Ensemble Learning based Supervised Learning approach for designing of 5G/B5G Classification System

Ramaswamy T[1] , S.P.V.Subba Rao[2] , C Raj Shubhang Reddy[3] , Shashi Bhanu Eklure[4]

Associate Professor[1] , Professor[2] , UG Scholar[3] , UG Scholar[4]

[1,2,3,4]Department of Electronics and Communication Engineering , Sreenidhi Institute of Science and Technology,Hyderabad

ramaswamyt@sreenidhi.edu.in, spvsubbarao@sreenidhi.edu.in , crajshubhangreddy@gmail.com,bhanushashi6@gmail.com

Abstract: The emergence of modern wireless technologies has posed a challenge in front of the network service providers as they need to ensure excellent quality of service as well as seamless experience to it's users. The focus on quality of user experience has increased with the introduction of the 5G technology. The Service Classification needs to be done using various parameters which includes several performance Indicators as well as quality indicators .Machine Learning techniques have become a powerful tool in variety of real time applications, for this project we have employed Ensemble learning based supervised learning approach to understand the feasibility of 5g service classification using ensemble learning . Use of separate machine learning algorithms has provided encouraging results but and ensemble learning approach using voting classifier provided us with use of various algorithms to provide us an efficiency of about 96**%.**

Keywords: 5G,Voting classifier, Ensemble Learning , quality of experience.

Introduction:

The next generation of wireless networks, such 5G and Beyond 5G, are developing quickly to deliver more dependable and compatible services at extremely high speeds utilising the existing internet infrastructure and technology. Due to its extremely fast transmission rates and great network throughput, 5G networks are seen as the next significant advancement in wireless communication. However, because of the large number of unique services and the lack of human expertise, it is challenging to manually manage these services given intricacy, adaptability, and vigour of 5G/B5G networks. Changes in behavioural patterns limit the ability to identify network activity. As a result, computational support is needed to manage the services and operations. But because of the complexity and diversity of the different services, finding the software or technology that is best suited for certain use cases is quite challenging for the network service providers and operators

The fact that 5G technology can be used to deliver a wide range of services with diverse requirements and characteristics leads to the necessity for the classification of distinct 5G services. In order to develop and optimise 5G networks to meet these objectives, different types of services have varied requirements for bandwidth, latency, dependability, and other performance indicators..

Since 5G is a particular kind of network that focuses on services, classifying those services requires an efficient strategy for allocating network resources. Due to their features and requirements, the expansion and variety of the many services provided in 5G networks have seen an explosive growth. Artificial intelligence (AI) has the potential to be employed in this situation to assist with the cognitive decisions regarding 5G/B5G[3].It can be tricky to monitor and regulate network resources, foresee and avoid Service Level Agreement (SLA) violations[1], It may harm both the QoS operation and the QoE consumers perceive, if these services aren't organised with the use of AI or ML. This makes the entire process of management punishing for the operators.

The right service categorization gives a technique to improve the network's QoS and QoE, which makes it crucial in the field of wireless communication and, in this case, in the implementation of 5G/B5G networks. Users look for services that need to be explicitly grouped so that network administrators may select network slices for various services in order to enhance the network's QoS, and improve users' impressions regarding the network's QoE, and establish SLAs for each slice. Both the user's perspective and the perspective of service providers can definitely be used to analyse the crucial 5G needs. KPIs are the primary element that 5G service classification systems now employ whilst carrying out the classification process. Network slicing, a key 5G concept, allows network operators to virtually divide several networks from a single physical framework by assigning resources logically and creating connections between the respective pieces. This makes it possible for the network and services to be deployed effectively to meet the needs and wants of the customers.[6]

We propose a Supervised Machine Learning (SML)-based classifier for 5G/B5G services approach, with a special focus on enhancing QoS and QoE, to solve and enhance the problem of managing 5G networks and the vast range of services. To provide enhanced classification, The KPIs and KQIs of the various services are used to form the foundation of this system[1].

The feedback provided after a new service is categorised, retraining proactively and then creating a fresh prognostic model by merging its KPI/KQI parameters in the data warehouse, is one contribution of our approach. to strengthen it in relation to the previously created prediction model from a structural perspective[1].

