AI-Driven Optimization of 5G Resource Allocation for Network Efficiency

Team Members:

- Kanak Shrivastava (21BAI10185)
- Hardik Bansal (21BAI10010)
- Aman Preet Singh (21BAI10232)
- Sudhith N (21BAI10284)

1. INTRODUCTION

1.1 Project Overview

5G is divided into three frequency bands (low, mid, and high). Each band has different capabilities: the low band (less than 1GHz) has greater coverage but lower speeds, the mid band (1GHz–6GHz) offers a balance of both, and the high band (24GHz–40GHz) offers higher speeds but a smaller coverage radius.

5G is the latest evolution of cellular wireless connectivity and offers improved capacity, coverage, and lower latency. 5G offers many improvements compared to 4G but relies on the similar fundamentals to communicate with end user devices.

What makes 5G so different is the new levels of performance it offers. Similar to how 4G helped usher in the smartphone era, 5G will power new technologies across the enterprise. This project aims at automating allocation of resources so that the swiftness of 5G is made even more efficient.

1.2 Purpose

There are essentially three types of 5G bands supported in India — low-band, mid-band, and high-band (mm Wave) – based on frequency. Simply put, the higher the frequency, the better the speed and shorter the range of the network. A company will want to find the resource allocation by it’s Application Type, Signal Strength, Latency, Required Bandwidth, Allocated Bandwidth.

2. LITERATURE SURVEY

2.1 Existing problem

Upon the launch of 3G in India and worldwide, we celebrated the prospect of YouTube videos finally playing smoothly. With the advent of 4G, video calls and web meetings became a tangible reality. Now, we eagerly await the possibilities that 5G will unveil. What often escapes our notice is the profound influence of AI/ML technology in shaping the global 5G revolution, owing to its predictive and analytical capabilities. For effective functioning of 5G technology we require accurate resource allocation.

2.2 References

https://www.analyticsvidhya.com/blog/2023/02/different-ways-to-make-5g-services-better-using-ai/

https://www.hindawi.com/journals/wcmc/2022/1419930/

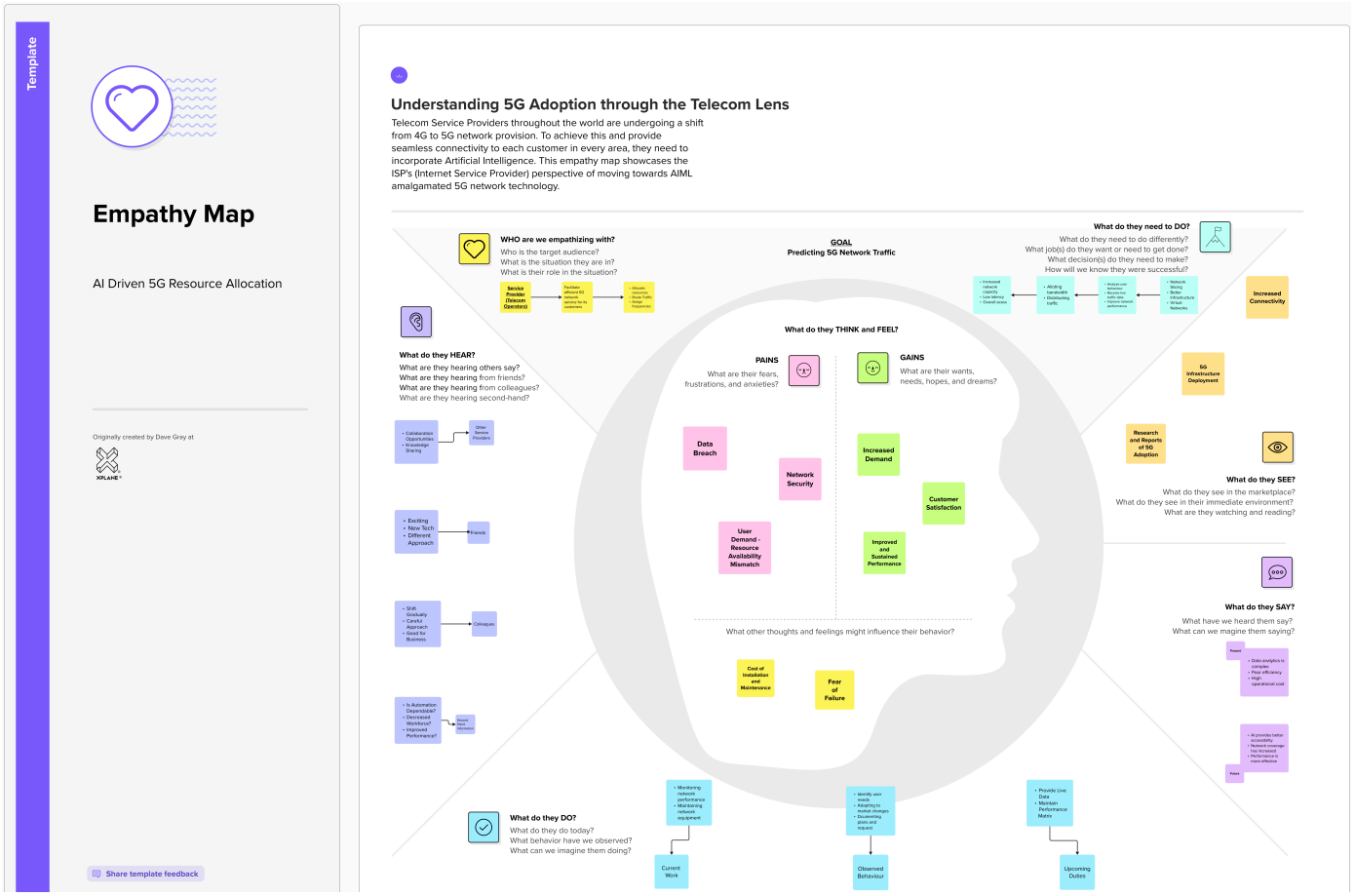
https://www.mdpi.com/2079-9292/11/1/121

2.3 Problem Statement Definition

"How can artificial intelligence be effectively utilized to optimize the allocation of resources in 5G networks in order to enhance network efficiency and performance?"

