Industrial Internship Report on

”Smart City Traffic Patterns Forecasting Project”

Prepared by

Shelin Vankawala

Executive Summary
This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT). This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks’ time. My project was Smart City Traffic Pattern Forecasting and I ensembles machine learning algorithms which exhibited varying degrees of effectiveness in forecasting traffic patterns. This internship gave me a very good opportunity to get exposure to Industrial problems and design/implement solution for that. It was an overall great experience to have this internship.

TABLE OF CONTENTS

1 Preface....3

2 Introduction....5

2.1 About UniConverge Technologies Pvt Ltd....5

2.2 About upskill Campus....9

2.3 Objective....10

3 Problem Statement....12

4 Existing and Proposed solution....13

5 Proposed Design/ Model....15

6 Performance Test....16

7 My learnings....17

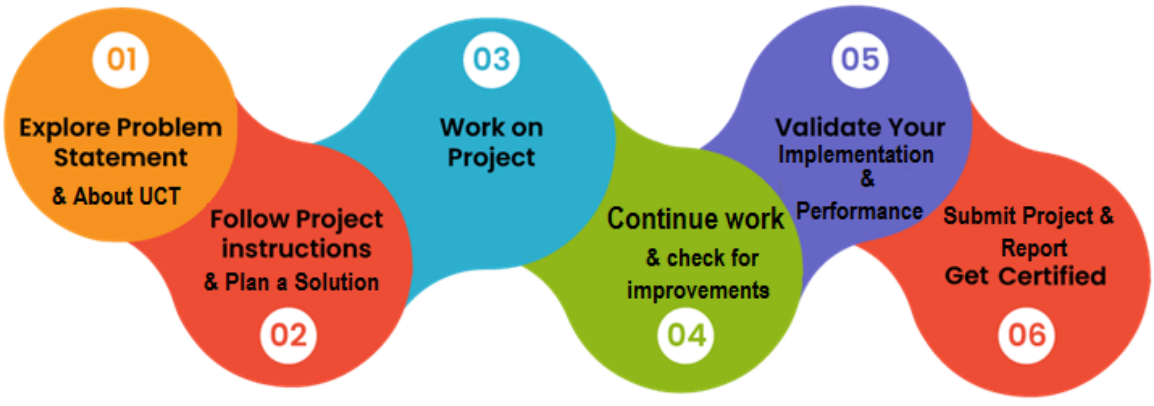
8 Future work scope....18

1 Preface

Over the span of six weeks, the internship has been a transformative journey, providing a blend of theoretical knowledge and hands-on experience crucial for career development. Engaging in a relevant internship has underscored the practical application of concepts learned in academic settings, emphasizing the importance of real-world problem-solving skills. The project, focusing on traffic pattern forecasting, served as a practical manifestation of machine learning principles, requiring a nuanced understanding of algorithms, data processing, and system integration.

The problem statement of the project was centered around developing an advanced traffic pattern forecasting system for smart city infrastructure. The challenge lay in surpassing conventional methods, addressing constraints such as memory and computational resources, and ensuring real-time adaptability. This real-world problem highlighted the complexity of urban traffic dynamics and provided insights into crafting efficient and scalable machine learning solutions.

The opportunity presented by the USC/UCT program has been instrumental in bridging the gap between academic learning and industry application. The structured nature of the program allowed for a holistic approach, with a carefully planned design flow that covered initialization, iteration, and implementation stages. Overall, the program has been a valuable stepping stone in the journey toward a career in data science and machine learning, offering practical insights that significantly contribute to professional growth.



The past six weeks have been an immensely enriching experience, offering a deep dive into the world of machine learning and data science. The practical application of concepts in the context of a real-world project has been invaluable, providing a tangible understanding of the challenges and intricacies involved. This internship has not only sharpened my technical skills but has also fostered a problem-solving mindset and a greater appreciation for the collaborative nature of the field.