III. Methodology

1. Efficient resource distribution: Since bandwidth, processing power, and energy in 5G networks are limited, it's critical to distribute these resources across various service types according to their needs. Network operators can prioritise resources depending on the demands of each

service by categorising services into separate groups, which can improve performance and make better use of network resources.

2. Guarantees for quality of service (QoS) In order to provide QoS to customers, it is crucial to identify the various types of services because they each have different QoS requirements. Network operators can establish and implement QoS regulations that match the demands of each service by categorising services into various groups, which can help to increase customer satisfaction and minimise network traffic.

3. Service differentiation: Service differentiation is the capacity to provide various levels of service performance and quality to various clients. Network operators can offer various service packages and pricing models to various sorts of clients based on their needs and preferences by categorising services into distinct groups. This can help to boost revenue and client loyalty.

4. Careful network planning and optimisation are necessary to offer the requisite performance and dependability on 5G networks, which are complex. Network operators can design and optimise networks to match the demands of each service by categorising services into distinct groups. This can help to lower network congestion, improve coverage, and boost capacity.

In conclusion, effective resource allocation, QoS guarantees, service differentiation, and network design and optimisation depend on the classification of different 5G services. Network operators may build and optimise 5G networks to fulfil the varying needs of many types of services by categorising services into different groups, which can help to improve overall network performance and customer satisfaction.

Network operators can build numerous virtual networks on top of a single physical network infrastructure thanks to a technology called 5G network slicing. Each virtual network, or "slice," is tailored to satisfy the particular needs of a certain use case or application, such as automated manufacturing, augmented reality, or driverless vehicles. Network operators can provide tailored network services with certain performance characteristics, such as low latency, high dependability, or high bandwidth, by generating specialised slices for various applications.

Although the idea of network slicing is not new, 5G networks offer a far more adaptable and scalable platform for its implementation, allowing network resources to be distributed dynamically and on-demand based on each slice's demands.[4]

Network functions like core network functions, radio access functions, and transport functions are specific to each network slice. The service level agreements (SLAs) for each slice, which outline the performance measurements and guarantees for each slice, can also be defined and enforced by network operators.[4]

For network operators and users alike, 5G network slicing has a number of advantages. It allows network operators to use network resources more effectively, lowers capital and operating costs, and enables them to provide specialised network services to various client and application types. It offers users a more dependable, secure, and adaptable network service that can accommodate their unique needs and demands.[5]

1. Smart cities: Using 5G network slicing, specific network slices may be created for various smart city applications, such as traffic control, public safety, and waste management.

2. Industrial automation: 5G network slicing can be used to generate specialised network slices for applications like remote monitoring, robotics, and factory automation.

3. Virtual reality and augmented reality: 5G network slicing can be used to design specific network slices for these applications, which need high bandwidth and low latency to provide a seamless user experience

The Voting Classifier has been used for deploying the ensemble learning algorithm for the classification system . The voting classifier has been mainly designed by using three best performing machine learning algorithm namely the Logistic Regression Algorithm, SVM and the Random Forest Algorithm.

III.I Database

The dataset used in this project is created with the values as per ITU standards, The dataset has 165 rows and 14 columns which are having various standard values and has Key performance and Quality Indicators of 5G network [1]

III.II Computational Parameters

Several significant performance metrics for 5G services are defined below:

1. Latency (ms): The amount of time it takes for data to travel across a network before it is received. Faster response times and improved real-time performance result from lower latency.

2. Jitter (ms): The fluctuation in latency with passing time. Real-time applications like video streaming and online gaming may have problems due to high jitter.

3. Bit Rate (Mbps): This is the rate in megabits per second at which data is sent across the network. Faster data transfer speeds result from higher bit rates.

4. Packet loss rate (%) The percentage of data packets that are lost during transmission is known as the packet loss rate (%). Lower data transfer speeds and performance problems can result from higher packet loss rates.

5. Peak Data Rate DL (Gbps): The highest download speed the network is capable of supporting, expressed in gigabits per second.

6. Peak Data Rate UL (Gbps):The greatest upload speed that the network can sustain, expressed in gigabits per second, is known as the peak data rate UL (Gbps).

7.Mobility (km/h): The fastest a user can go while still being connected to the network. Better support for applications like self-driving automobiles or high-speed trains comes with greater mobility.