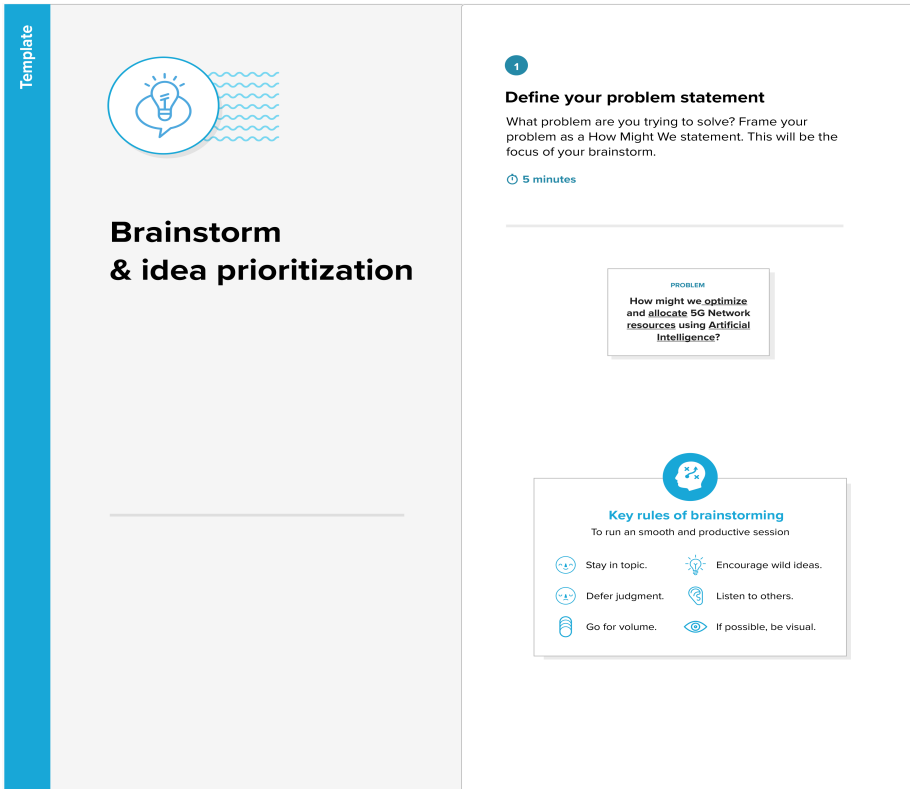
3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

