Industrial Internship Report on

” Forecasting of Smart City Traffic Patterns”

Prepared by

[Bhoomika Sai Neerasa]

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. I have to finish the project including the report in 6 weeks’ time. The project **“Forecasting of Smart City Traffic Patterns”** focuses on analyzing and predicting vehicle traffic flow in a smart city using data science and machine learning techniques. With growing urbanization, traffic congestion has become a significant challenge affecting commute times, pollution, and city planning. This project aimed to provide data-driven insights to support traffic management by analyzing data from four major junctions, extracting features such as hour, day, month, day of the week, and weekend indicators, and identifying patterns across weekdays, weekends, and peak hours. Predictive models were developed using XGBoost regression, and performance was evaluated using MAE, RMSE, R^2, and MAPE metrics. The results showed that the model successfully captured hourly traffic trends and accurately predicted vehicle counts for each junction. Visualizations of actual vs predicted traffic highlighted the model’s effectiveness in identifying peak traffic periods. This project demonstrates the potential of machine learning in smart city applications, enabling authorities to make informed decisions for congestion management and infrastructure planning. Future work could include integrating real-time traffic data, weather, and neighboring junction influence to further improve prediction accuracy. 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 six weeks, I had the opportunity to work on the project **"Forecasting of Smart City Traffic Patterns"** as part of my industrial internship facilitated by **upskill Campus (USC)** and **UniConverge Technologies (UCT)**.

This internship gave me valuable exposure to real-world industrial problems and helped me apply my academic knowledge to practical scenarios. The problem statement was centered around forecasting traffic patterns in a smart city, with the aim of assisting the government in improving traffic management and infrastructure planning.

The program was well-structured, starting with understanding the problem statement, collecting and preprocessing data, exploring forecasting models, and finally evaluating results. Weekly progress and guidance ensured that the work was aligned with industry standards.

Through this internship, I gained practical skills in Python, data preprocessing, visualization, and machine learning model where I used XGBoost. I also learned how external events like holidays and festivals impact traffic data and how forecasting can support smart city planning.

I would like to thank **Upskill Campus**, **The IoT Academy**, and **UniConverge Technologies Pvt. Ltd.** for providing this wonderful learning opportunity. I am also grateful to my mentors, faculty members, and peers for their guidance and support throughout this internship.

To my juniors and peers, I would strongly recommend taking such industry-oriented internships, as they bridge the gap between theoretical knowledge and practical applications, while also boosting career growth and confidence.

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.

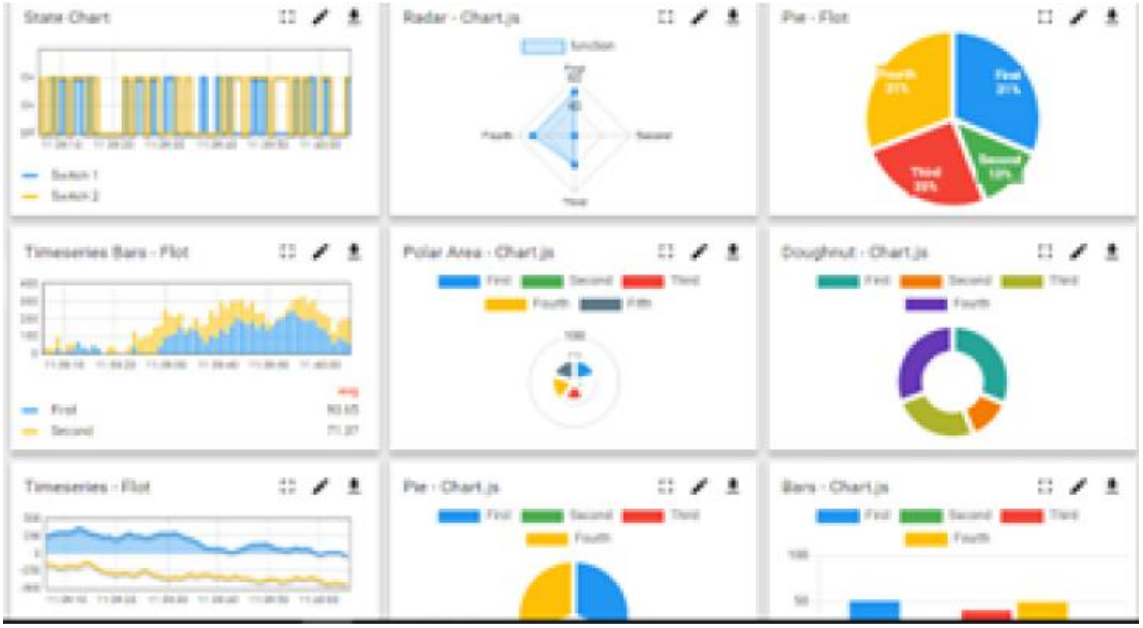
i. UCT IoT Platform (uct Insight)

