Industrial Internship Report on

”Forecasting of Smart city traffic patterns”

Prepared by

SUJALSINH GOHEL

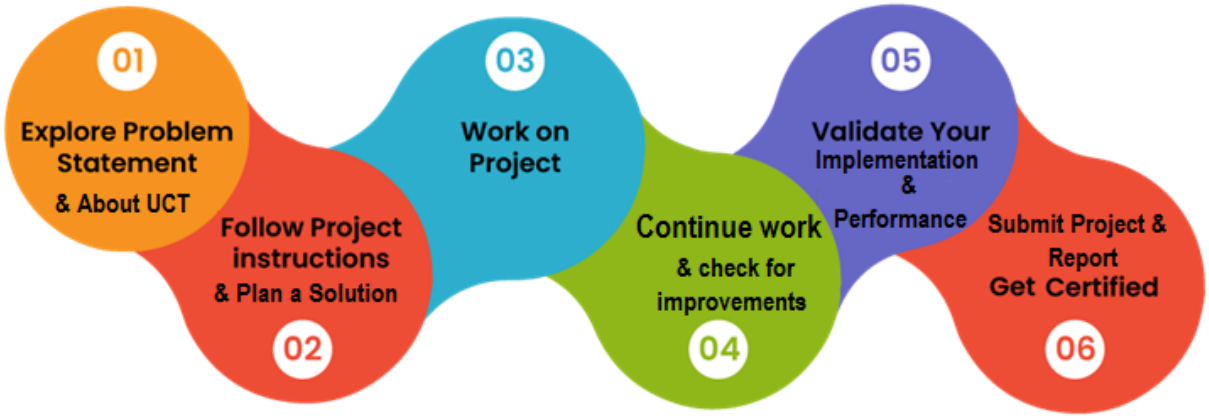
Executive Summary
Project involves developing advanced predictive models to analyze and anticipate traffic flows within urban environments. Utilizing a combination of historical traffic data, real-time sensor inputs, and machine learning algorithms, the project aims to enhance traffic management and reduce congestion. By integrating these models with smart city infrastructure, the goal is to provide city planners and traffic authorities with actionable insights to optimize traffic control measures, improve public transportation efficiency, and ultimately enhance the overall urban mobility experience. This project not only addresses current traffic challenges but also lays the groundwork for future innovations in smart city development.

TABLE OF CONTENTS

1 Preface

Over the past six weeks, my internship at UniConverge Technologies Pvt Ltd has been an enriching experience, primarily focused on the development of advanced predictive models to forecast smart city traffic patterns. The work involved a comprehensive analysis of historical traffic data, real-time sensor inputs, and the application of machine learning algorithms to predict traffic flows. The end goal was to create a system that could assist in optimizing traffic management and reducing congestion in urban environments. Each week brought new challenges and learning opportunities, from data preprocessing and model selection to system integration and performance testing.

Internships play a critical role in career development, providing practical experience and exposure to real-world problems. This particular internship allowed me to apply theoretical knowledge gained from academic courses to tangible issues in traffic management and smart city planning. It helped me develop technical skills in data analysis and machine learning, as well as soft skills such as problem-solving, project management, and effective communication.



The core objective of the project was to develop predictive models for forecasting traffic patterns in a smart city environment. The problem statement involved analyzing various factors affecting traffic flow, including historical traffic data and real-time sensor inputs. The challenge was to create models that could accurately predict traffic congestion and provide actionable insights for traffic authorities and city planners.

UniConverge Technologies Pvt Ltd provided me with a fantastic opportunity to work on a cutting-edge project that addresses a significant urban issue. The company offered a supportive environment with access to industry experts and advanced technological resources. This experience allowed me to gain hands-on experience in the field of smart city solutions and contribute to innovative developments in urban traffic management.

The internship program was meticulously planned to ensure a structured learning experience. The first week involved onboarding and orientation, where I was introduced to the company, its values, and the project objectives. Subsequent weeks were divided into specific tasks, including data collection, data preprocessing, exploratory data analysis, model selection, and development, followed by testing and integration. Regular meetings with my supervisor and team ensured that I received continuous feedback and guidance, allowing for iterative improvements to the project. The final week was dedicated to performance testing, documentation, and presentation of the project outcomes.

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.