I extend heartfelt gratitude to all those who have contributed directly or indirectly to my learning journey. Special thanks to UpSkills and UCT whose guidance and insights have been instrumental in navigating complex challenges and refining the project. The support from peers and colleagues has been equally commendable, creating an environment conducive to shared learning and collaboration. Each interaction, whether it be a discussion, feedback session, or shared resource, has contributed to my growth, and I am genuinely appreciative.

To my juniors and peers embarking on similar journeys, I offer the advice to embrace challenges with curiosity and persistence. The world of data science is dynamic and ever-evolving, and the key lies in continuous learning and adaptability. Seek out diverse perspectives, collaborate with enthusiasm, and approach problems with a solution-oriented mindset. The journey may be demanding, but the rewards are immeasurable, and every hurdle is an opportunity to learn and grow. Best of luck on your respective paths, and may you find as much fulfilment and excitement as I have in this enriching experience.

2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies e.g. Internet of Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LoRaWAN), Java Full Stack, Python, Front end** etc.

i. UCT IoT Platform (uct Insight)

UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

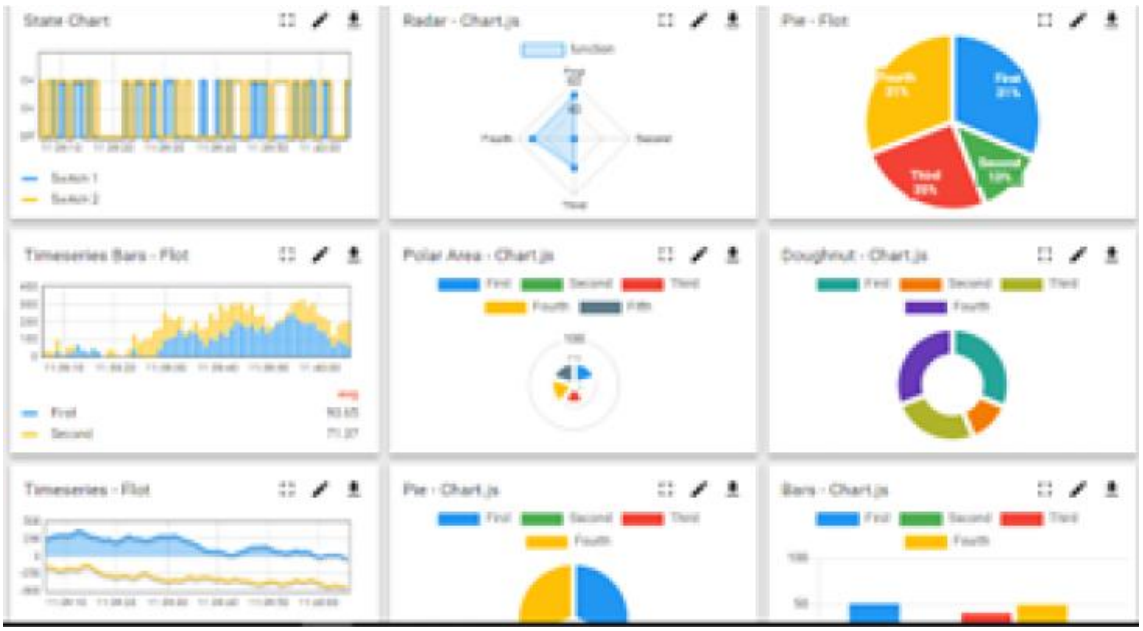
• Build Your own dashboard

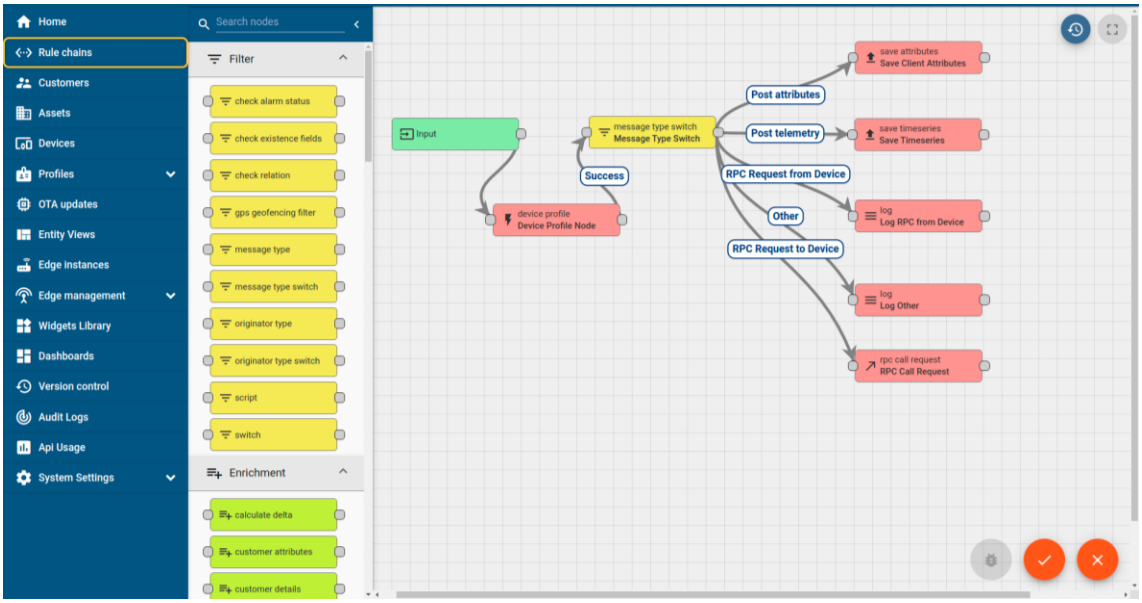
• Analytics and Reporting

• Alert and Notification

• Integration with third party application(Power BI, SAP, ERP)

• Rule Engine





ii. Smart Factory Platform (FACTORY WATCH)

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- with a scalable solution for their Production and asset monitoring
- OEE and predictive maintenance solution scaling up to digital twin for your assets.
- to unleased the true potential of the data that their machines are generating and helps to identify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they what to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.

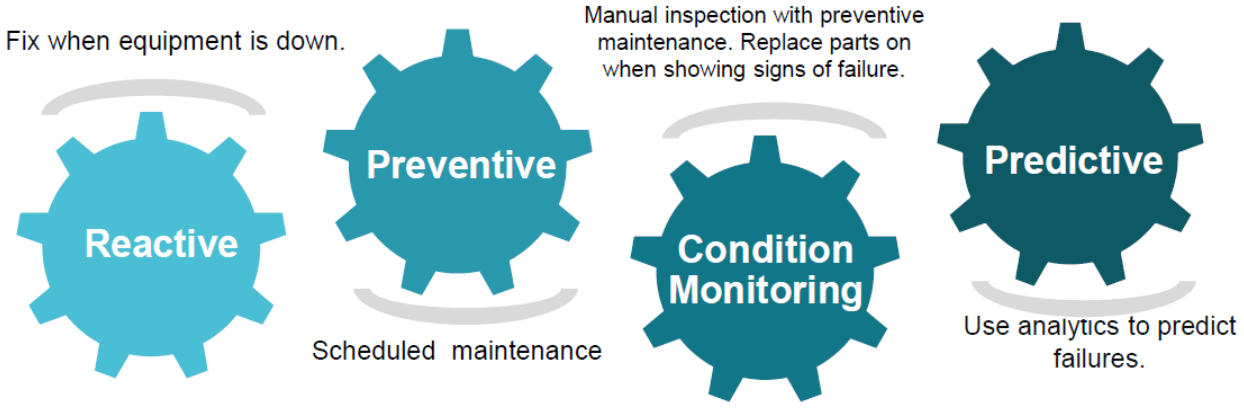
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i

iii. LoRaWAN based Solution

UCT is one of the early adopters of LoRAWAN teschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

