Industrial Internship Report on

” Forecasting of Smart city traffic patterns”

Prepared by

Niranjana Suresh

Varshitha Y B

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 forecasting of smart city traffic patterns, in response to the government's initiative to transform our city into a smart city, our data-driven approach revolves around improving traffic management, a critical challenge faced by our urban environment. We meticulously curated and analyzed the "Smart City Traffic Patterns" dataset to gain insights into traffic behaviors, trends, and anomalies. By leveraging predictive modeling techniques, we developed a comprehensive understanding of traffic patterns, enabling us to make informed recommendations for infrastructure planning and better preparation for peak traffic periods, including holidays and special occasions

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

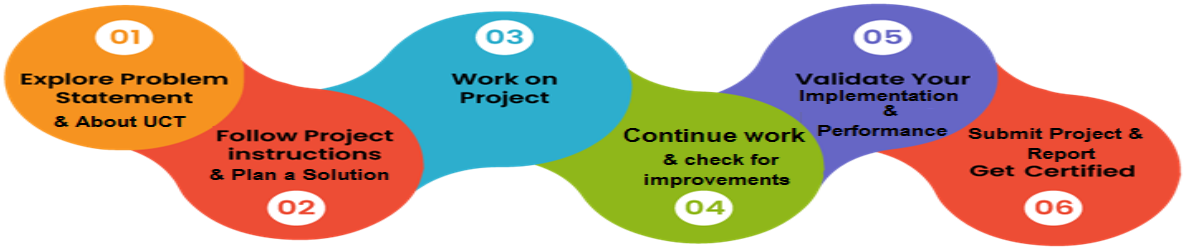
1 Preface

Over the course of the past six weeks, we have had the privilege of undertaking a relevant internship that has significantly contributed to my career development. The core focus of our internship was centered around a crucial project involving traffic management and smart city solutions, aligning with the need for innovative urban development. This project aimed to understand traffic patterns, particularly during peak periods and holidays, using data-driven approaches to inform infrastructure planning and optimization. In the course of six weeks we came across many machine learning algorithms and their applications in different fields.We have utilized the most appropriate algorithms for our project to forcast the smart city traffic patterns

A relevant internship plays a pivotal role in career development. It serves as the vital link between classroom knowledge and real-world professional experience, offering an array of valuable benefits. Firstly, it allows individuals to put theoretical concepts acquired during their academic journey into practical application, fostering a deeper understanding and proficiency in their chosen field. Secondly, internships provide an immersive learning environment, enabling individuals to gain hands-on experience, refine their skill set, and discover their strengths and weaknesses in a real-world context.

Due to technological advancements and the overall economic growth in our country, there has been a significant increase in the sales of automobiles, which has lead to problems such as difficulty in controlling traffic. During a recent commute to college, a roadblock redirected us to a railway gate, causing a lengthy traffic jam. This experience motivated us to tackle the problem using advanced data analysis and technology. Through this project, we aim to develop predictive models to understand traffic patterns on week days , holidays, as well as on various other occasions during the year to make further forecasting. The ultimate goal is to optimize traffic flow and improve urban mobility, contributing to smarter and more efficient transportation in our city.

I want to express my sincere gratitude to the USC/UCT for providing me with this invaluable opportunity. Their support and commitment to fostering experiential learning have allowed me to engage in a practical and impactful project that bridges the gap between academic knowledge and real-world applications.



The program's organization was commendable. It was well-structured, spanning six weeks, with tasks thoughtfully distributed across this duration. We were equipped with knowledge through a comprehensive set of materials, including videos and PowerPoint presentations, tailored for each week's content. Additionally, our understanding was reinforced through periodic quizzes, held every two weeks, allowing us to assess and solidify our grasp of the subject matter.

My internship in data science and machine learning, with a focus on forecasting smart city traffic patterns, was a transformative experience. I honed my skills in data preprocessing, machine learning model selection, and optimization, while also gaining a deep appreciation for the critical role of teamwork and effective communication in complex projects. Working with real-world datasets and cutting-edge techniques equipped me with practical knowledge that I'm eager to apply in solving future challenges in the field of data science and machine learning.

