Vol 18, Fall 1404

Information Technology in Engineering Design

https://sanad.iau.ir/journal/ited

Ranking of cells and estimation of defective cells with TOPSIS technique and Machine learning in 5G&B5G

Reza Moammer Yami(1) Ali Akbar Khazaei(2)* Saeed Rahati Quehani(3)

(1) Department of Electrical Engineering, Ma.C. , Islamic Azad University, Mashhad, Iran
(2) Department of Electrical Engineering, Ma.C., Islamic Azad University, Mashhad, Iran*
(3) Department of Biomedical Engineering, CT.C., Islamic Azad University, Tehran, Iran

(Date received: 1403/08/29 Date accepted: 1403/11/27)

Abstract

The intelligence of 5G&B5G networks in the field of management and configuration, optimization and self-healing to reduce operational costs and capital costs without human intervention is one of the requirements of 5G&B5G. The self-healing process, as one of the main sub-sets of these networks, identifies faulty and out-of-service cells and compensates for these cells. In this research, a hybrid learning algorithm called stack classification, which is one of the types of hybrid methods in group learning models, has been used to estimate the faulty cells of the network. For this purpose, first, data set labeling and grid cell ranking are performed in the data preparation section using the TOPSIS algorithm, which is a rule-based approach. Then, using the stack classification algorithm, the faulty cells are classified and estimated at the data level. The results of applying the proposed algorithm showed its high efficiency in detecting faulty cells. Therefore, precision, recall, accuracy of the training and testing parts of the learning model showed a high rate of correct detection in determining the target class.

Keywords: Multi-criteria-decision-making, self organizing networks, self-healing of defective cells, management of defective cells, TOPSIS

* Corresponding author:

Ali Akbar Khazaei
Address: Department of Electrical Engineering, Ma.C., Islamic Azad University, Mashhad, Iran

Email: khazaei@mshdiau.ac.ir

1- Introduction

New technologies of 5G and B5G telecommunication networks are one of the most important research areas today [1-3]. Many researchers and companies are active in this field. Considering the progress and increasing penetration of artificial intelligence and the two main arms of this field, namely machine learning and data mining, in various areas of the new generation of mobile communications, the change and development of the infrastructure of these networks to provide more features and services, management and optimization and intelligent configuration in this generation of networks is required and is based on artificial intelligence [4-6]. Mobile networks have made significant progress since 1983, the first generation (1G) to the current generation (6G). Table 1 and Figure 1 show the evolution of the capabilities and services available in mobile wireless networks for each of the existing generations. The fifth generation has provided users with technologies such as augmented and artificial intelligence and complex high-speed systems and heterogeneous network structures, virtual reality [7]. For the sixth generation, blockchain was added as a main technology for use in networks. In addition to the blockchain capability in the sixth generation and above [8], other services are also expected as other capabilities as shown in Figure 2 [7]. With the increase of such capabilities for use in networks, many challenges have also been raised in the management and configuration of optimization and self-healing in the infrastructure of these networks without human intervention. Artificial intelligence algorithms, especially highly flexible machine learning methods, have been studied and used as the first and most important idea in solving the challenges in these three areas according to the intended goals of these networks [9].

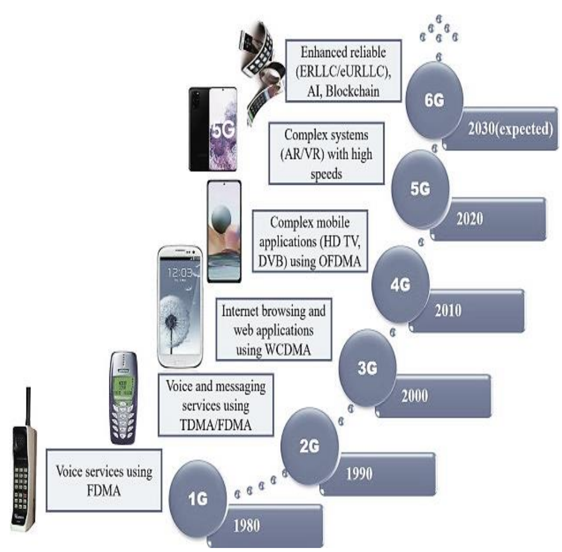

Figure 1: Evolution of infrastructure capabilities in mobile networks from 1G to 6G [9]