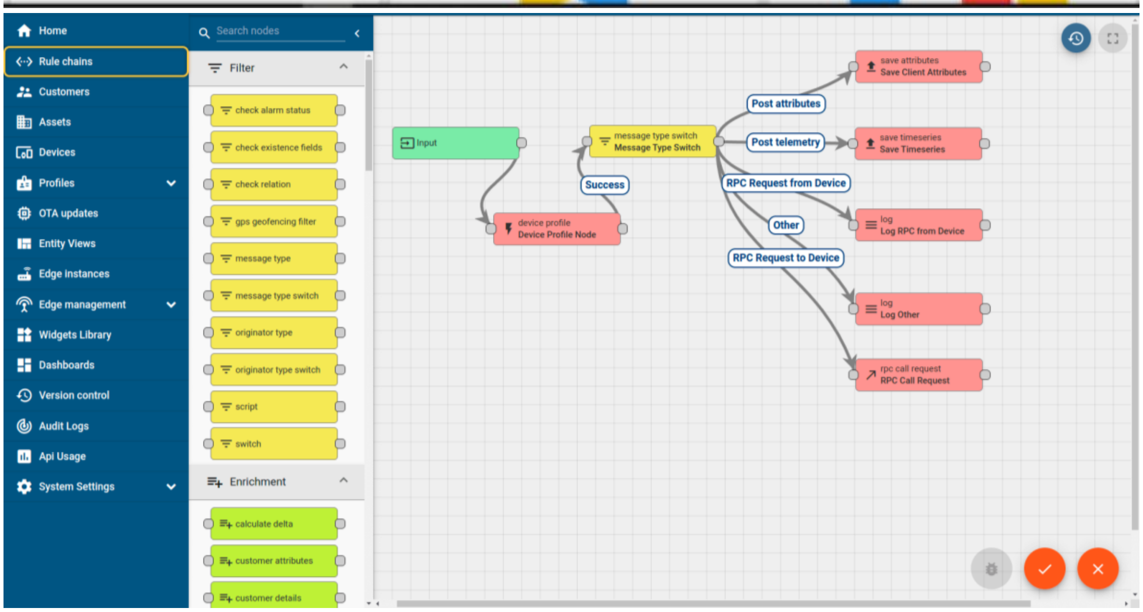
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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


State Chart
Radar - Chart.js
Pie - Flot
Timeseries Bars - Flot
Polar Area - Chart.js
Doughnut - Chart.js
Timeseries - Flot
Pie - Chart.js
Bars - Chart.js


Home
Rule chains
Customers
Assets
Devices
Profiles
OTA updates
Entity Views
Edge instances
Edge management
Widgets Library
Dashboards
Version control
Audit Logs
Api Usage
System Settings
Search nodes
Filter
check alarm status
check existence fields
check relation
gps geofencing filter
message type
message type switch
originator type
originator type switch
script
switch
Enrichment
calculate delta
customer attributes
customer details
Input
device profile
Device Profile Node
Success
message type switch
Message Type Switch
Post attributes
Post telemetry
RPC Request from Device
Other
RPC Request to Device
save attributes
Save Client Attributes
save timeseries
Save Timeseries
log
Log RPC from Device
log
Log Other
rpc call request
RPC Call Request

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.