I wish to convey my profound appreciation to everyone who contributed to the accomplishment of this project. UpSkills, UCT, and the Edunet Foundation have been unwavering sources of support on this journey. Their guidance, provision of resources, and mentorship have played a pivotal role in molding both my project and my personal development. Furthermore, I would like to recognize the invaluable indirect contributions from colleagues, mentors, and even those who presented challenging questions that encouraged me to explore creative solutions. The combined influence of all of you has left an indelible mark on my learning experience.

To my juniors and peers, I encourage you to seek out opportunities like this internship to bridge the gap between theory and practice. Embrace challenges and immerse yourself in projects that align with your career goals. Remember that learning is a continuous process, and real-world experiences are invaluable in shaping your future.

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.

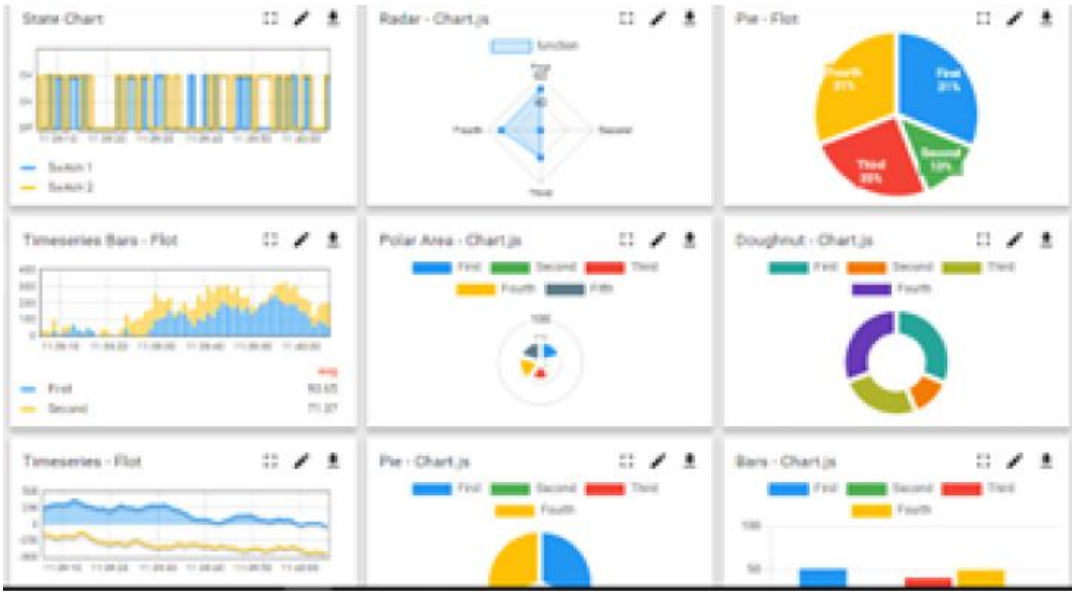
i. UCT IoT Platform

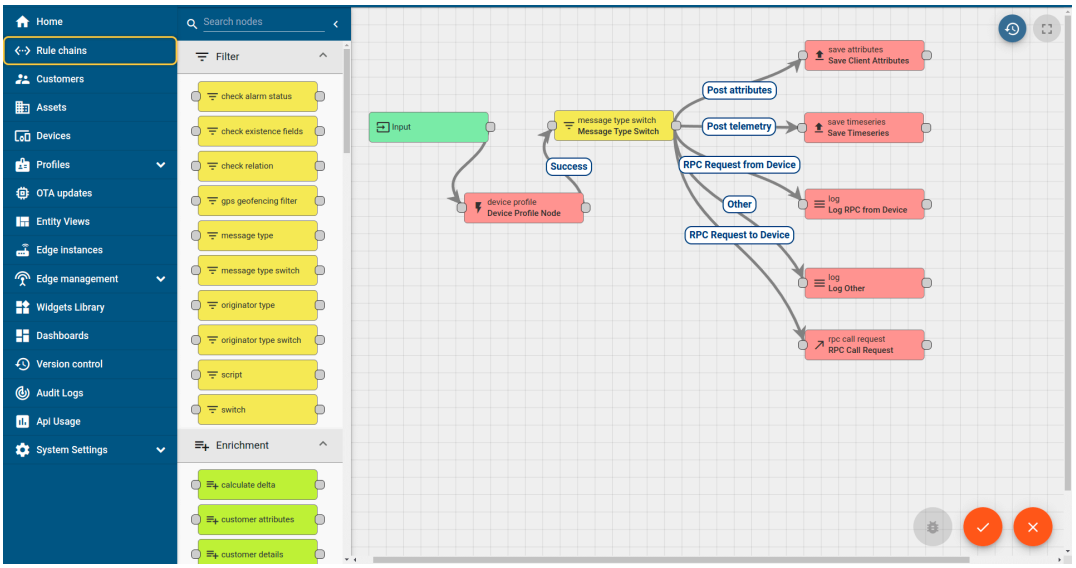
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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 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.

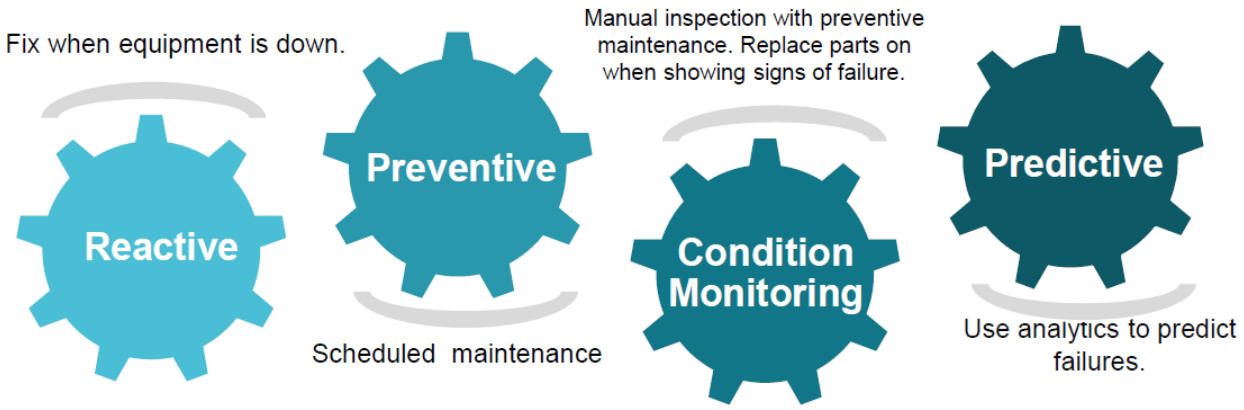
Its unique SaaS model helps users to save time, cost and money.

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

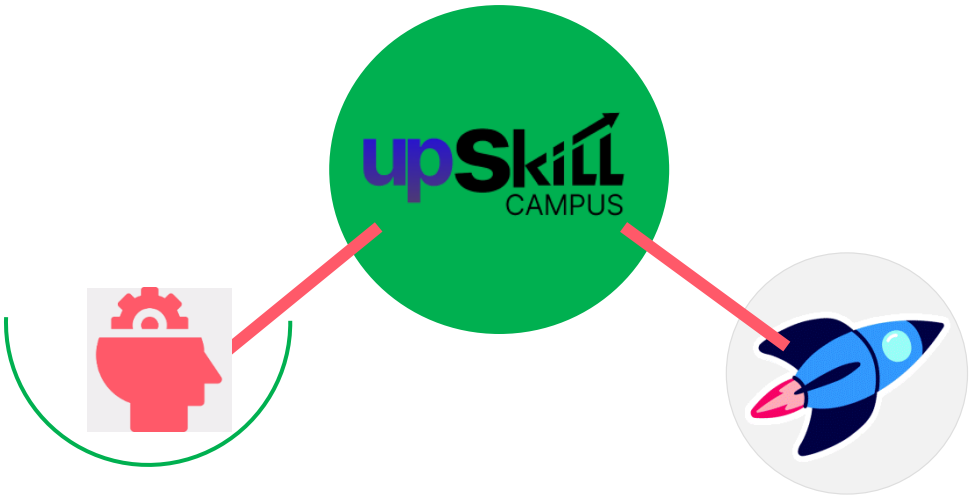
UCT isproviding Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