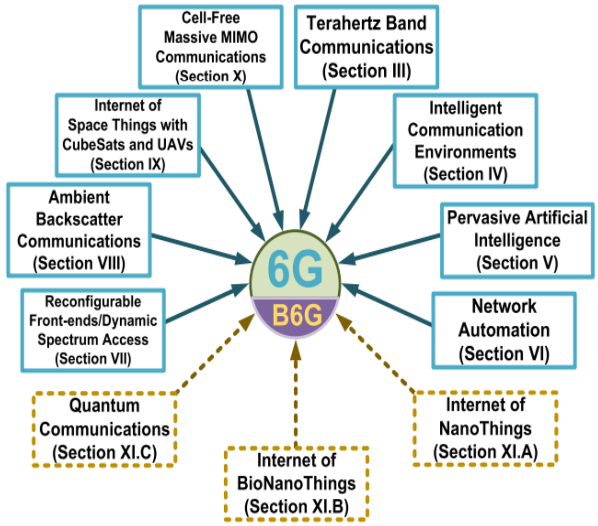

Figure 2: Capabilities of the sixth generation of mobile communications [7]

Table 1: Evolution of mobile communications [10-11]

Generation	Year	Main capabilities of mobile network generations
<1G	< 1983	voice service analogue single side band analogue systems
1G	1983_	voice service analogue low capacity mobile systems FDMA
2G	1990_	sms service high-capacity voice with limited data service TDMA,GSM,EDGE,GRPS
2.5G	1995_	CDMA system using 1.25 MHz bandwidth
3G	1999_	web service application UMTS, TD-SCDMA CDMA2000, WCDMA
4G	2013_	mobile internet application WiMAX , LTE-A
5G	2021_	internet of things Augmented Reality , Virtual Reality BDMA, SDN, Mm-Wave Embb , URLLC, mMTC
6G	2023_	Blockchain technology, THz and optical wireless communications, integrated sensor and communication, AI-based smart connectivity, multi-party trust models with new cryptographic technologies, integrated terrestrial and non-terrestrial networks for global access, and green and sustainable network for low TCO and global sustainability."

can be seen in Table 1 and Figure 1, one of the main advances in what can be identified as the characteristics of each generation is the increase in the speed and volume of data that must be transmitted until the production of LTE. It can be said that after achieving the IMT-2020 requirements for 5G, the research community had already considered the B5G/B5G era and anticipated the requirements of this generation. Figure 3 shows the key technologies predicted to drive the mobile industry towards 5G/B5G.

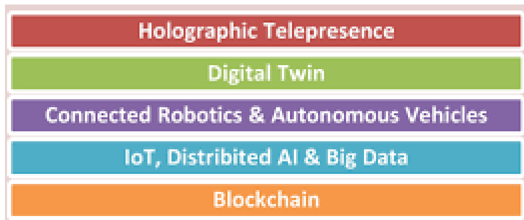

Figure 3: Expected capabilities and services in 5G/B5G generations [12]

With the development of generations of mobile networks, one of the most important concepts that is being considered today and many studies are being conducted on it is the self-organization of mobile networks [11]. In these networks, the goal is to manage, configure and modify the network structure automatically [10]. In this context, machine learning techniques and artificial intelligence will play a fundamental role. The main goal is to repair the network structure itself in the event of defects in the network cells. Network management is for four main operations: 1. Identification, 2. Diagnosis, 3. Compensation and 4. Improving network efficiency. In fact, considering the field of self-healing, the main focus of this technology can be considered the management of defective cells, which in two parts deals with the detection and identification of defective cells and the other main sub-section of this field deals with the ability to compensate for disorders and damage to cells [10]. A detailed examination of each of the above can be done using different statistical algorithms. But the important point is the lack of capability and efficiency of statistical algorithms using proven theories in networks with complex and heterogeneous structural models. Therefore, to increase the accuracy and efficiency in examining networks in terms of identifying, diagnosing and managing their cells, the use of artificial intelligence and learning algorithms using approximate numerical approaches has been considered [13-14]. Machine learning algorithms with different approaches can examine the cells of networks with their desired goals with higher flexibility than classical statistical algorithms based on proven theorems. The three main machine learning models used in the analysis and investigation of self-organizing networks in each field of this technology include unsupervised learning, supervised learning and reinforcement learning approach [15]. To continue the work ahead, in the second II, activities related to identifying defective cells using learning algorithms are discussed.

The third part explains the proposed method. In the fourth part, the results of the proposed algorithm are reviewed and analyzed. Finally, in the fifth section, a summary of the implementation operations and general overview of the presented activity is discussed.