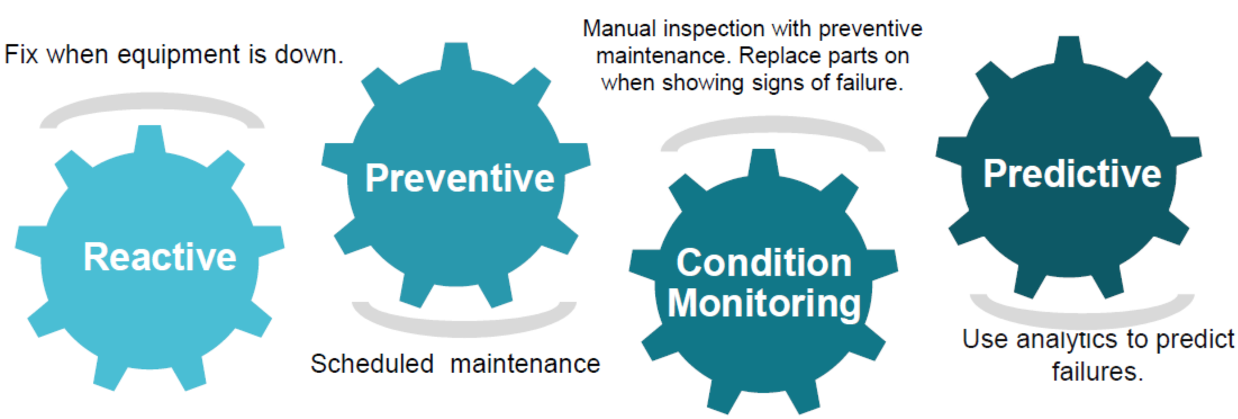
Its unique SaaS model helps users to save time, cost and money.

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

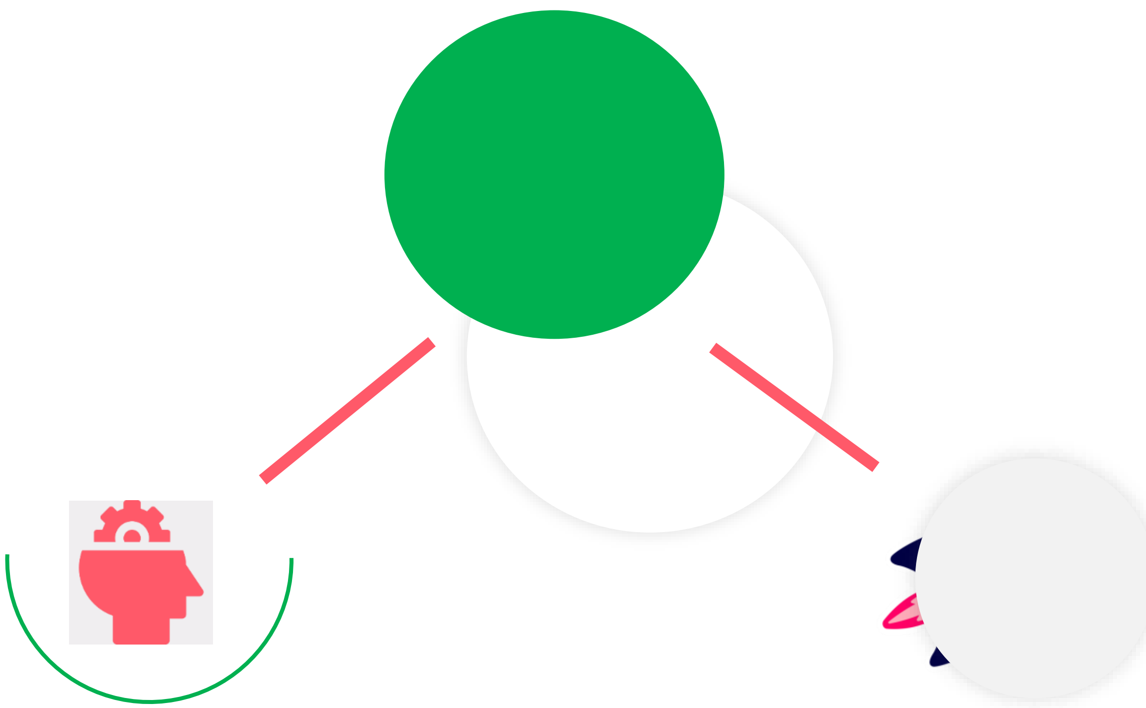
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