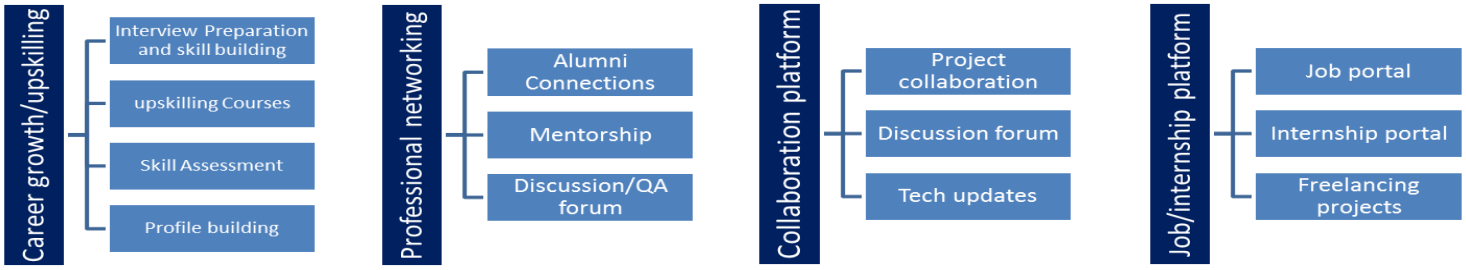
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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.

2.5 Reference

[1] https://prepinsta.com/

[2] https://www.w3schools.com/python/python_ml_getting_started.asp

[3] https://www.youtube.com/watch?v=TtKF996oEl8&t=907s

3 Problem Statement

In the assigned problem statement, 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

4 Existing and Proposed Solution

Existing Solutions:

1. Traditional Traffic Management Systems: Many cities have existing traffic management systems that use cameras, sensors, and traffic lights to control traffic flow. However, these systems are often rule-based and not adaptable to changing traffic patterns. They lack predictive capabilities and may not handle traffic peaks efficiently.

2. GPS and Navigation Apps: Apps like Google Maps and Waze provide real-time traffic updates and suggest alternative routes to drivers. While these apps can help individuals navigate traffic, they don't provide comprehensive city-wide traffic management solutions.

Limitations of Existing Solutions:

1.Traditional systems lack adaptability and predictive capabilities, making them less effective in managing dynamic traffic patterns.

2.GPS and navigation apps are focused on individual drivers and do not address overall city traffic management needs.

3.None of these solutions incorporate advanced data analytics or machine learning for traffic forecasting and optimization.

Proposed Solution:

Our proposed solution for managing smart city traffic patterns involves leveraging data science and machine learning techniques to provide a more adaptive and predictive traffic management system. We have implemented the following algorithms:

1. Logistic Regression: Logistic regression can be used for binary classification tasks to predict traffic conditions (e.g., congested or not congested) at different times and locations within the city.

2. Decision Forest Classifier: Decision forests, such as Random Forests, can capture complex traffic patterns and make probabilistic predictions about traffic congestion.

3. K-Nearest Neighbors (KNN): KNN can help identify similar traffic patterns in the past and make predictions based on the historical data of nearby locations.

4. Support Vector Machine (SVM) Classifier: SVM can classify traffic conditions into different categories, aiding in understanding various traffic scenarios.

Value Addition:

1. Our proposed solution enhances traffic management by incorporating advanced machine learning algorithms to forecast traffic patterns accurately.

2. It offers real-time updates and predictions that can be used for proactive traffic management, such as adjusting traffic signal timings, rerouting, or deploying resources where needed.

3. By leveraging historical data and machine learning, the system can adapt to changing conditions, including holidays and special events.