2- Introduction to theoretical foundations and research history

Identification of defective cells in networks is usually determined based on three general approaches. In the diagram of Figure 4, the generality of these approaches is shown separately.

Table2: General methods of solving the problem of 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

Considering the field of self-healing, whose main goal is to identify and diagnose failures and compensate for defective cells, it is very important to use a method that can have sufficient accuracy, speed, and stability in field of performance. Usually, the use of Machine Learning Technique has a high compatibility with parametric mapping for use in any part of this field due to the mechanisms they have. An important issue in machine learning methods is their complete dependence on the type of data that exists to solve the problem. The problem model in these methods is directly dependent on the sample space in the dataset, because the data used for this purpose is received and analyzed from three heterogeneous levels. 1: Site level, 2: Cell level, and 3: User level. Also, the received data is generally in the areas of determining voice and data quality and network signaling, and a healthy cell must be in all three areas. In good condition, the three unsupervised learning models, supervised learning, and reinforcement learning form the basis of learning problems, and less research has been done in all three areas to determine whether the network is healthy or defective. What problems can be solved with what structure and with what model is one of the most important issues that has been investigated [10]. In general, the application of each learning pattern to solve the problem of identifying and compensating for defective cells in networks, according to the conditions and purpose of the problem, can be according to the framework shown in Figure 4.

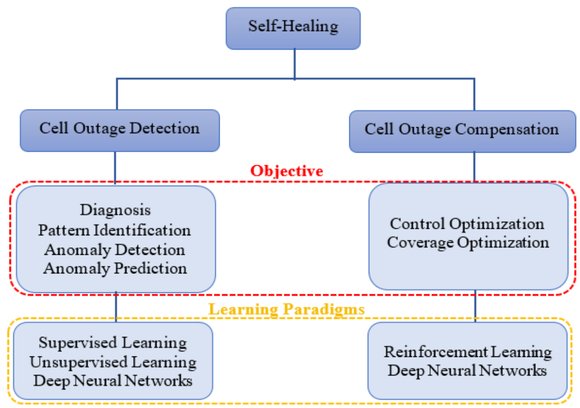

Figure 4: The generality of learning patterns in the management of cells for the field of self-healing of the network

Supervised learning is usually the main approach in identifying these cells. One of the algorithms used in networks for fault detection in other target operations in the network is the decision tree [16]. 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. [17]. Using this algorithm, radio link failures in networks have been investigated. Another activity by Hangli and Zhuqiao [18] 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 [19], which was used to classify and predict faulty cells at the network level. Logistic regression [20] 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 [21-22] is the application of logistic regression in Internet networks. For example, Kulkarni et al. [21] used the logistic regression algorithm to identify failures and errors in the performance of 5G Internet network antennas. By using this algorithm, network indicators were simulated to estimate failures and errors in antenna performance.

The random forest method [23] 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 [24-26], 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 [24] 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 [27-29]. For example, in the research [27] and [28], researchers used a combination of KNN and SVM 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 [30-32], 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, supervised 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 [33-34] 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. Therefore, what is important in solving the problem in question is natural.

1. Imbalance in the available data in terms of the considered class
2. Insufficient amount of data in activities with intended goals
3. Multi-source (the presence of high variance in the data level)

4. Low sensitivity of the model in determining the exact class (bias problem)
5. Lack of real-time efficiency

Knowing the exact approach to confronting each of the existing challenges can improve the efficiency of the implementation algorithms to solve the desired problem. Considering the challenges mentioned above, the aim of this research is to use a hybrid learning approach using goal-oriented algorithms. Therefore, the problem of high data variance, imbalance in the number of data samples with different classes, and also the problem of model bias can be reduced to the optimum level. Also, based on the announced materials, all research in this field is based on receiving or not receiving network-level indicators or changes in the indicators of neighboring cells. In this research, by using the TOPSIS technique as one of the multi-criteria decision-making methods, all the indicators received from different network levels and their positive and negative effects on the performance of a network can be considered with appropriate actions. The cell weights of the proposed method are explained in the next section.

3- The proposed algorithm

In this section, the proposed method for identifying defective cells in the network is described. Figure 5 shows the diagram of the proposed method. As shown in the diagram, the proposed algorithm consists of two main parts. The first part is related to data preprocessing and preparation, and the second part is related to data sample classification operations. In the second part, the model training process is performed using the stacking learning algorithm [35] in two levels. The first level is related to the training of basic models, and the second level is the training of the main model, which is the adaptive reinforcement method [36]. Each section is explained below.