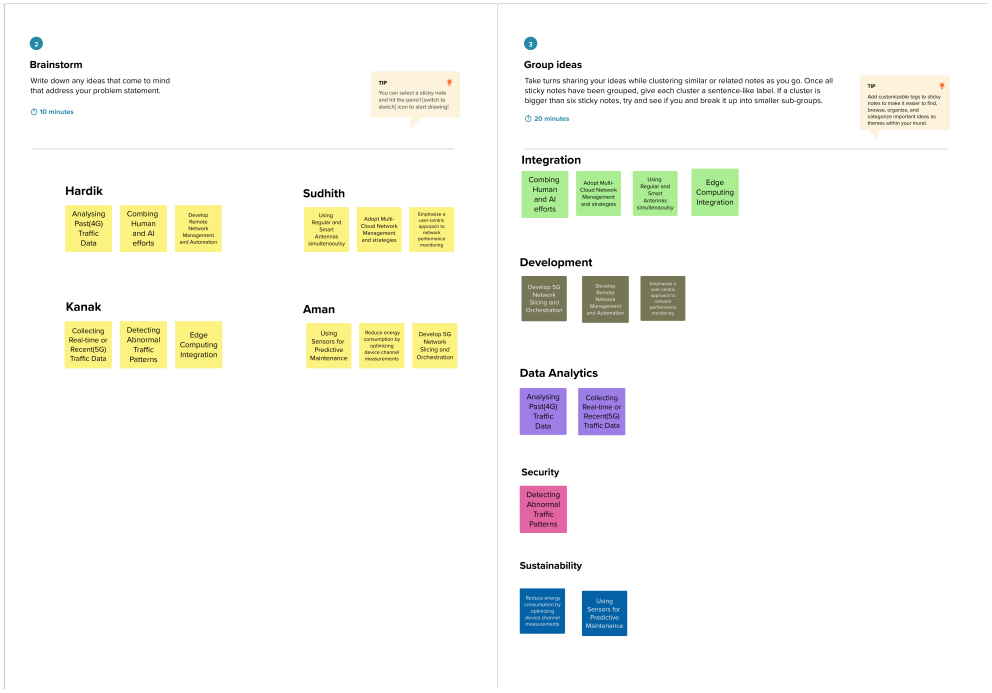


3.2 Ideation & Brainstorming

Step 1: Collaborate and Select the Problem Statement



Step 2: Brainstorm, Idea Listing and Grouping



Step 3: Idea Prioritization

4

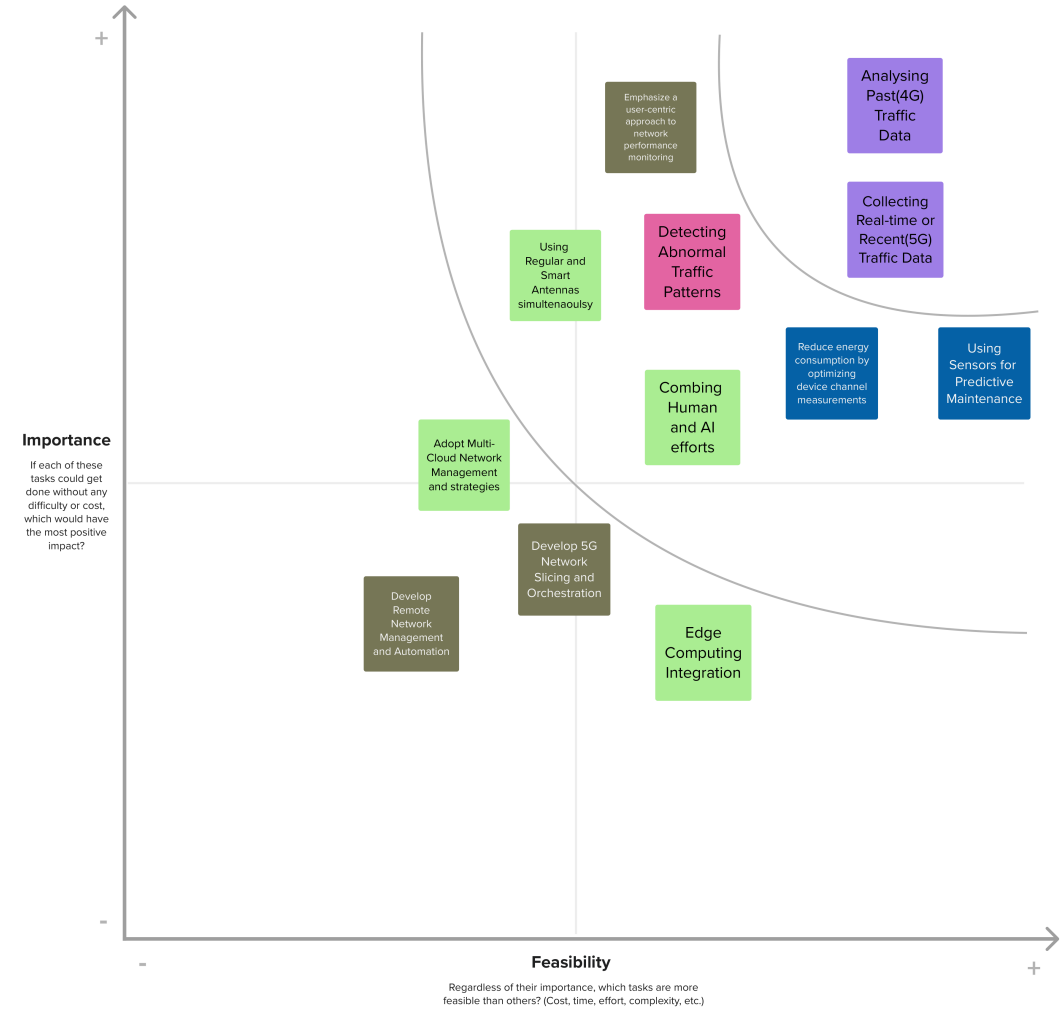
Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

20 minutes

TIP

Participants can use their cursors to point at where sticky notes should go on the grid. The facilitator can confirm the spot by using the laser pointer holding the **H key** on the keyboard.



4. REQUIREMENT ANALYSIS

4.1 Functional Requirements

The key functional requirements for the AI-driven 5G resource allocation system are:

- Predictive modeling - The system shall implement a machine learning model capable of predicting optimal resource allocation amounts given inputs like bandwidth, latency, signal strength etc.

- Dynamic optimization - The system shall continuously optimize allocations over time based on usage patterns and network conditions.

- Traffic analysis - The system shall ingest and process live network traffic data to understand usage.

- User profiling - The system shall utilize subscriber data like application usage to personalize allocations.

- Edge deployment - The system shall allow optimized model deployment on edge devices like base stations.

- Simulation testing - The system shall provide the ability to simulate allocation scenarios and data.

- Monitoring dashboard - The system shall provide visualizations and analytics of model performance, traffic, and allocations.

- API integration - The system shall expose APIs for integration into operator networks and provisioning systems.

4.2 Non-functional Requirements

- Scalability: The system shall scale to handle large networks with millions of subscribers and high traffic volumes.
- Latency: The system shall provide sub-millisecond latency for time-sensitive 5G applications.
- Availability: The system shall provide 99.999% uptime with redundancy and failover mechanisms.
- Security: The system shall secure data via encryption, role-based access, and other controls.
- Interoperability: The system shall integrate with common telco protocols like SNMP, Netconf, and 3GPP standards.
- Usability: The system shall provide an intuitive web interface and APIs for ease of use.
- Maintainability: The system shall be easy to maintain with modular architecture and test coverage.