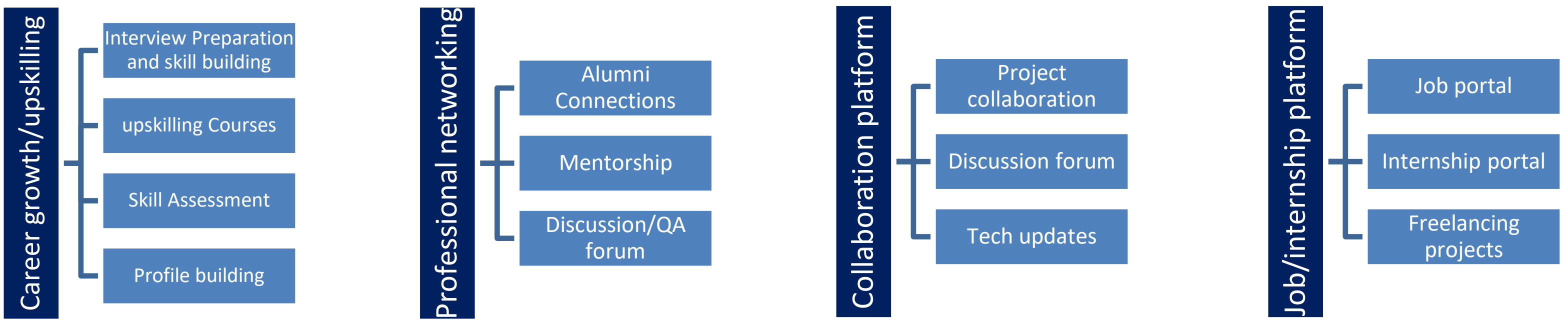
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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.

2.5 Reference

[1]. Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). *Internet of Things for SmartCities*. IEEE Internet of Things Journal, 1(1), 22–32.

[2]. Gaur, A., Scotney, B., Parr, G., & McClean, S. (2015). *Smart City Architecture and Its Applications Based on IoT*. Procedia Computer Science, 52, 1089–1094.

[3]. Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y. (2015). *Traffic Flow Prediction With Big Data: A Deep Learning Approach*. IEEE Transactions on Intelligent Transportation

2.6 Glossary

Terms	Acronym
Internet	IoT
Machine Learning	ML
Artificial Intelligence	AI
Root Mean Square Error	RSME
Mean Absolute Error	MAE

3 Problem Statement

In the assigned problem statement

Traffic congestion has become one of the most pressing challenges faced by rapidly developing cities. With increasing population and vehicle density, the existing infrastructure is often unable to handle the fluctuating traffic demands efficiently. Peak hours, holidays, and special occasions bring about irregular traffic surges, leading to delays, fuel wastage, and increased pollution. To overcome these challenges, there is a growing need for smart, data-driven solutions that can anticipate traffic patterns and help in proactive city planning.

In this project, the government aims to develop a robust traffic management system by forecasting traffic flow across four major junctions of the city. As traffic behavior varies significantly between weekdays, weekends, and holidays, incorporating such variations into the forecasting models becomes crucial. By leveraging machine learning techniques and analyzing historical traffic data, the system will provide actionable insights to predict traffic peaks. These predictions will aid in infrastructure development, optimize signal timings, and ensure better preparedness for congestion during high-demand periods. Ultimately, this solution will contribute to the broader vision of building a sustainable and intelligent smart city ecosystem.

4 Existing and Proposed solution

Existing Solution:

Several cities currently rely on conventional traffic management systems such as manually controlled traffic lights, fixed-time signal scheduling, and CCTV monitoring. More advanced cities use sensor-based systems and GPS data from vehicles to track congestion in real time. While these methods provide valuable visibility into traffic flow, they are primarily reactive in nature. They can detect congestion but cannot forecast or prevent it. Moreover, most existing solutions do not adapt to exceptional circumstances such as public holidays, festivals, or accidents, where traffic patterns deviate significantly from the norm.