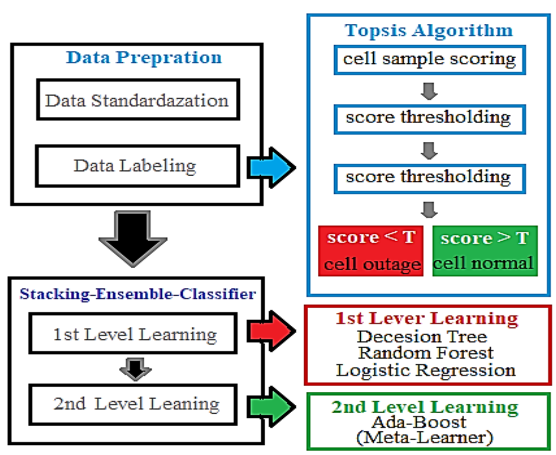

Figure 5: overview of the proposed method

3-1- Data preparation

Data preparation is performed in most learning operations with different goals and different learning domains, which is mostly considered as preprocessing before starting the main part of the training process. Data preprocessing involves manipulating features such as equalization and mapping a value to a specific time interval. Another thing that is common for collected data is labeling the samples and assigning them a class for use in a supervised learning model.

3-2- Data standardization

The main purpose of preprocessing is to clean the data structure from noise. For numerical data, data cleaning usually involves processes such as normalization and standardization of the data. With normalization and standardization operations, the spurious effect of features that have values that are much larger or much smaller than usual compared to other features is removed.

3-3- Data labeling

According to the purpose of the problem of detecting defective cells, the dataset used must have binary labels. Multi-criteria decision-making methods include multi-objective and multi-criteria decision-making categories. Decision-making involves selecting the best option or considering decision-making factors. Each decision-making method has a specific task, one is to measure the criteria, one is to rank the options and the other is to evaluate the criteria. For labeling the dataset, a multi-criteria technique called TOPSIS [37] has been used. The reason for using the TOPSIS technique is to weight each criterion to determine the quality of a cell, and also the options are ranked based on their similarity to the ideal state, so that the more a cell is similar to the ideal state, the higher it is. The ranking of this decision-making method has strong mathematical support. Cell labeling consists of three general steps: 1. Cell scoring, 2. Cell ranking and 3. Determining a value as a separator between healthy and defective cells and labeling the desired data sample. Here, the threshold value set was considered to be 0.3. This value was obtained by considering the weight of each indicator used and the positive or negative coefficient of that indicator in the type of network performance.

3-4- Stacked hybrid learning algorithms (first level)

The basic algorithms described in the second part perform the classification operation of a sample as basic models and independently of each other in the stack combination approach. Each model estimates the output class of the sample. Finally, the output class is determined using the majority vote at the first level. The stack combination method, unlike the ensemble combination method, uses different algorithms to classify the data. The aim is to deal with the problem of bias. For the basic algorithms selected for training at the first level of the algorithm stack, the following main features were considered:

- The decision tree method is robust to noise.
- The logistic regression method works well for two-class data. It also has effective learning compared to the unbalanced number of data samples in terms of classes.
- The random forest method uses the ensemble combination approach due to its structural nature. Therefore, the variance problem in this algorithm is solved.

3-5- Adaptive learning algorithm (second level)

Using an adaptive learning algorithm to improve the weight of features during learning. This learning model is used according to the way it works, which gives more importance to more difficult examples at each stage of training. The adaptive method eliminates the effect of noise in this way. On the other hand, by giving importance to the examples, it solves the problem of variance in the training process. Another point is that since the model structure is considered more complex at each stage, the model bias is also automatically adjusted. Finally, considering the stated structure, the adaptive method is very suitable for binary classification. Considering what was described as the proposed method, the following important points in choosing the method and algorithms used can be stated:

Algorithms used are all resistant to noise.

- The logistic regression classification method has a good performance in situations where the dataset is mixed (one class has a much higher sample size than the other).
- Each of the random forest and decision tree algorithms is highly compatible with each other's training structure.
- The random forest classification method in the first level of training, and the adaptive method in the second level of training each solve the variance challenge.
- The generality of the stack method solves the model bias problem for the proposed method. Due to the use of random and adaptive forest algorithms, an attempt has been made to solve the challenges of data variance as well as bias.