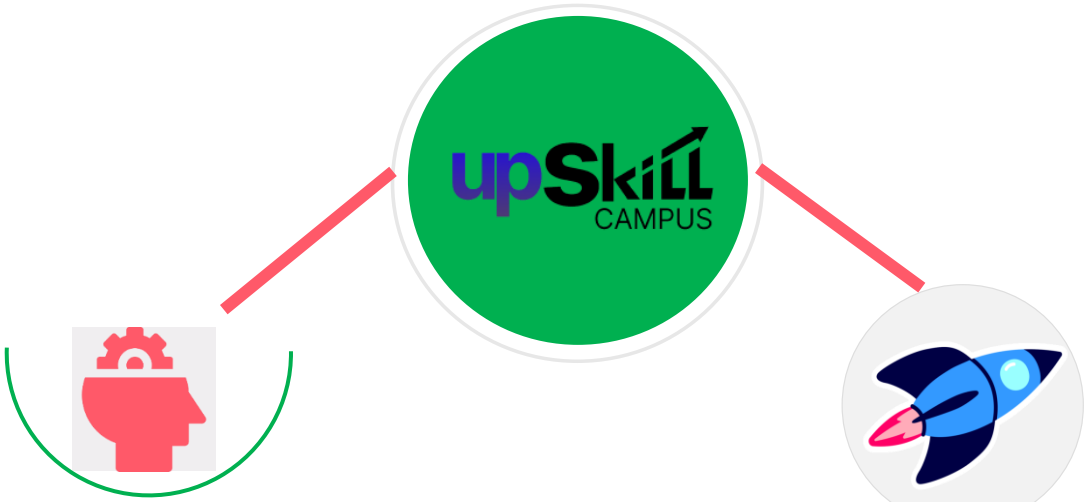
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

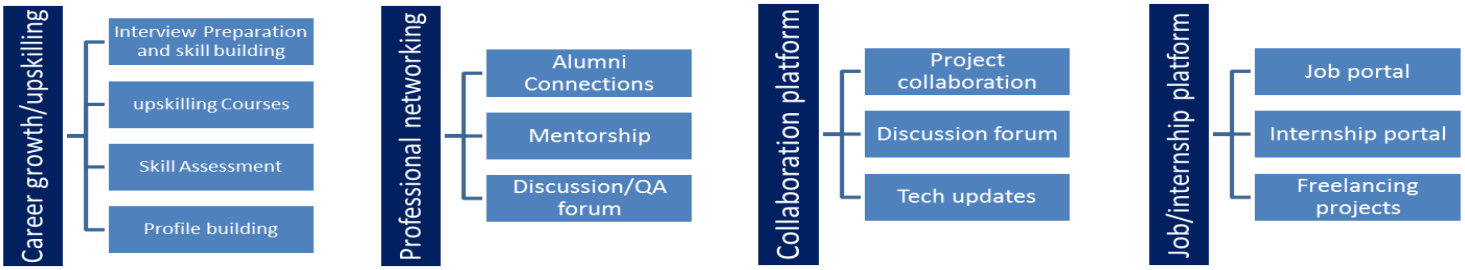
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

upSkill Campus aiming to upskill 1 million learners in next 5 year

https://www.upskillcampus.com/



2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

2.4 Objectives of this Internship program

The objective for this internship program was to

- get practical experience of working in the industry.
- to solve real world problems.
- to have improved job prospects.
- to have Improved understanding of our field and its applications.
- to have Personal growth like better communication and problem solving.

3 Problem Statement

Project Title: Smart City Traffic Pattern Forecasting

We are working with the government to transform various cities into a smart city. The vision is to convert it into a digital and intelligent city to improve the efficiency of services for the citizens. One of the problems faced by the government is traffic. You are a data scientist working to manage the traffic of the city better and to provide input on infrastructure planning for the future.