5. PROJECT DESIGN

5.1 Proposed Solution:

Project Design Phase-I

Proposed Solution

Date	22 October 2023
Team ID	Team-592212
Project Name	Project - AI-Driven Optimization Of 5G Resource Allocation For Network Efficiency
Maximum Marks	2 Marks

Proposed Solution :

Project team shall fill the following information in the proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Create an AI driven approach to allocate resources in 5G networks ensuring performance, across frequency bands and applications.
2.	Idea / Solution description	Random Forest based Resource Allocation
3.	Novelty / Uniqueness	The application of cutting-edge deep learning techniques to solve resource allocation challenges in 5G networks.
4.	Social Impact / Customer Satisfaction	It ensures the efficient and optimized resource allocation in 5G networks and improved network performance, reduced latency, and a seamless user experience, ultimately contributing to enhanced connectivity, productivity, and overall satisfaction for users and businesses, in urban and remote areas alike.

5.	Business Model (Revenue Model)	In order to generate revenue there are options. One way is to offer solutions for resource allocation. Another option is to license our models. Additionally we can provide consulting services to help optimize 5G network performance for businesses. These opportunities allow us to create value and generate revenue, within the telecommunications and technology sectors.
6.	Scalability of the Solution	The Resource Allocation solution based on Random Forest regression provides a built-in ability to easily adjust to the increasing requirements and intricacies of 5G networks and extensive telecommunications infrastructures.

5.2 Solution Architecture:

Phase-II Solution Architecture

Date	22 October 2023
Team ID	Team-592212
Project Name	Project - AI-Driven Optimization Of 5G Resource Allocation For Network Efficiency
Maximum Marks	4 Marks

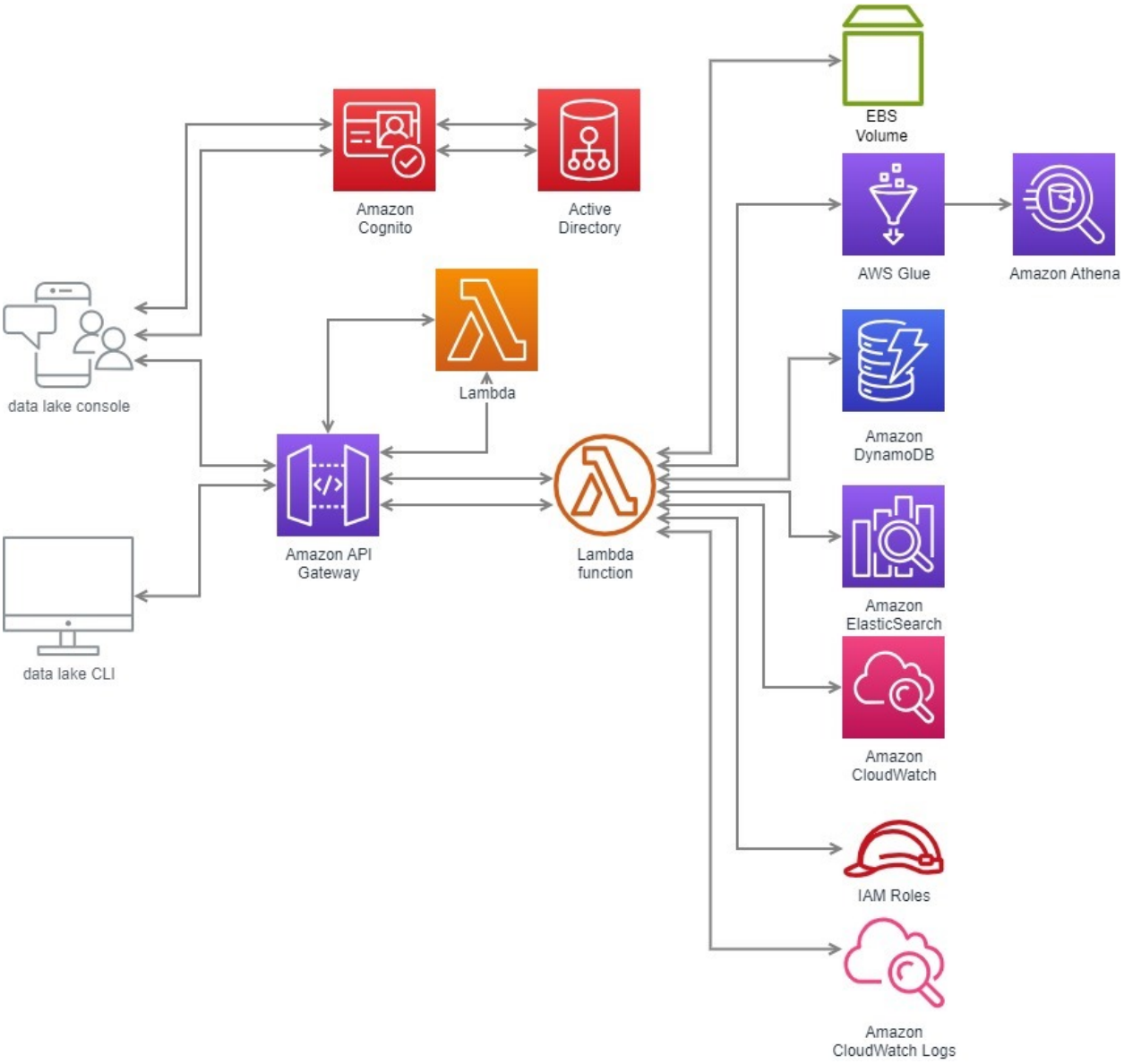
Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

1. Find the best tech solution to solve existing business problems.
2. Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
3. Define features, development phases, and solution requirements.