8. Reliability: The likelihood that a data transmission will be successful (in percentage terms). Less mistakes and improved performance result from higher reliability.

9. Service Availability (%): The proportion of time that users can access a service. Fewer service interruptions and improved user satisfaction result from higher availability.
10. Survival Time (ms): The amount of time a connection can last before it is shut down. Better support for applications like video calls or online gaming results from longer survival times.

11. Experienced Data Rate DL (Mbps): The average download speed users experienced over a specific time period, taking into account elements like packet loss and latency.

12. Experienced Data Rate UL (Mbps): The average upload speed users experienced during a specific time period, taking into consideration elements like packet loss and delay.

13. Connection interruption or unavailability duration, measured in milliseconds (ms). Better performance and fewer service interruptions result from shorter interruption times.

(actual)	false	(predicted)
true	negative	positive
Negative	a	b
positive	c	d

The efficiency of the algorithm has been computed on the basis of four main parameters[2]

i)Accuracy: The metric Accuracy is the simplest and most direct metric of evaluation of the system model as it refers to the total number of accurate prediction to the total number of predictions is called accuracy.[2]

$$Accuracy = \frac{a + d}{a + b + c + d}$$

ii) Precision: Precision is the ratio of accurate positive prediction to total number of correctly predicted outcomes.[2]

$$Precision = \frac{d}{b + d}$$

iii) Recall: Recall refers to the ratio of accurate positive prediction to the total number of Positive predictions.[2]

$$Recall = \frac{d}{c + d}$$

iv) F1 Score: F1 score refers to the average of precision and recall ,It is also a important parameter which helps is accuracy prediction of overall model.[1]

$$F1\ score = \frac{2 \times Precision \times recall}{Precision + recall}$$

v) Support: Support refers to as the number of samples of true response in each class of target values.

III.III Machine Learning Algorithms Deployed:

In our Analysis we have deployed a variety of machine learning algorithms which are found to be highly accurate inorder to form a ensemble learning voting classifier.

Logistic Regression Algorithm

In binary classification issues, where the objective is to predict a binary output variable (0 or 1) based on one or more input features, the supervised learning approach known as logistic regression is used. It is a statistical model that depicts the relationship between the input data and the output variable using the logistic function.When a binary classification is used, the two classes are sometimes referred to as the positive class and the negative class. This is how the logistic regression algorithm operates: it estimates the chance that the output variable belongs to a specific class. The linear combination of input data is transformed into a probability value between 0 and 1 using the logistic function, also referred to as the sigmoid function.

Random Forest Classifier Algorithm

The random forest algorithm is a reliable machine learning algorithm which actually contains a number of decision trees and the output of the model is the average of the decisions from the subsets.[7]
On basis of the defined parameters as in above the outcome for this algorithm is as follows,

Decision Tree Classifier Algorithm

A decision tree is a form of supervised learning technique that works best for classification problems. It is a treestructured classifier, with core nodes that represent the features and elements of the input data set, branches that reflect the decision rules, and leaf nodes that all represent the classification outcome.[1]

```
 precision         recall    f1-score support

           1       1.00       1.00       1.00
       4
           2       1.00       1.00       1.00
       5
           3       0.57       1.00       0.73
       4
           4       1.00       1.00       1.00
       4
           5       1.00       1.00       1.00
       5
           6       0.80       1.00       0.89
       4
           7       1.00       0.86       0.92
       7
           8       1.00       1.00       1.00
      12
           9       1.00       0.40       0.57
       5

    accuracy                             0.92
      50
   macro avg       0.93       0.92       0.90
      50
weighted avg       0.95       0.92       0.92
      50
```

Fig-II- Performance Metrics for Decision Tree Classifier Algorithm

Voting Classifier Model:

