure prediction for 5G communication reliability," ITU Journal on Future and Evolving Technologies, vol. 3, no. 2, pp. 142–156, Jul. 2022, doi: 10.52953/lzlj8762. https://doi.org/10.52953/lzlj8762

[13] Saeed, U., Jan, S. U., Lee, Y. D., & Koo, I. (2021). Fault diagnosis based on extremely randomized trees in wireless sensor networks. *Reliability engineering & system safety*, *205*, 107284. https://doi.org/10.1016/j.ress.2020.107284

[14] Fernandes, A. A. T., Figueiredo Filho, D. B., Rocha, E. C. D., & Nascimento, W. D. S. (2021). Read this paper if you want to learn logistic regression. *Revista de Sociologia e Política*, *28*, 006. https://doi.org/10.1590/1678-987320287406en

[15] Kulkarni, A. M., Saini, G., & Pattnaik, S. S. (2023, March). Antenna array fault detection using logistic regression technique. In *International Conference on Artificial Intelligence of Things* (pp. 13-29). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-48781-1_2

[16] Sharma, V., Hong, T., Cecchi, V., Hofmann, A., & Lee, J. Y. (2023). Forecasting weather-related power outages using weighted logistic regression. *IET Smart Grid*, *6*(5), 470-479. https://doi.org/10.1049/stg2.12109

[17] Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest. In Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3 (pp. 246-252). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34038-3

[18] Abd El-Aziz, L., Amr, E., Yehia, H., Mostfa, H., Hisham, M., Shenawy, A., ... & El-Akel, H. (2020, October). Cell outage detection and degradation classification based on alarms and kpi's correlation. In *2020 2nd Novel intelligent and leading emerging sciences conference (NILES)* (pp. 230-235). IEEE. https://doi.org/10.1109/niles50944.2020.9257920

[19] Samidi, F. S., Mohamed Radzi, N. A., Mohd Azmi, K. H., Mohd Aripin, N., & Azhar, N. A. (2022). 5G technology: ML hyperparameter tuning analysis for subcarrier spacing prediction model. *Applied Sciences*, *12*(16), 8271. https://doi.org/10.3390/app12168271

[20] Ali-Tolppa, J., Kajo, M., Gajic, B., Malanchini, I., Schultz, B., & Liao, Q. (2020). Cognitive Autonomy for Network Self-Healing. *Towards Cognitive Autonomous Networks: Network Management Automation for 5G and Beyond*, 345-384. https://doi.org/10.1002/9781119586449.ch9

[21] Moysen, J. (2016). A cell outage management framework for dense heterogeneous networks. *IEEE transactions on vehicular technology, 2016, vol. 65, núm. 4, p. 2097-2113*. https://doi.org/10.1109/tvt.2015.2431371

[22] Zoha, A., Saeed, A., Imran, A., Imran, M. A., & Abu-Dayya, A. (2016). A learning-based approach for autonomous outage detection and coverage optimization. *Transactions on Emerging Telecommunications Technologies*, *27*(3), 439-450. https://doi.org/10.1002/ett.2971

[23] Asghar, M. Z., Ahmed, F., & Hämäläinen, J. (2021, December). Artificial intelligence enabled self-healing for mobile network automation. In *2021 IEEE Globecom Workshops (GC Wkshps)* (pp. 1-6). IEEE. https://doi.org/10.1109/gcwkshps52748.2021.9681937

[24] Yu, P., Wang, H., & Chen, Z. (2022, June). Active Cell Outage Detection Algorithm for Broadband Services in 5G Cloud Radio Access Networks. In *2022 IEEE International Symposium on*

Broadband Multimedia Systems and Broadcasting (BMSB) (pp. 1-5). IEEE. https://doi.org/10.1109/bmsb55706.2022.9828778

[25] Yu, P., Yang, X., Zhou, F., Li, H., Feng, L., Li, W., & Qiu, X. (2020). Deep reinforcement learning aided cell outage compensation framework in 5G cloud radio access networks. *Mobile Networks and Applications*, *25*(5), 1644-1654. https://doi.org/10.1007/s11036-020-01574-8

[26] Iwamoto, M., Suzuki, A., & Kobayashi, M. (2023, May). Deep Reinforcement Learning Based Antenna Selection for Cell Outage Compensation. In *ICC 2023-IEEE International Conference on Communications* (pp. 3945-3950). IEEE. https://doi.org/10.1109/icc45041.2023.10279017