According to the above points, it can be said that the adopted approach will somehow solve the three main challenges. The obtained results will show this issue. In the next section, the results are analyzed and reviewed.

4- Evaluation of the proposed solution

For the current part, the results obtained from the proposed algorithm are presented. At first, the dataset used in the research are introduced. Then, in the first part of the results review, the comparison of the output obtained from the execution algorithm compared to the basic methods is stated.

4-1- Dataset

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 TOPSIS 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 2. Also, the coordinates of the data are stated in Table 3.

Table 3: Features used in the data

Features(Kpi)s	
ERAB succes 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.
S1Signal_E-RAB_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.
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).
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.
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.
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.
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.
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.
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.
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.
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.
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).

Table 4: Coordinates of dataset samples

defective	normal	number	group
2696	51000	53696	Training
703	12721	13424	Testing
3399	63721	67120	Total

4-2- Comparison

In this section, we will express the results obtained on the data used in the proposed method compared to the basic methods of machine learning, as well as the results of some related methods. Table 4 shows the results obtained for the four most important criteria in identifying defective cells. We examined the results of the proposed method according to the imbalance of the data on the criteria of accuracy, precision and recall. We compared it with other methods for accuracy criteria. We have also shown the confusion matrix for the proposed method, which shows the high accuracy of the proposed method in determining the target data class, which is defective cells. As shown in Table 4, the proposed method with stacking approach has been more accurate than other methods. The main reason for increasing the accuracy of the proposed model with the stacking approach is the use of the adaptive model in the second level of the hybrid approach algorithm. Also, the used approach reduces the amount of variance and bias in the proposed model.

A comparison between other algorithms and the proposed method has been made in four criteria, which are:

●Accuracy:

Accuracy is one of the most widely used metrics in evaluating the performance of machine learning algorithms, especially in classification problems. It indicates what percentage of the model's predictions were correct.

●Complexity:

Algorithm complexity refers to the amount of computational resources required to run an algorithm in relation to the size of the inputs. This metric is typically divided into two parts:

Time Complexity

Space Complexity

●Scalability:

Scalability refers to the ability of a machine learning algorithm to handle and work with large or growing data. In other words, a scalable model should be able to effectively maintain or even improve its performance as the data size or problem complexity increases.

● Adaptivity (for unbalanced dada):

Adaptivity refers to the ability of machine learning algorithms to adjust and improve their performance in specific situations. One particularly challenging situation is unbalanced data, where the number of examples in different classes varies significantly.

Table 5: The results of the proposed method with basic methods and a number of related activities[42]

Method (simple base method)	Accuracy	Complexity	Scalability	Adaptivity (for unbalanced dada)
KNN	87.13	Low	Low	Low
SVM	89.66	High	Fair	Fair
Logistic Regression	81.71	High	Low	Low
Multi Layer Perceptron	88.76	High	Fair	Fair
Decision Tree	88.05	Low	Fair	Fair
Random Forest	93.11	High	Fair	Fair
Naïve Bayes	92.38	Low	Fair	Fair
Kulkarni et. all[19]	95	Low	Fair	Fair
Abd El-Aziz [24]	99	High	Fair	Fair
Erdogan & biling [38]	87.61	Low	Low	Low
Yu et. All [39]	91.77	Low	Low	Fair
Thing [40]	86.71	High	Low	Low
Babalola et. all [41]	92.71	Low	Fair	Fair
(Proposed Method) Stacking Ensemble	99.92	High	High	High

As seen, the proposed method has a very high accuracy in identifying defective cells at the network level. Due to the unbalanced sample of the classes and the very small number of defective cell classes in the data, we calculated the values of accuracy ,recall, and f-measure and precision along with the accuracy criteria to validate the accuracy of the proposed model. According to the obtained results, the precision value was equal to 99.40, the recall value was equal to 98.82 and the f-measure value was also equal to 99.86. Also, the confusion matrix for the test samples is shown in Figure ۶. As it is known, the correct detection value of the target class for the defective class is very ideal.