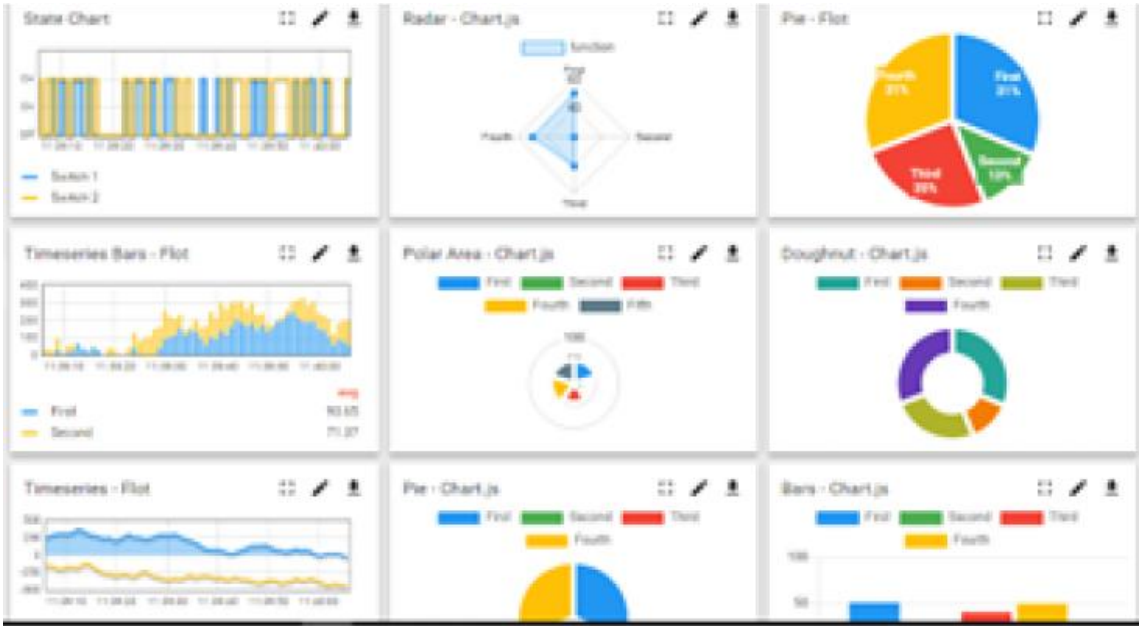
i. UCT IoT Platform (UCT Insight)

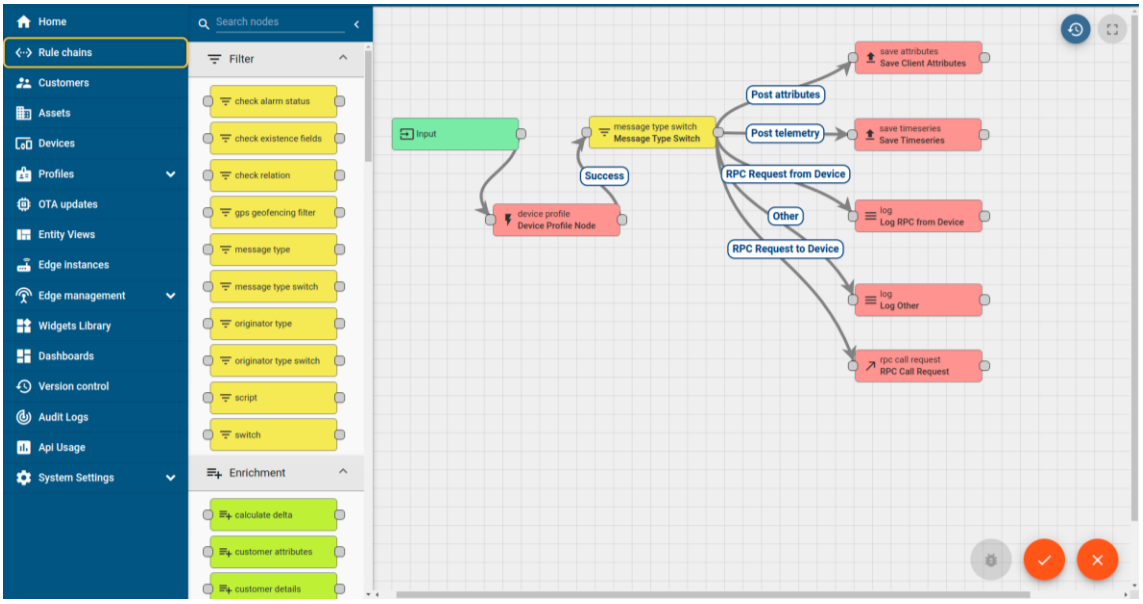
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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.