Limitations of Existing Systems:

- Limited to real-time monitoring without predictive capability.
- Cannot effectively handle sudden spikes in traffic caused by special events or holidays.
- Lack of integration with future infrastructure planning, restricting long-term effectiveness.
- Manual intervention often required, reducing scalability and efficiency.

Proposed Solution:

The proposed solution uses machine learning algorithms, specifically XGBoost regression models, trained on historical traffic data. By engineering time-based features such as hour of day, day of week, month, and weekend indicators, the model can capture seasonal and temporal variations in traffic flow. Predictions are generated separately for each of the four major city junctions to ensure higher accuracy. The system goes beyond real-time monitoring by forecasting future traffic peaks, enabling authorities to take preventive measures in advance.

Value Addition:

- Provides predictive insights instead of only real-time monitoring.
- Helps government authorities with infrastructure planning by understanding long-term traffic trends.
- Improves traffic signal scheduling and reduces congestion during peak hours.
- Contributes to reduced fuel consumption and emissions, supporting sustainability goals.
- Lays the foundation for a scalable smart city traffic management system that can integrate with IoT sensors and dashboards in the future.

4.1 Code submission (Github link)

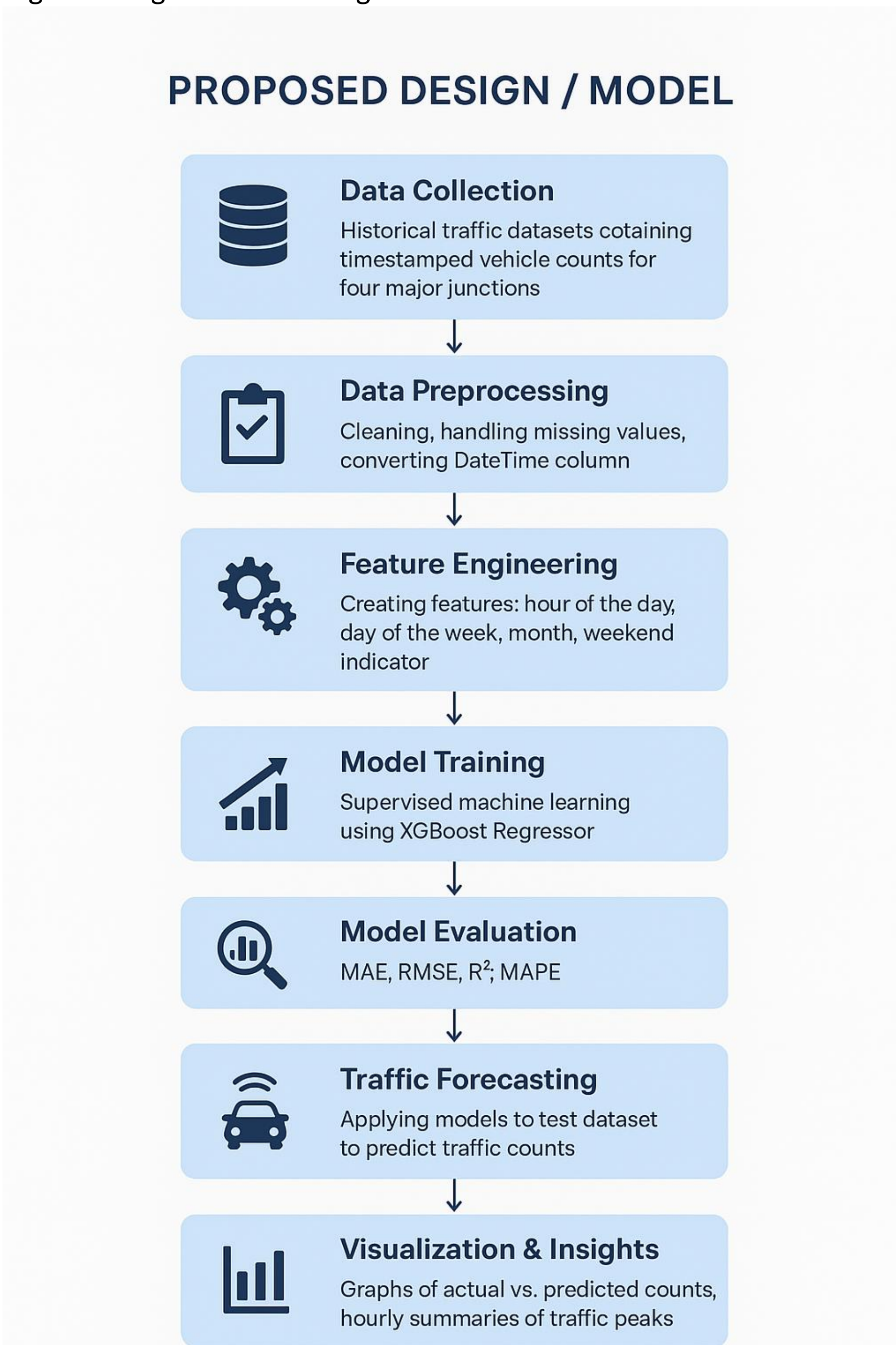
https://github.com/BhoomikaNeerasa/upskillcampus/blob/main/ForecastingofSmartcitytrafficpatterns.py