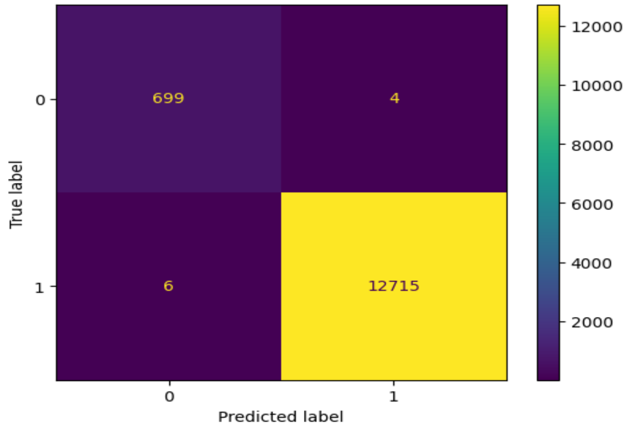

Figure 6: Confucion matrix for test samples

5- Conclusion and future works

In this research, using a machine learning algorithm with a combined learning approach called stacking, the estimation of defective cells at the network level was done. According to the type and structure of the available data, the purpose of presenting the mentioned learning algorithm was to solve the three existing challenges. 1. Imbalance in the number of samples with existing classes, which can affect the training process. 2. The second existing challenge that arose as a result of the first challenge was the problem of bias in the proposed model. It means what algorithm to use and with what requirements. 3. Finally, the third challenge is related to the dispersion and the amount of data variance that exists in the data collection procedure. To solve the three mentioned challenges together, the first step was to use learning algorithms that can have the necessary efficiency due to the imbalance of data abundance. The second step was to use an approach that could completely solve the problem of model bias. For this purpose, the solution of using learning of combined models such as a stack was done. In addition to the bias problem, another problem was related to the variance of the data. For this purpose, in the classification process, the stack algorithm was used in the first level and the adaptive algorithm was used in the second level to control the bias. The results of running the algorithm on the available data showed that the proposed method can successfully manage the three existing challenges. To validate the results, in addition to the accuracy criterion that was reported in general, accuracy precision recall, and f-measure were also analyzed. Also, the correct positive and correct negative detection rates were expressed using the clutter matrix to show the efficiency of the algorithm in accurately detecting the class of defective cells. The performed activity has a fundamental advantage over related activities performed in the field of deep learning with high accuracy. Unlike deep learning methods that are often not interpretable and require a lot of data, in the proposed implementation method, high data was not used in the classification process. On the other hand, the speed of training in existing learning was much higher than in deep learning. Also, as another advantage, it can be stated that the proposed method compared to the type and use of the basic algorithms used caused the interpretability of the method compared to the used algorithms and data.

References

[1] U. Ghafoor, M. Ali, H. Z. Khan, A. M. Siddiqui, and M. Naeem, "NOMA and future 5G & B5G wireless networks: A paradigm," Journal of Network and Computer Applications, vol. 204, p. 103413, Aug. 2022, doi: 10.1016/j.jnca.2022.103413.

[2] A. Sufyan, K. B. Khan, O. A. Khashan, T. Mir, and U. Mir, "From 5G to beyond 5G: A Comprehensive Survey of Wireless Network Evolution, Challenges, and Promising Technologies," Electronics, vol. 12, no. 10, p. 2200, May 2023, doi: 10.3390/electronics12102200.

[3] K. Samdanis and T. Taleb, "The Road beyond 5G: A Vision and Insight of the Key Technologies," IEEE Network, vol. 34, no. 2, pp. 135–141, Mar. 2020, doi: 10.1109/mnet.001.1900228.

[4] A. Nauman, T. N. Nguyen, Y. A. Qadri, Z. Nain, K. Cengiz, and S. W. Kim, "Artificial Intelligence in Beyond 5G and 6G Reliable Communications," IEEE Internet of Things Magazine, vol. 5, no. 1, pp. 73–78, Mar. 2022, doi: 10.1109/iotm.001.2100140.

[5] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-Intelligence-Enabled Intelligent 6G Networks," IEEE Network, vol. 34, no. 6, pp. 272–280, Nov. 2020, doi: 10.1109/mnet.011.2000195.

[6] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge Artificial Intelligence for 6G: Vision, Enabling Technologies, and Applications," IEEE Journal on Selected Areas in Communications, vol. 40, no. 1, pp. 5–36, Jan. 2022, doi: 10.1109/jsac.2021.3126076.

[7] I. F. Akyildiz, A. Kak, and S. Nie, "6G and Beyond: The Future of Wireless Communications Systems," IEEE Access, vol. 8, pp. 133995–134030, 2020, doi: 10.1109/access.2020.3010896.