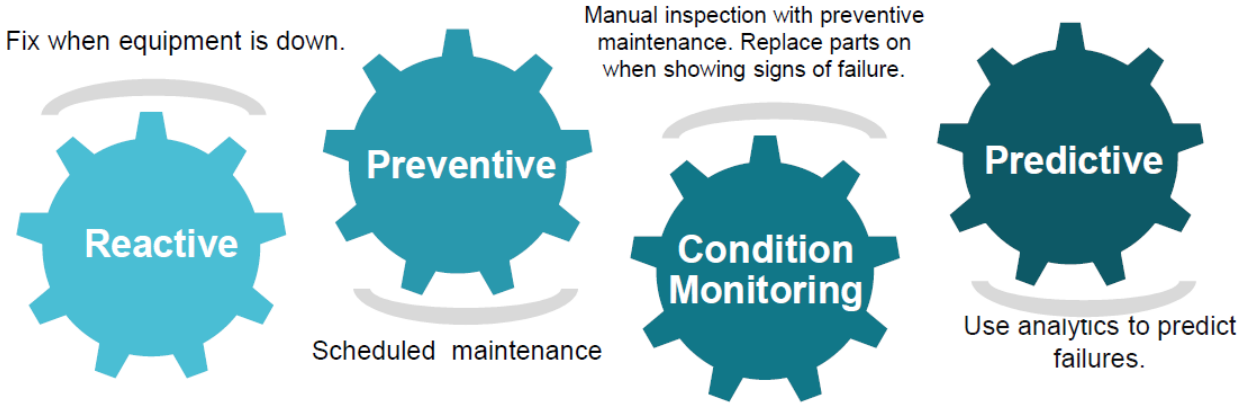
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

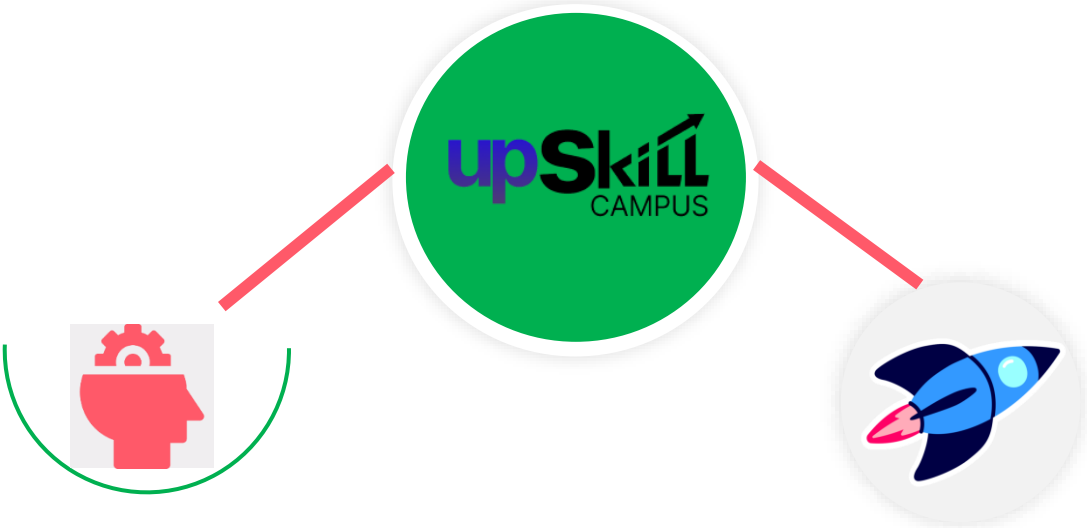
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