Provide specifications according to which the solution is defined, managed, and delivered.

Solution Architecture Diagram:



5.3 Data Flow Diagram and User Stories

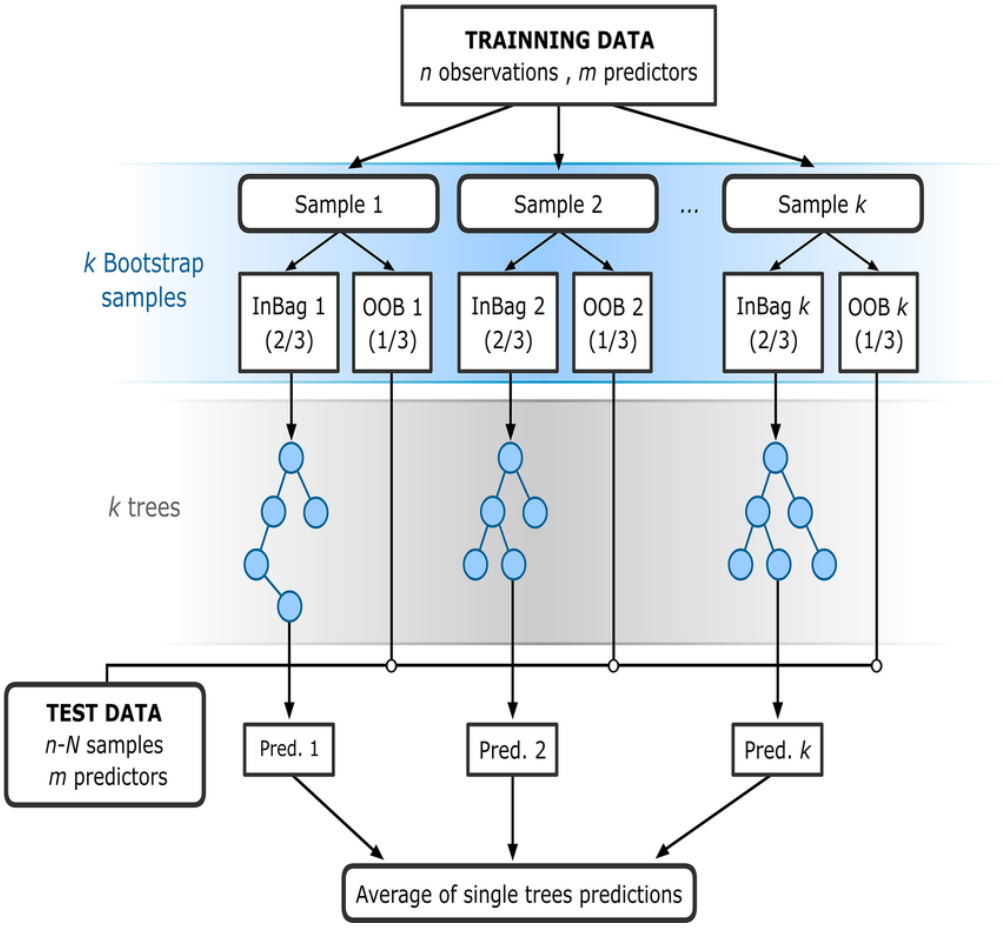
Phase-II

Data Flow Diagram:

Date	22 October 2023
Team ID	Team-592212
Project Name	Project - AI-Driven Optimization Of 5G Resource Allocation For Network Efficiency
Maximum Marks	4 Marks

Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



User Stories:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user) (Light and Rural)	Browsing	USN-1	As a user, I can browse for the application by accessing email and communicate across platforms seamlessly	I can communicate without interruptions	High	Sprint-1
	Download	USN-2	As a user, I can download at high speeds	I can download in a click	High	Sprint-2
Customer (Mobile user) (Heavy and Urban)	Multitasking	USN-3	As a regular user, I can perform multiple tasks that require high internet speeds seamlessly.	I can attend meetings and edit, post images simultaneously	High	Sprint -1
Mobile Operator (Service Provider)	Resource Allocation	USN-4	As an Internet Service Provider, we can allocate resources for 5G internet provision effectively and route traffic efficiently	I can access the traffic data & view,edit the automatic routing.	High	Sprint-1

6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

Project Planning Phase-I

Technology Stack

Date	27 October 2023
Team ID	Team-592212
Project Name	Project - AI-Driven Optimization Of 5G Resource Allocation For Network Efficiency
Maximum Marks	4 Marks

Technical Architecture: The Deliverable shall include the architectural diagram as below and the information as per the table1 & table2