[8] K. M. Bin Hasan, M. Sajid, M. A. Lapina, M. Shahid, and K. Kotecha, "Blockchain technology meets 6 G wireless networks: A systematic survey," Alexandria Engineering Journal, vol. 92, pp. 199–220, Apr. 2024, doi: 10.1016/j.aej.2024.02.031.

[9] S. A. Abdel Hakeem, H. H. Hussein, and H. Kim, “Vision and research directions of 6G technologies and applications,” Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 6, pp. 2419–2442, Jun. 2022, doi: 10.1016/j.jksuci.2022.03.019.

[10] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, “A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks,” IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2392–2431, 2017, doi: 10.1109/comst.2017.2727878.

[11] T. Zhang, K. Zhu, and E. Hossain, “Data-Driven Machine Learning Techniques for Self-Healing in Cellular Wireless Networks: Challenges and Solutions,” Intelligent Computing, vol. 2022, Jan. 2022, doi: 10.34133/2022/9758169.

[12] F. B. Saghezchi, J. Rodriguez, Z. Vujicic, A. Nascimento, K. M. S. Huq, and F. Gil-Castiñeira, “Drive Towards 6G,” in Enabling 6G Mobile Networks, Cham: Springer International Publishing, 2021, pp. 3–35. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-74648-3_1

[13] S. Sridharan, “Machine Learning (ML) In a 5G Standalone (SA) Self Organizing Network (SON),” arXiv.org, Nov. 24, 2020. https://arxiv.org/abs/2011.12288

[14] H. Fourati, R. Maaloul, L. Chaari, and M. Jmaiel, “Comprehensive survey on self-organizing cellular network approaches applied to 5G networks,” Computer Networks, vol. 199, p. 108435, Nov. 2021, doi: 10.1016/j.comnet.2021.108435.

[15] A. G. Papidas and G. C. Polyzos, “Self-Organizing Networks for 5G and Beyond: A View from the Top,” Future Internet, vol. 14, no. 3, p. 95, Mar. 2022, doi: 10.3390/fi14030095.

[16] B. Charbuty and A. Abdulazeez, “Classification Based on Decision Tree Algorithm for Machine Learning,” Journal of Applied Science and Technology Trends, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[17] W. Li and Z. Qi, “Network selection algorithm based on decision tree in heterogeneous wireless networks,” MATEC Web of Conferences, vol. 189, p. 04010, 2018, doi: 10.1051/matecconf/201818904010.

[18] Nethraa Sivakumar, Pooja Srinivasan, Nikhil Viswanath, and Venkateswaran N, “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.

[19] U. Saeed, S. U. Jan, Y.-D. Lee, and I. Koo, “Fault diagnosis based on extremely randomized trees in wireless sensor networks,” Reliability Engineering & System Safety, vol. 205, p. 107284, Jan. 2021, doi: 10.1016/j.ress.2020.107284.

[20] A. A. T. Fernandes, D. B. Figueiredo Filho, E. C. da Rocha, and W. da S. Nascimento, “Read this paper if you want to learn logistic regression,” Revista de Sociologia e Política, vol. 28, no. 74, 2020, doi: 10.1590/1678-987320287406en.

[21] A. M. Kulkarni, G. Saini, and S. S. Pattnaik, “Antenna Array Fault Detection Using Logistic Regression Technique,” in Communications in Computer and Information Science, Cham: Springer Nature Switzerland, 2023, pp. 13–29. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1007/978-3-031-48781-1_2

[22] V. Sharma, T. Hong, V. Cecchi, A. Hofmann, and J. Y. Lee, “Forecasting weather-related power outages using weighted logistic regression,” IET Smart Grid, vol. 6, no. 5, pp. 470–479, Aug. 2023, doi: 10.1049/stg2.12109.

[23] Y. Liu, Y. Wang, and J. Zhang, “New Machine Learning Algorithm: Random Forest,” in Information Computing and Applications, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 246–252. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-34062-8_32

[24] L. A. El-Aziz et al., “Cell Outage Detection and Degradation Classification Based on Alarms and KPI’s Correlation,” in 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), Oct. 2020. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/niles50944.2020.9257920

[25] A. K. Balyan et al., “A Hybrid Intrusion Detection Model Using EGA-PSO and Improved Random Forest Method,” Sensors, vol. 22, no. 16, p. 5986, Aug. 2022, doi: 10.3390/s22165986.