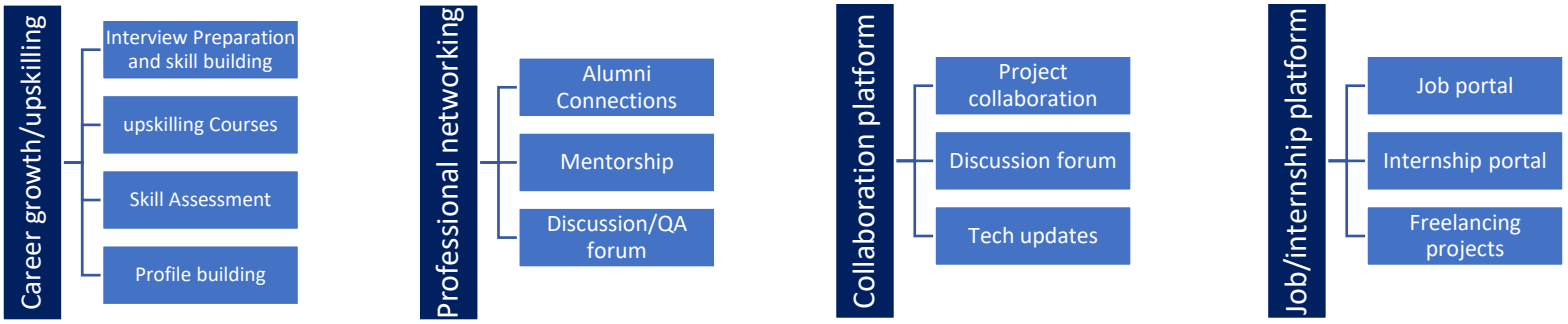
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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

UCT 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. As 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

1. Traditional Traffic Management Systems:

Description: These systems use traffic signals, road sensors, and cameras to monitor and manage traffic flow. They rely on predefined rules and manual adjustments by traffic authorities.

Limitations: Traditional systems often struggle with real-time responsiveness and adaptability. They may not efficiently handle unexpected traffic surges or incidents, leading to delays and congestion.

2. GPS and Navigation Systems:

Description: Systems like Google Maps and Waze use real-time GPS data from users to provide traffic updates and suggest optimal routes.

Limitations: While effective for individual navigation, these systems do not offer comprehensive traffic management solutions for city-wide optimization. They can also lead to unintended consequences like increased traffic on alternative routes.

3. Basic Predictive Models:

Description: Some cities use basic statistical models and historical data to forecast traffic patterns and plan accordingly.

Limitations: These models often lack accuracy and fail to account for real-time changes and complex traffic dynamics. They may not integrate well with modern smart city infrastructure.

Proposed Solution

Our proposed solution aims to enhance traffic forecasting and management through advanced predictive modeling and seamless integration with smart city infrastructure. The solution leverages a combination of historical traffic data, real-time sensor inputs, and machine learning algorithms to provide accurate and actionable traffic predictions.

1. Data Integration:

Sources: Integrate data from various sources, including historical traffic records, real-time sensor data (such as cameras and IoT devices), and external factors like weather conditions and events.

Technology: Use data fusion techniques to create a comprehensive and real-time traffic dataset.

2. Advanced Predictive Models:

Algorithms: Implement machine learning algorithms (such as regression models, time series analysis, and neural networks) to analyze and predict traffic patterns.

Adaptability: Ensure models are adaptive and continuously learn from new data, improving their accuracy over time.

3. Smart City Integration:

Communication: Develop interfaces for seamless integration with existing smart city infrastructure, such as traffic lights, public transportation systems, and emergency services.

4.1 Code submission (https://github.com/Sujalsinh12345/upskillcampus/blob/main/Forecasting_of_Smart_city_traffic_patterns.ipynb)

4.2 Report submission (https://github.com/Sujalsinh12345/upskillcampus/blob/main/Forecasting_of_Smart_city_traffic_patterns_Sujalsinh_USC_UCT.pdf)

5 Proposed Design/ Model

Data Preprocessing:

- **Data Cleaning:** Removing duplicates, handling missing values, and correcting errors.
- **Data Normalization:** Scaling features to a standard range to improve model performance.
- **Data Integration:** Merging data from different sources to create a unified dataset.

Feature Engineering:

- **Feature Selection:** Identifying and selecting relevant features that impact traffic patterns (e.g., time of day, weather conditions, public events).
- **Feature Creation:** Generating new features that might help improve the model, such as traffic flow trends, congestion levels, and route popularity.

Predictive Model:

- **Algorithm Selection:** Choosing appropriate machine learning algorithms (e.g., time series models, neural networks, ensemble methods) based on the nature of the data and the problem.
- **Model Training:** Training the model on historical data using a portion of the data for training and another for validation.
- **Model Evaluation:** Evaluating model performance using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared.

Traffic Forecasting Output:

- **Forecast Generation:** Using the trained model to generate traffic forecasts for specific routes and times.
- **Visualization:** Presenting the forecasted data through dashboards, heat maps, and alerts to city planners and traffic authorities.

6 Performance Test

6.1 Test Plan/Test Cases

Memory Usage Test:

Objective: Evaluate memory consumption during model training and prediction.

Test Cases: Utilize memory_profiler to profile memory usage at different stages: data loading, preprocessing, model training, and prediction.