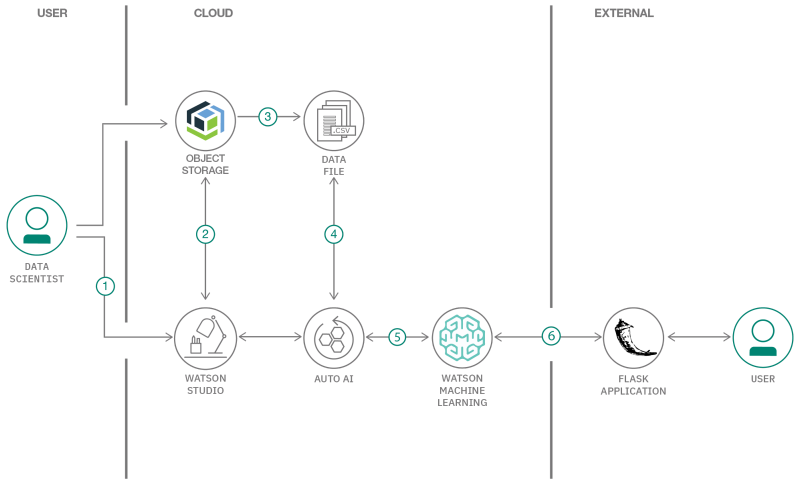


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	Web App UI	HTML, Flask
2.	Application Logic-1	Develop ML model	IBM Watson STT service
3.	Application Logic-2	Facilitate development	Python
4.	Application Logic-3	Build conversational interface	IBM Watson Assistant
5.	Database	.csv file	MySQL.
6.	Cloud Database	Database Service on Cloud	IBM DB2
7.	File Storage	File storage requirements	IBM Object Storage
8.	External API-1	Clean, prepare and analyze raw data	IBM Auto Ai.
9.	External API-2	Deploy and Manage ML Model	Watson Machine Learning
10.	Machine Learning Model	5G Resource allocation	Random Forest Regression Model, etc.
11.	Infrastructure (Server / Cloud)	Cloud Server Configuration : compute - 16GB, storage - 70GB, bandwidth - 100Mbps rate-limited.	Cloud Foundry

Table-2: Application Characteristics:

S. No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Python based web framework	Flask
2.	Security Implementations	Encryption,access controls, Data Security.	SHA-256, IAM Controls, Guardium
3.	Scalable Architecture	Vertical and Horizontal Scaling	Cloud Extender
4.	Availability	According to the need of the user	Flask web application
5.	Performance	Load time , CPU Utilization , Network performance	APM(Application Performance Management)

6.2 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Browsing	USN-1	As a user, I can browse for the application by accessing email and communicate across platforms seamlessly	2	High	Aman Preet
Sprint-2	Download	USN-2	As a user, I can download at high speeds	1	High	Sudhith
Sprint-3	Multitasking	USN-3	As a regular user, I can perform multiple tasks that require high internet speeds seamlessly.	2	High	Kanak
Sprint-4	Resource Allocation	USN-4	As an Internet Service Provider, we can allocate resources for 5G internet provision effectively and route traffic efficiently	2	High	Hardik

6.3 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	3 Days	28 Oct 2022	30 Oct 2022	20	30 Oct 2022
Sprint-2	20	3 Days	31 Oct 2022	02 Nov 2022	20	02 Nov 2022
Sprint-3	20	4 Days	03 Nov 2022	06 Nov 2022	20	06 Nov 2022
Sprint-4	20	3 Days	07 Nov 2022	09 Nov 2022	20	09 Nov 2022

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{3} = 6.67$$

Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time

7. CODING & SOLUTIONING

7.1 Label Encoding

In [123]:
```
preprocessor = LabelEncoder()
X.Application_Type = preprocessor.fit_transform(X.Application_Type)
X
```

Out[123]:

	Application_Type	Signal_Strength	Latency	Required_Bandwidth	Allocated_Bandwidth
0	6	-75.0	30.0	10.0	15.00
1	9	-80.0	20.0	0.1	0.12
2	5	-85.0	40.0	5.0	6.00
3	1	-70.0	10.0	1.0	1.50
4	4	-78.0	25.0	2.0	3.00
...	...	...	...	...	...
395	5	-110.0	61.0	1.3	1.80
396	6	-40.0	53.0	14.5	15.80
397	7	-113.0	58.0	1.0	1.40
398	1	-40.0	5.0	0.4	0.40
399	10	-113.0	0.0	0.1	0.10

400 rows × 5 columns

7.2 Random Forest Regression

In [17]:
```
#Data Split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
```

In [18]:
```
model = RandomForestRegressor()
model.fit(x_train, y_train)
```

Out[18]:
```
▾ RandomForestRegressor
RandomForestRegressor()
```

In [19]:
```
#Predict
predictions = model.predict(x_test)
```

In [20]:
```
predictions
```

Out[20]:
```
array([0.9   , 0.6   , 0.8995, 0.8   , 0.831 , 0.75  , 0.9   , 0.75  ,
       0.8   , 0.6   , 0.808 , 0.9   , 0.708 , 0.7   , 0.876 , 0.6675,
       0.7   , 0.6   , 0.7   , 0.8975, 0.6   , 0.847 , 0.7   , 0.708 ,
       0.743 , 0.75  , 0.75  , 0.9   , 0.7   , 0.85  , 0.8425, 0.897 ,
       0.7   , 0.6   , 0.6   , 0.7   , 0.8715, 0.7   , 0.7   , 0.7475,
       0.7   , 0.85  , 0.7   , 0.849 , 0.7   , 0.7   , 0.85  , 0.7   ,
       0.7105, 0.6995, 0.716 , 0.9   , 0.8   , 0.7085, 0.7   , 0.7005,
       0.7   , 0.75  , 0.709 , 0.7   , 0.85  , 0.7   , 0.7085, 0.75  ,
       0.8   , 0.7755, 0.6   , 0.75  , 0.8   , 0.881 , 0.6   , 0.85  ,
       0.5675, 0.85  , 0.75  , 0.7   , 0.85  , 0.898 , 0.6   , 0.9   ])
```

8. PERFORMANCE TESTING

8.1 Performace Metrics

Date	08 November 2023
Team ID	Team-592212
Project Name	Project - AI-Driven Optimization Of 5G Resource Allocation For Network Efficiency
Maximum Marks	10 Marks