4.2 Report submission (Github link) :

https://github.com/BhoomikaNeerasa/upskillcampus/blob/main/ForecastingofSmartcitytrafficpatterns_Bhoomika_USC_UCT.pdf

5 Proposed Design/ Model

The design flow of the solution follows a structured data science pipeline, starting from raw data collection to generating actionable insights.



The steps are explained as follows:

1. **Data Collection:**
   Historical traffic datasets containing timestamped vehicle counts for four major junctions were collected. Each record included information about the junction ID, date, and time of vehicle entry.
2. **Data Preprocessing:**
   The raw data was cleaned to handle missing values and inconsistencies. Duplicate records were removed, and the DateTime column was converted into a structured format suitable for analysis. This ensured high-quality input for the modeling stage.
3. **Feature Engineering:**
   From the timestamp field, new features such as hour of the day, day of the week, month, and weekend indicator were created. These features allowed the model to capture temporal traffic variations such as rush hours, weekday vs. weekend behavior, and seasonal trends.
4. **Model Training:**
   The dataset was split into training and validation sets. A supervised machine learning approach using XGBoost Regressor was implemented. The model was trained separately for each junction to capture junction-specific traffic behavior.
5. **Model Evaluation:**
   The trained models were evaluated using metrics such as Mean Absolute Error (MAE)**,** Root Mean Squared Error (RMSE)**,** R^2 score**,** and Mean Absolute Percentage Error (MAPE)**.** These ensured that the model's predictive performance was quantified and reliable.
6. **Traffic Forecasting:**
   Once validated, the models were applied to the test dataset to generate traffic forecasts for each junction. The predictions provided estimates of vehicle flow for different hours of the day and across different days.
7. **Visualization & Insights:**
   Forecasted results were visualized using graphs comparing actual vs. predicted traffic counts. Hourly summaries were also generated to highlight traffic peaks, which serve as valuable inputs for infrastructure planning and traffic signal optimization**.**

6 Performance Test

Constraints & Design Considerations:

- **Accuracy:** Achieved using XGBoost and temporal features (hour, day, weekend).
- **Speed / Efficiency:** Preprocessing and chunked predictions enabled near real-time forecasting.
- **Memory Usage:** Data handled in chunks; unnecessary columns removed.
- **Scalability:** Separate models per junction allow parallel processing.
- **Durability:** Missing/inconsistent data handled via imputation.

Recommendations:

- Use distributed computing or GPU acceleration for larger-scale deployment.
- Apply incremental learning for continuous updates.
- Periodically retrain models to adapt to evolving traffic patterns.

6.1 Performance Outcome

Test Results by Junction:

Junction	MAE	RMSE	R^2	MAPE
1	10.89	13.51	0.6503	28.36%
2	3.50	4.57	0.6104	28.83%
3	4.13	7.27	0.5194	40.26%
4	1.84	2.57	0.5276	30.48%

Interpretation:

- **Junction 1**: Highest absolute errors (MAE, RMSE) due to high traffic volume but R^2 of 0.65 indicates good variance explanation.
- **Junction 2**: Low absolute errors; model captures traffic trends well.
- **Junction 3**: High MAPE (40.26%) indicates more irregular traffic patterns; prediction errors are relatively larger.
- **Junction 4**: Low traffic junction; low errors in both absolute and relative terms.