Execution Time Test:

Objective: Measure the time taken for various operations.

Test Cases: Utilize Python's time module or Jupyter Notebook's %time and %%time magic commands to record execution times for data loading, preprocessing, model training, and prediction.

Accuracy Assessment:

Objective: Assess the accuracy of traffic predictions.

Test Cases: Compare predicted traffic patterns with ground truth data using evaluation metrics such as Mean Absolute Error (MAE) or Root Mean Square Error (RMSE).

6.2 Test Procedure

Memory Usage Test:

Use memory_profiler to profile memory consumption.

Profile memory usage before and after each major operation (data loading, preprocessing, training, prediction).

Record memory consumption for different dataset sizes and complexities.

Execution Time Test:

Utilize Python's time module or Jupyter Notebook's magic commands.

Record execution times for data loading, preprocessing, training, and prediction on sample datasets.

Run the model multiple times to obtain average execution times.

Accuracy Assessment:

Train the model on a training dataset and evaluate its performance on a separate validation dataset.

Calculate evaluation metrics (MAE, RMSE) to quantify prediction accuracy.

6.3 Performance Outcome

Memory Usage:

Monitor memory consumption at each stage and ensure it remains within acceptable limits.

Analyze memory usage patterns for different dataset sizes and complexities.

Execution Time:

Ensure execution times meet desired performance benchmarks.

Identify any bottlenecks and optimize code for improved efficiency if necessary.

Accuracy Assessment:

Validate model performance using evaluation metrics (MAE, RMSE).

Confirm that predicted traffic patterns align well with ground truth data.

7 My learnings

During the course of this internship, I gained invaluable experience and knowledge that will significantly contribute to my career growth. Here are the key learnings from this project:

Machine Learning and Data Science:

Algorithm Proficiency I developed a strong understanding of various machine learning algorithms, including Random Forest, Gradient Boosting, and Neural Networks. This knowledge is crucial for developing predictive models and will be beneficial in any data-centric role.

Feature Engineering: I learned the importance of feature engineering and how to effectively transform raw data into meaningful features that improve model performance.

Data Handling and Preprocessing:

Data Cleaning I gained hands-on experience in cleaning and preprocessing large datasets, which is essential for ensuring data quality and reliability.

Data Visualization Tools like Seaborn and Matplotlib were used to visualize data, which is critical for understanding data distributions and identifying patterns.

Model Evaluation and Optimization:

Cross-Validation and Hyperparameter Tuning These techniques are vital for model evaluation and optimization, ensuring that models generalize well to new data.

Performance Metrics I learned to evaluate models using various performance metrics, which is key for selecting the best model for a given problem.

Real-World Application:

Industry Relevance This project provided a practical application of theoretical knowledge, bridging the gap between academia and industry. Understanding real-world constraints and how to address them is crucial for delivering viable solutions.

These learnings have equipped me with the skills and knowledge necessary for a successful career in data science and machine learning. I am now better prepared to tackle complex data-driven problems and develop innovative solutions that can make a tangible impact.

8 Future work scope

1. Incorporating Additional Data Sources:

- Weather Data: Integrating weather data could improve the accuracy of traffic predictions, as weather conditions can significantly impact traffic patterns.

- Event Data: Including data on public events, road closures, and construction could enhance the model's ability to predict unusual traffic spikes.

2. Advanced Modeling Techniques:

- Deep Learning: Exploring deep learning models, such as Long Short-Term Memory (LSTM) networks, could further improve prediction accuracy, especially for time-series data.

- Ensemble Methods: Combining multiple models into an ensemble could yield better performance by leveraging the strengths of different algorithms.

3. Real-Time Implementation:

- Edge Computing: Implementing the model on edge devices could enable real-time traffic predictions, reducing latency and improving responsiveness.

- Stream Processing: Using stream processing frameworks could allow for continuous data ingestion and real-time analysis.

4. User Interface and Visualization:

- Dashboard Development: Creating an interactive dashboard for city planners and traffic authorities could make it easier to visualize traffic predictions and derive actionable insights.

- Mobile App Integration: Developing a mobile application for real-time traffic updates could benefit commuters directly.

5. Scalability and Deployment:

- Cloud Integration: Utilizing cloud platforms for scalable data processing and storage could enhance the system's ability to handle larger datasets and more complex computations.

- API Development: Developing APIs for easy integration with other smart city infrastructure systems could extend the usability and impact of the solution.