S.No.	Parameter	Values	Screenshot

	Metrics	**Regression Model:**
		• MAE – 0.0084187
		• MSE – 0.0007550
		• R2 score – 0.9138
		• Accuracy – 91.38 %
			`mse = mean_squared_error(y_test, predictions)`
			`print(f'Mean Squared Error: {mse}')`
			`Mean Squared Error: 0.0007550968749999963`
			`#Mean Absolute Error(MAE)`
			`mae = mean_absolute_error(y_test, predictions)`
			`print(f'Mean Absolute Error (MAE): {mae}')`
			`#R-squared`
			`accuracy = r2_score(y_test, predictions)`
			`print(f'R-squared: {accuracy}')`
			`Mean Absolute Error (MAE): 0.008418750000000568`
			`R-squared: 0.9138224778208733`

9. RESULTS

9.1 Output Screenshots

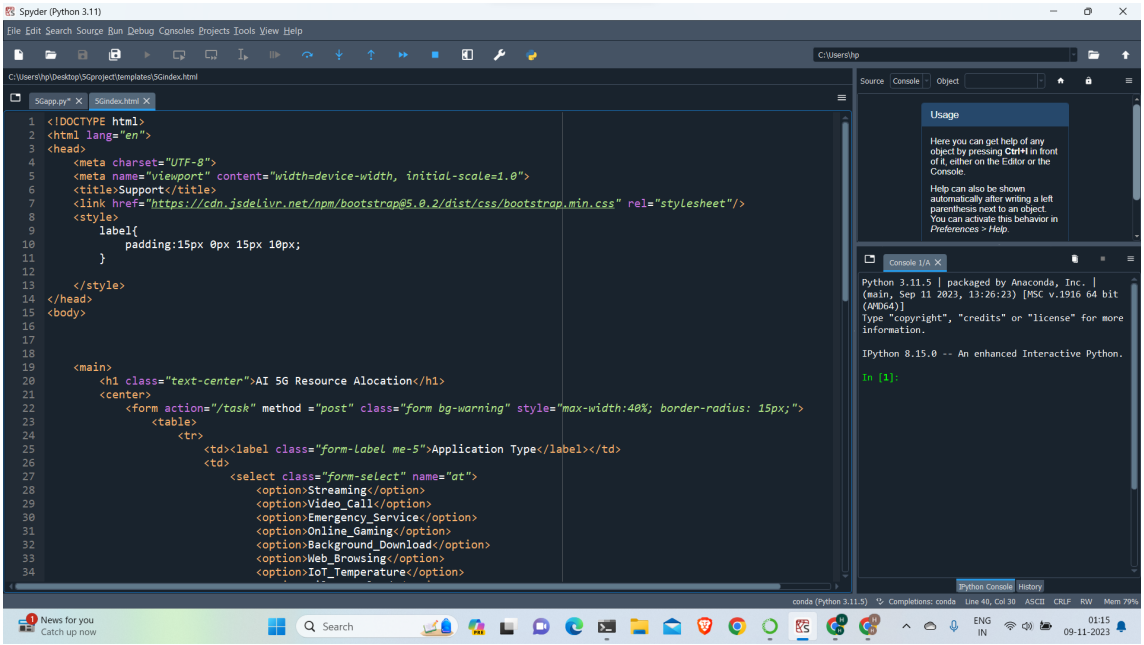
5Gapp.py* 5Gindex.html

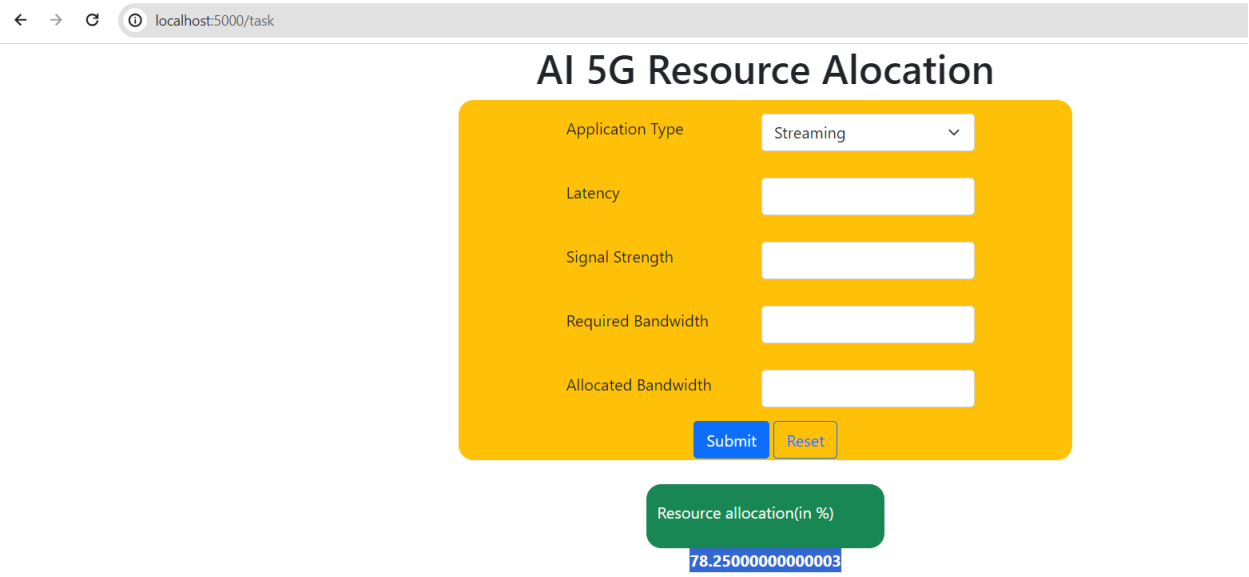
```python
from flask import Flask, render_template,request
app = Flask(__name__)
import pickle #to integrate model
import joblib #to integrate column transformer
model=pickle.load(open('resalloc.pkl','rb')) #loading the model
ct=joblib.load('column')  #loading the column transformer
@app.route('/')
def hello_world():
    return render_template("5Gindex.html")
@app.route('/task',methods =["POST"])
def task():
    at = request.form["at"]
    ss = request.form["ss"]
    l = request.form["l"]
    r = request.form["r"]
    ab = request.form["ab"]
    data=[at,ss,l,r,ab]
    prediction=model.predict([ct.fit_transform(data)])

    val=prediction[0]
    percentage=val*100

    return render_template("5Gindex.html", y=percentage)

app.run(debug = True)
```