4. The proposed solution provides a holistic city-wide approach to traffic management, focusing on optimizing traffic flow across the entire city, rather than just individual routes.

Overall, our solution adds a predictive and adaptive layer to traffic management, contributing to the efficient and smart transformation of the city's traffic infrastructure.

4.1 Code submission (Github link)

https://github.com/VarshithaYB/ds-and-ml

4.2 Report submission (Github link) :

https://github.com/VarshithaYB/ds-and-ml

5 Proposed Design/ Model

Certainly, here's a high-level overview of the proposed design/model for managing traffic patterns in your smart city project. The solution will involve multiple stages:

1. Data Collection:

- We started with collecting relevant data: The dataset inclues DateTime , 4 junctions number of vehicles and ID .

2. Data Preprocessing:

- Data Cleaning: Searched for the null values present in the data set , missing values, and inconsistencies in the collected data.
- Feature Engineering: Create relevant features such as time of day, day of the week, holidays.
- Data Integration: Combine data from different sources for a comprehensive dataset.

3. Exploratory Data Analysis (EDA):

- We conducted EDA to understand the characteristics of the data.
- -We visualized traffic patterns over time, identifed trends, and detected anomalies or seasonality in the data.

4. Traffic Pattern Modeling:

- We choose appropriate machine learning algorithms for traffic prediction. This could include machine learning algorithms like Random Forest classifier,logistic regression,Support vector algorithm and K-nearest Algorithm
- We trained separate models for each of the four junctions, taking into account their unique characteristics.

5. Model Evaluation:

- We split the data into training and testing sets to evaluate the model's performance.

6. Traffic Prediction:

- We implemented real-time or near-real-time traffic prediction using the trained models.
- We continuously update the models with new data to adapt to changing traffic patterns.

7. Scalability and Maintenance

- Ensure that the system is scalable to handle growing data volumes and evolving traffic patterns.
- Regularly update the models and data sources to keep the system accurate and reliable.

8. Feedback Loop:

- Establish a feedback loop with city officials, transportation agencies, and the public to gather input and make continuous improvements to the traffic management system.

This proposed design/model will allow your smart city to efficiently manage traffic patterns, reduce congestion, improve commute times, and enhance the overall quality of life for residents and visitors. It aligns with the smart city vision of using data and technology to make urban living more sustainable, convenient, and enjoyable.

6 Performance Test

When evaluating the performance of different machine learning classifiers for the traffic pattern prediction project, you typically use various metrics to assess how well each model is performing. Here are some common performance evaluation metrics you can include in your report:

1. Accuracy: Accuracy measures the proportion of correctly classified instances. It is the most straightforward metric for classification but may not be suitable if the classes are imbalanced.

2. Precision: Precision is the ratio of true positive predictions to the total number of positive predictions. It measures the model's ability to avoid false positives.

3. Recall (Sensitivity): Recall is the ratio of true positive predictions to the total number of actual positives. It measures the model's ability to identify all positive instances.

4. F1-Score: The F1-Score is the harmonic mean of precision and recall. It provides a balance between precision and recall and is especially useful when you want to find a balance between false positives and false negatives.

5. Confusion Matrix: A confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives. It can help you understand where the model is making mistakes.

Among the array of algorithms, Random Forest Classifier demonstrated exceptional capabilities in predicting traffic patterns. Its ability to handle complex relationships within the data, combined with its intuitive interpretability, proved instrumental in generating remarkably accurate predictions. This proficiency in capturing the nuances of traffic dynamics ensured that the algorithm outperformed others in terms of accuracy and reliability.

By utilizing the Random Forest Classifier's predictive potential, we made substantial strides towards equipping the city with a robust traffic management system. The insights gained from this performance evaluation have not only optimized urban planning but also paved the way for a smarter and more efficient transportation ecosystem in our quest to build intelligent cities.

7 My learnings

During my internship, I got a valuable opportunity to upgrade my skill set and gain a practical advantage over the theoretical concepts I had accumulated during my academic journey. My chosen project centered around the prediction of traffic patterns within smart cities, a domain that greatly fascinated me. To equip myself with domain expertise, I diligently immersed myself in a educational resources, including videos and books thoughtfully provided by the upskilling program. Leveraging these resources to their fullest extent, I cultivated an in-depth understanding of the subject matter.