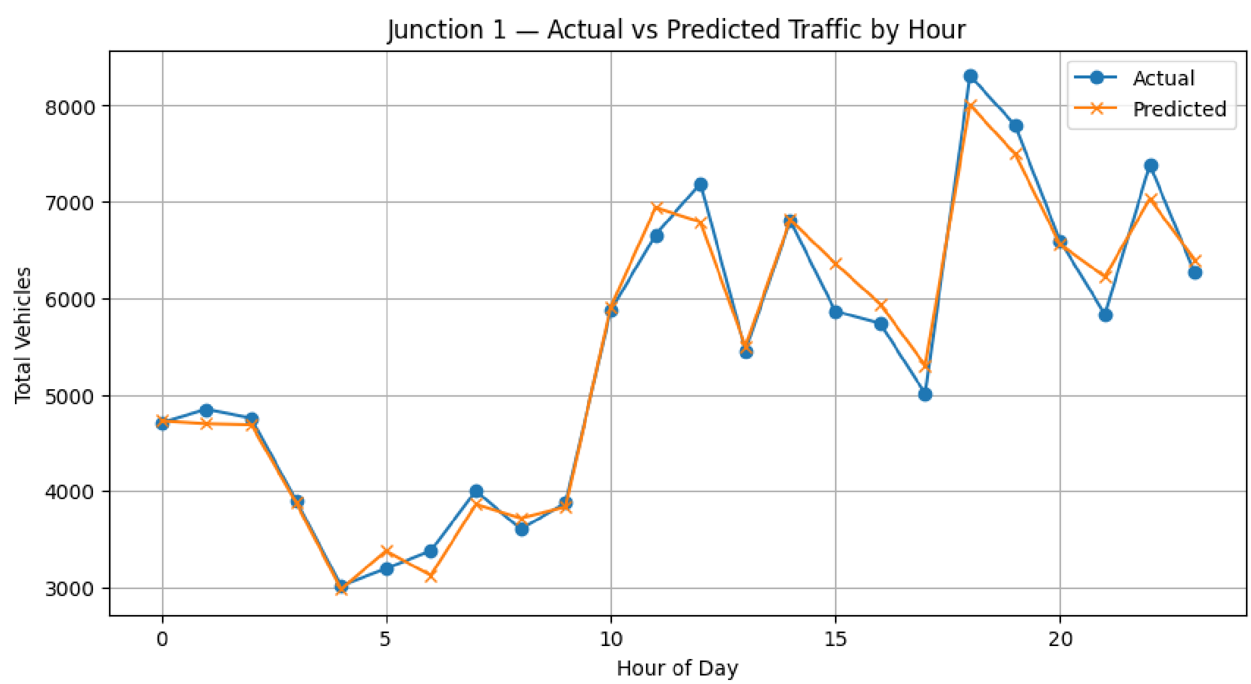


Figure 1

Figure 1 : Comparison of actual and predicted vehicle counts for Junction 1. The model captures the hourly traffic trends well, including morning and evening peaks, with minor deviations at peak hours.