A voting classifier is an ensemble learning technique that enhances the overall efficiency and precision of the classification model by combining the predictions of various individual classifiers. A voting classifier works by combining the results of numerous base classifiers, which are often made up of several algorithms or variations of the same algorithm with varied parameters.Voting classifiers are frequently applied when noisy data, class imbalance, or overfitting prevent a single classifier from performing adequately on its own. A voting classifier can frequently outperform any solo classifier in terms of accuracy and generalisation performance by combining the predictions of numerous classifiers.

```
 precision    recall   f1-score    support


           1       1.00       1.00       1.00
  4
           2       1.00       1.00       1.00
  5
           3       0.57       1.00       0.73
  4
           4       1.00       1.00       1.00
  4
           5       1.00       1.00       1.00
  5
           6       1.00       1.00       1.00
  4
           7       1.00       1.00       1.00
  7
           8       1.00       0.75       0.86
 12
           9       1.00       1.00       1.00
  5


    accuracy                             0.94
 50
   macro avg       0.95       0.97       0.95
 50
weighted avg       0.97       0.94       0.94
 50
```

Fig-III- Performance Metrics of Voting Classifier Model

Analysis:

The proposed system is a voting classifier based system in which the prediction of the output is to be done on the basis of ensemble learning based on the best efficient algorithms i.e the logistic regression algorithm along with decision tree algorithm and the random forest classifier. The use of the an ensemble learning approach is to avoid overfitting of models. If a machine learning model predicts over hundred percent of efficiency it is said to be an overfitting model but with an ensemble learning approach we can overcome the problem also we can utilise the algorithms with excellent efficiency and accuracy. The compassion of F1 scores of all the models is defined as follows,

S.No	Machine Learning Model	F-1 score	Weighted average
1	Logistic regression	0.89	0.91
2	Random Forest Algorithm	0.962	0.969

3	Decision Tree Classifier	0.90	0.92
4	Voting Classifier Algorithm	0.95	0.94

Fig: Comparison of the Machine learning classifier models on basis of F-1 Score.

Conclusion & Future Scope :

With the emergence of new technologies like the Internet of Things, there is a service requirement of millions of channels in a single substation, increasing the need for improvement in service capability. The 5G and beyond 5G wireless communication technology would need to meet high expectations from the customer's perspective. The classification of the many vital services, which is based primarily on the various metrics as per requirements such as latency Jitter, can be done in a number of ways to improve the system's service capabilities. Although the findings of using machine learning techniques are promising, an ensemble learning strategy may be more effective for categorising these numerous techniques.

5G network slicing is a potent technology that gives network operators the ability to design unique network services for various customer and application kinds. It has several advantages, such as higher performance and dependability, more flexible and adaptable network infrastructure, and more effective use of network resources. Network slicing is anticipated to become more crucial in determining the future of the telecoms sector as 5G networks develop and mature.

References

[1] Preciado-Velasco, J.E.; Gonzalez-Franco, J.D.; Anias-Calderon, C.E.; Nieto-Hipolito, J.I.; Rivera-Rodriguez, R. 5G/B5G Service Classification Using Supervised Learning. *Appl. Sci.* **2021**, *11*, 4942. https://doi.org/10.3390/app11114942

[2] M. T. Khan and A. Adholiya, "Machine Learning-Based Application for Predicting 5G/B5G Service," *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 2023, pp. 613-619, doi: 10.1109/Confluence56041.2023.10048885.

[3] C. Zhang, Y. L. Ueng, C. Studer and A. Burg, "Artificial Intelligence for 5G and beyond 5G: Implementations Algorithms and Optimizations", *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 10, no. 2, pp. 149-153, 2020.

[4] S. Khan, S. Khan, Y. Ali, M. Khalid, Z. Ullah and S. Mumtaz, "Highly Accurate and Reliable Wireless Network Slicing in 5th Generation Networks: A Hybrid Deep Learning Approach", *J. Netw. Syst. Manag.*, vol. 30, no. 2, 2022.

[5] Klaine, P.V.; Imran, M.A.; Onireti, O.; Souza, R.D. A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 2392 2431.

[6] Mullins, M.; Taynann, R. Cognitive Network Management for 5G. 5GPPP Work. *Gr. Netw. Manag. QoS* **2017**, *1*, 1–40. Available online: **https://5g-ppp.eu/wp-content/uploads/2016/11/NetworkManagement_WhitePaper_1.0.pdf** (accessed on 26 May 2021).

[7] Demestichas, P.; Tsagkaris, A.G.K.; Vassaki, K.S. Service Classification in 5G Networks. November. Seoul, Korea. 2016, p. 13..