[26] F. S. Samidi, N. A. Mohamed Radzi, K. H. Mohd Azmi, N. Mohd Aripin, and N. A. Azhar, “5G Technology: ML Hyperparameter Tuning Analysis for Subcarrier Spacing Prediction Model,” Applied Sciences, vol. 12, no. 16, p. 8271, Aug. 2022, doi: 10.3390/app12168271.

[27] J. Ali-Tolppa, M. Kajo, B. Gajic, I. Malanchini, B. Schultz, and Q. Liao, “Cognitive Autonomy for Network Self-Healing,” Towards Cognitive Autonomous Networks, pp. 345–384, Oct. 2020, doi: 10.1002/9781119586449.ch9.

[28] O. Onireti et al., “A Cell Outage Management Framework for Dense Heterogeneous Networks,” IEEE Transactions on Vehicular Technology, vol. 65, no. 4, pp. 2097–2113, Apr. 2016, doi: 10.1109/tvt.2015.2431371.

[29] A. Zoha, A. Saeed, A. Imran, M. A. Imran, and A. Abu-Dayya, “A learning-based approach for autonomous outage detection and coverage optimization,” Transactions on Emerging Telecommunications Technologies, vol. 27, no. 3, pp. 439–450, Aug. 2015, doi: 10.1002/ett.2971.

[30] M. Z. Asghar, F. Ahmed, and J. Hamalainen, “Artificial Intelligence Enabled Self-healing for Mobile Network Automation,” in 2021 IEEE Globecom Workshops (GC Wkshps), Dec. 2021. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/gcwkshps52748.2021.9681937

[31] Y. Ruan, Y. Wang, and Y. Tang, “An Intelligent Cell Outage Detection Method in Cellular Networks,” in 2021 16th International Conference on Computer Science & Education (ICCSE), Aug. 2021. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/iccse51940.2021.9569511

[32] P. Yu, H. Wang, and Z. Chen, “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), Jun. 2022. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/bmsb55706.2022.9828778

[33] P. Yu et al., “Deep Reinforcement Learning Aided Cell Outage Compensation Framework in 5G Cloud Radio Access Networks,” Mobile Networks and Applications, vol. 25, no. 5, pp. 1644–1654, Jul. 2020, doi: 10.1007/s11036-020-01574-8.

[34] M. Iwamoto, A. Suzuki, and M. Kobayashi, “Deep Reinforcement Learning Based Antenna Selection for Cell Outage Compensation,” in ICC 2023 - IEEE International Conference on Communications, May 2023. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/icc45041.2023.10279017

[35] S.-A. N. Alexandropoulos, C. K. Aridas, S. B. Kotsiantis, and M. N. Vrahatis, “Stacking Strong Ensembles of Classifiers,” in IFIP Advances in Information and Communication Technology, Cham: Springer International Publishing, 2019, pp. 545–556. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-19823-7_46

[36] P. Beja-Battais, “Overview of AdaBoost : Reconciling its views to better understand its dynamics,” arXiv.org, Oct. 06, 2023. https://arxiv.org/abs/2310.18323

[37] S. Chakraborty, “TOPSIS and Modified TOPSIS: A Comparative Analysis,” Open Engineering Inc, Apr. 2021. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.31224/osf.io/y39j7

[38] S. Z. Erdogan and T. T. Bilgin, “A data mining approach for fall detection by using k-nearest neighbour algorithm on wireless sensor network data,” IET Communications, vol. 6, no. 18, pp. 3281–3287, Dec. 2012, doi: 10.1049/iet-com.2011.0228

[39] X. Yu, H. Chen, W. Zhao, and L. Xie, “No-Reference QoE Prediction Model for Video Streaming Service in 3G Networks,” in 2012 8th International Conference on Wireless Communications, Networking and Mobile Computing, Sep. 2012. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/wicom.2012.6478588

[40] V. L. L. Thing, “IEEE 802.11 Network Anomaly Detection and Attack Classification: A Deep Learning Approach,” in 2017 IEEE Wireless Communications and Networking Conference (WCNC), Mar. 2017. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1109/wcnc.2017.7925567

[41] O. P. Babalola and V. Balyan, “Cell Outage Detection in 5G Self-organizing Networks Based on FDA-HMM,” in Proceedings of Fourth International Conference on Computing, Communications, and Cyber-Security, Singapore: Springer Nature Singapore, 2023, pp. 119–130. Accessed: May 21, 2024. [Online]. Available: http://dx.doi.org/10.1007/978-981-99-1479-1_9

[42] [63] 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.