The government wants to implement a robust traffic system for the city by being prepared for traffic peaks. They want to understand the traffic patterns of the four junctions of the city. Traffic patterns on holidays, as well as on various other occasions during the year, differ from normal working days. This is important to take into account for your forecasting.

3.1 **Data set Link:**

https://drive.google.com/file/d/1y61cDyuO9Zrp1fSchWcAmCxk0B6SMx7X/view?usp=sharing

4 Existing and Proposed solution

Numerous existing solutions in the domain of traffic pattern forecasting rely on conventional methods, leveraging statistical models and historical data. While these approaches provide valuable insights, their limitations become apparent when dealing with the complexity and dynamism of urban traffic. Conventional models often struggle to adapt to the evolving nature of cities, unable to capture intricate patterns influenced by varying factors such as special events, urban development, and unexpected incidents. Additionally, their reliance on historical data may hinder their effectiveness in forecasting sudden changes or accommodating emerging trends, limiting their practical utility in a rapidly evolving smart city environment.

Our proposed solution revolves around the refinement of analysis techniques and the incorporation of advanced machine learning algorithms. By leveraging a meticulously curated dataset and implementing sophisticated algorithms like Random Forest, Decision Tree, LightGBM, among others, we aim to enhance the precision of traffic pattern forecasts. This approach seeks to address the limitations of traditional methods by offering a more dynamic and adaptable model. The focus is on capturing nuanced patterns, adapting to real-time changes, and providing a comprehensive understanding of traffic dynamics that goes beyond the capabilities of existing solutions.

The proposed solution brings substantial value addition in multiple dimensions. Firstly, by refining analysis techniques and employing advanced algorithms, we aim to provide a forecasting system that is not only accurate but also adaptive to the ever-changing urban landscape. This adaptability ensures a more proactive response to emerging traffic trends and city developments. Secondly, the incorporation of machine learning algorithms allows for a more granular understanding of complex traffic patterns, offering insights that conventional models might overlook. Ultimately, our solution strives to elevate the efficiency of urban traffic management, enhancing citizen well-being and the overall quality of life in our smart city.

4.1 Code submission (Github link):
https://github.com/ShelinVankawala/Smart_City_Traffic_Pattern_Forecasting

4.2 Report submission (Github link) : first make placeholder, copy the link.

5 Proposed Design/ Model

The proposed design for the traffic pattern forecasting system incorporates a diverse set of Machine Learning algorithms, leveraging the strengths of each to enhance predictive accuracy and adaptability. The ensemble of algorithms includes XGBoost, Linear Regression, Support Vector Regression (SVR), Gradient Boosting, LightGBM, Neural Network Regressor, Decision Tree, and Random Forest. Each algorithm contributes unique capabilities to address different facets of the traffic prediction challenge.

XGBoost and Gradient Boosting are powerful ensemble methods that sequentially build a series of weak learners, learning from errors to improve accuracy. These algorithms excel in capturing complex relationships within the data and are well-suited for scenarios with non-linear dependencies. Linear Regression provides a baseline model, offering simplicity and interpretability in capturing linear trends.

Support Vector Regression (SVR) is chosen for its effectiveness in handling high-dimensional data and capturing non-linear patterns. LightGBM, known for its efficiency and speed, is particularly suitable for large datasets, contributing to the scalability of the solution.

The inclusion of a Neural Network Regressor brings deep learning capabilities into play, enabling the model to capture intricate patterns and relationships within the data that may be challenging for traditional algorithms. Decision Tree and Random Forest models, known for their interpretability and robustness, further contribute to the ensemble, providing a diversified set of models to handle various traffic scenarios.