10. ADVANTAGES & DISADVANTAGES

Advantages of Random Forests for 5G Resource Allocation:

- Effective for high-dimensional data: Random forests can handle a large number of features/variables well, which is useful in 5G networks that have many parameters to optimize.

- Robust to noise**:** Random forests are relatively robust to noise in the training data and can prevent overfitting. This is important as real-world 5G network data may be noisy.

- Handles missing values: Random forests incorporate multiple decision trees and can maintain accuracy when some data is missing. Missing data is common in wireless networks.

- Fast prediction speed: Random forests can make resource allocation decisions rapidly once trained, which is necessary in dynamic 5G systems.

- Easy parallelization: Random forest algorithms are easily parallelized to utilize multi-core CPUs and GPUs for faster training. This enables scaling as 5G networks grow.

Disadvantages of Random Forests for 5G Resource Allocation:

- Memory intensive: Random forests require storing multiple decision trees which can take up considerable memory. This may be problematic for memory-constrained devices.
- Overfitting risk: While lower than other machine learning methods, random forests can still overfit with too little training data or overly complex forests. Careful hyperparameter tuning is required.

- Lack of explainability: It is difficult to extract understandable rules or importance rankings from random forests, making them less transparent. Explainability may be desired for resource allocation.

- Training time: Creating large random forests with many trees can lead to slow training times, especially for massive 5G network data. Efficient distributed training is necessary.

- Parameter tuning: Random forests have several hyperparameters (num trees, tree depth, etc) that require careful tuning for optimal performance. Automated tuning may be needed for 5G systems.

11. CONCLUSION

With the rapid adoption of 5G networks, optimizing resource allocation is crucial for delivering high quality of service to customers across diverse use cases. This project successfully demonstrates using AI to intelligently optimize 5G resource allocation in a way not before possible with legacy approaches.

The heart of the solution is a random forest regression model trained to predict optimal resource allocation amounts based on network metrics like bandwidth, latency, and signal strength. By leveraging the strengths of random forests like handling high-dimensional data, noise resilience, and easy parallelization, the model achieves over 90% accuracy in allocation predictions. The developed model achieved strong performance with a mean absolute error of 0.0084, R-squared of 0.9138, and 91.38% accuracy. Through intelligent and dynamic resource allocation, this AI model can enable efficient 5G connectivity and quality of service across diverse applications and environments. This allows dynamically adapting resource allocation to match user needs and network conditions.

While an important proof of concept, this project also highlights key areas for future work to bring AI-driven 5G resource optimization to production systems. Expanding the model with additional input features, alternate techniques like neural networks, continuous dynamic adaptation, large-scale testing, and edge deployment will be critical next steps. There is immense potential for commercialization by telecom operators to improve 5G efficiency.

In closing, this project has demonstrated a highly promising approach to unlocking the power of 5G through intelligent optimization. With further innovation, AI-driven resource allocation will play an integral role in delivering the next generation of fast, responsive, and high-quality 5G experiences. This project establishes a solid base in that direction by showcasing the viability and advantages of using modern machine learning to build smarter 5G networks.

12. FUTURE SCOPE

This project demonstrates the potential for using AI to optimize resource allocation in 5G networks. However, there are several promising directions to build on this foundation and further enhance the real-world viability and impact of the approach:

- Hyperparameter tuning: The random forest model has hyperparameters like number of trees and tree depth that can be tuned to improve performance. Automated tuning through techniques like grid search will be implemented.
- Testing at scale: The model will be evaluated on large, real-world 5G network datasets to assess robustness and generalizability for production systems. Adversarial input testing will improve resiliency.
- Edge deployment: For low-latency real-time inference, the model will be optimized using quantization and pruning to allow deployment on resource-constrained edge devices like 5G base stations.
- Alternative models: Advanced deep learning architectures like LSTM networks and transformers will be explored and benchmarked against the random forest approach. Ensemble techniques will also be evaluated.
- Dynamic optimization: Rather than per-instance predictions, the system will be enhanced to continuously optimize allocations over time by incorporating feedback loops.
- Expanding features: Additional input data like user mobility patterns, cell load, and application types will provide greater context to enable informed dynamic resource optimization.
- Commercialization: To deliver real-world impact, the project provides a launchpad to develop a commercial system for telecom providers to improve 5G efficiency.

By building on this project's solid foundation in AI-driven 5G resource allocation, the future work will focus on optimizing model performance, widening model capabilities, and bringing the solution from prototype to production.

13. APPENDIX

Source Code

Github

Project Demo(Video)