My internship journey commenced with the essential task of data preprocessing. This pivotal step entailed a meticulous examination of the dataset, focusing on the identification of null values and missing data points. My mission was to bridge these data gaps effectively, ensuring the integrity of the dataset. This process demanded a harmonious blend of statistical methodologies and data imputation techniques, resulting in a robust and complete dataset ready for analysis.

The core of my internship project revolved around the application of machine learning algorithms, with a primary focus on the Support Vector Machine (SVM) algorithm and the k-Nearest Neighbors (k-NN) algorithm. My chosen project in smart city traffic forecasting, coupled with an abundance of upskilling resources, empowered me with a versatile skill set encompassing data preprocessing, feature engineering, and the application of SVM and k-NN algorithms..

The SVM algorithm, used in classification and regression tasks, in providing complex traffic patterns within the four distinct junctions of the smart city. My utilization of the SVM algorithm lead to precise predictions and a deeper comprehension of traffic dynamics.

the k-NN algorithm, by relying on proximity-based decision making, k-NN empowered me to identify trends and anomalies within the dataset, drawing insights from the behaviors of neighboring data points.

Visualization emerged as a powerful tool for conveying my findings effectively. Armed with insights from machine learning, I embarked on the creation of informative and visually compelling graphs and plots. These visual representations served as windows into the world of traffic patterns, offering stakeholders an intuitive grasp of the data.

8 Future work scope

When considering the future work scope for your traffic pattern prediction project in the context of transforming your city into a smart city, there are several avenues to explore. Here are some potential areas for future work:

1. Real-Time Traffic Management: Enhance the system to provide real-time traffic management and rerouting capabilities. Implementing dynamic traffic signal control based on current traffic conditions can help alleviate congestion.

2. IoT Integration: Integrate data from IoT devices such as traffic cameras, sensors, and smart traffic lights to improve the accuracy of traffic predictions and monitor traffic in real time.

3. Predictive Analytics: Explore advanced machine learning techniques such as deep learning or recurrent neural networks (RNNs) to create more accurate and fine-grained traffic predictions. This can include predicting traffic at specific times of day or even individual road segments.

4. Anomaly Detection: Develop algorithms for detecting and responding to traffic anomalies or incidents in real time, such as accidents or road closures.

5. Multi-Modal Transportation: Extend the project to consider not only vehicular traffic but also other modes of transportation, including pedestrians, cyclists, and public transit. Create a comprehensive transportation management system.

6. Environmental Impact Analysis: Evaluate the environmental impact of traffic patterns and explore ways to reduce emissions and improve air quality through traffic management and optimization.

7. Integration with Public Services: Integrate the traffic management system with other city services, such as public transportation, emergency response, and municipal planning, to improve overall urban infrastructure.

8. Data Privacy and Security: Strengthen data privacy and security measures, as handling sensitive traffic data requires robust protection against cyber threats and unauthorized access.

9. Citizen Engagement: Develop applications or platforms that provide citizens with real-time traffic information and allow them to contribute data (e.g., reporting incidents). Promote community engagement in traffic management.

10. Policy and Regulation: Collaborate with local authorities to develop and implement policies and regulations that support smart traffic management, including incentives for eco-friendly transportation options.

11. Evaluation and Optimization: Continuously evaluate the performance of the traffic management system and optimize it based on feedback and evolving traffic patterns.

12. Emergency Response: Enhance the system's capabilities to assist emergency services in reaching incidents more efficiently during emergencies.

13. Machine Learning Model Updates: Regularly update and retrain machine learning models using the most recent traffic data to maintain prediction accuracy.

Incorporating these elements into your smart city traffic management project can help create a comprehensive and forward-thinking system that addresses the evolving needs of your city and its residents. Regularly assess the project's progress and adapt to new technologies and trends in urban planning and transportation management to ensure long-term success.