This ensemble approach ensures a comprehensive understanding of traffic patterns, allowing the system to adapt to the nuances of urban dynamics. During the iterative phase of the design flow, the performance of each algorithm is rigorously assessed, and hyperparameter tuning is applied to optimize their individual contributions. The final ensemble is expected to provide a robust and adaptable forecasting system capable of addressing the complexity of urban traffic management.

6 Performance Test

In the performance testing phase, the ensemble of machine learning algorithms exhibited varying degrees of effectiveness in forecasting traffic patterns. Notably, Random Forest emerged as the top-performing algorithm, demonstrating the highest predictive accuracy and reliability. Its ability to handle complex relationships within the data and maintain robustness against overfitting contributed to its outstanding performance.

Following closely behind was Decision Tree, showcasing reliable performance and interpretability. Decision Tree's simplicity and ability to capture non-linear patterns were evident in its consistent forecasting capabilities.

XGBoost, known for its gradient boosting technique and ability to handle diverse data, also delivered commendable performance. Its sequential learning process and adaptability to nuanced patterns within the data contributed to accurate traffic predictions.

The other algorithms, including Linear Regression, SVR, Gradient Boosting, LightGBM, and Neural Network Regressor, demonstrated satisfactory performance, albeit with varying degrees of accuracy. Each algorithm contributed unique insights into traffic patterns, with their strengths aligning with specific data characteristics.

Overall, the ensemble approach allowed for a comprehensive evaluation of each algorithm's performance, highlighting the versatility and adaptability of Random Forest, the interpretability of Decision Tree, and the gradient boosting capabilities of XGBoost. These results will inform the final selection of algorithms for the traffic pattern forecasting system, ensuring a well-rounded and effective solution.

7 My learnings

This machine learning project has been an invaluable learning experience, offering insights that transcend theoretical knowledge and extend into practical applications. From data preparation and model selection to addressing real-world constraints and designing a comprehensive testing strategy, the project has provided a holistic view of the end-to-end machine learning process. This hands-on experience has not only deepened my understanding of algorithms and methodologies but has also honed my problem-solving skills in the context of real-world challenges.

One key takeaway is the significance of iterative development and continuous refinement. Additionally, tackling constraints, such as memory and computational limitations, has underscored the need for resource-efficient solutions and a keen understanding of the practical implications of algorithmic choices.

The project has equipped me with practical skills in data analysis, feature engineering, and model implementation, enhancing my proficiency in machine learning. The experience of addressing challenges in a real-world scenario has fostered a problem-solving mindset and the ability to navigate complexities in a structured manner.

As I reflect on this project, I am confident that the skills and insights gained will significantly contribute to my career growth. The practical knowledge acquired, coupled with the ability to translate theoretical concepts into tangible solutions, positions me well for tackling complex challenges in the field of machine learning. The iterative development approach and problem-solving skills honed during this project will undoubtedly be assets in future endeavors, fostering continuous learning and innovation in my professional journey.

8 Future work scope

While the machine learning project provided a comprehensive exploration of traffic pattern forecasting, there were certain ideas and avenues that, due to time constraints, couldn't be fully realized. One such area is the integration of external data sources, such as weather patterns, public events, or social media trends, to enhance the predictive capabilities of the model. The inclusion of these dynamic factors could provide a more holistic understanding of the urban environment's influence on traffic.

Additionally, the project focused primarily on supervised learning techniques, and an exploration into unsupervised learning for anomaly detection or clustering could further enrich the analysis. The application of reinforcement learning algorithms to optimize traffic management strategies in real-time is another promising avenue for future exploration.

In the future, expanding the scope to consider multiple cities or regions could provide a broader perspective on the scalability and generalizability of the developed solution. Collaborations with urban planners and policymakers to integrate the forecasting system seamlessly into city management practices would be a pivotal next step.

Overall, these ideas represent potential directions for future work, building upon the foundation laid during this project. The time limitations serve as opportunities for ongoing exploration and innovation, ensuring that the project's impact can continue to grow and evolve in addressing the dynamic challenges of urban traffic management.