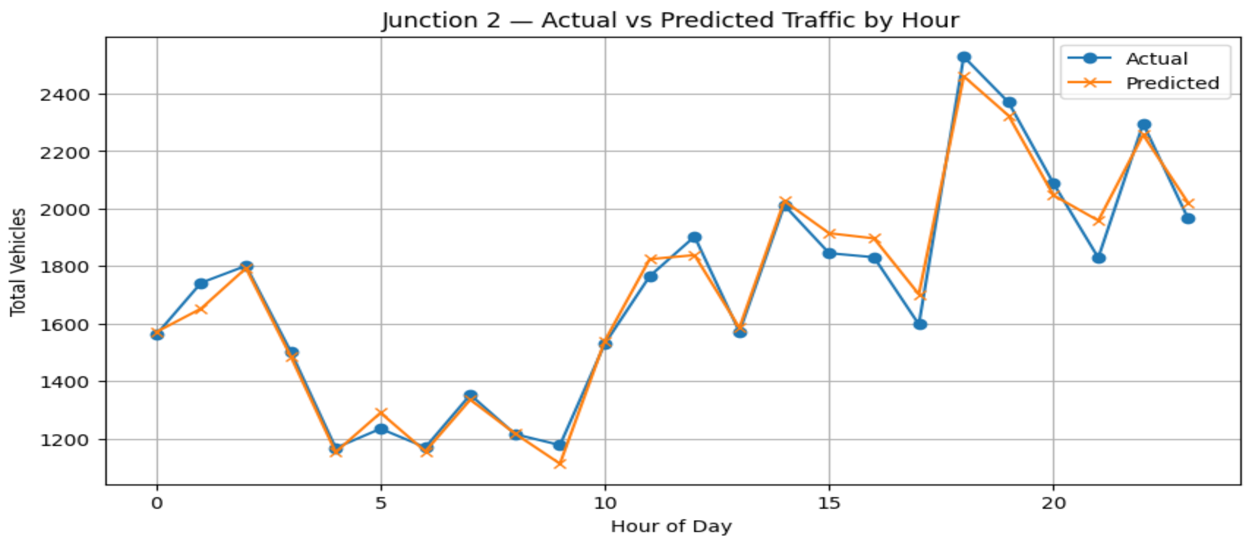


Figure 2

Figure 2: Hourly vehicle traffic at Junction 2: predicted values closely follow the actual counts, showing accurate representation of traffic fluctuations throughout the day.

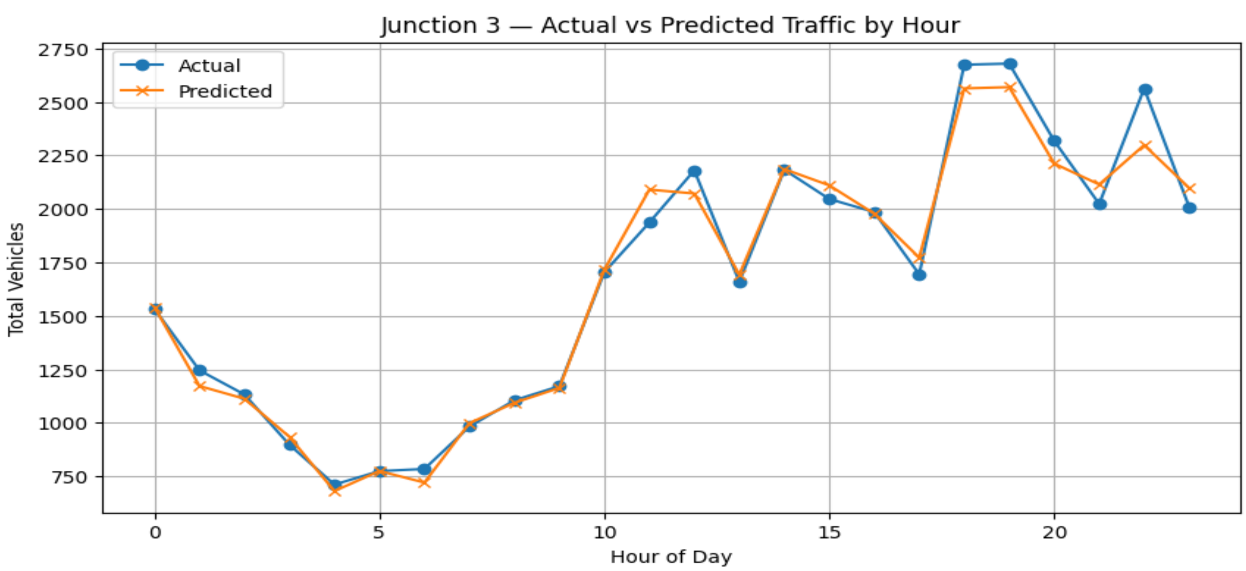


Figure 3

Figure 3 : The model captures general trends for Junction 3 but deviations are observed during irregular traffic periods, reflecting higher variability at this junction.

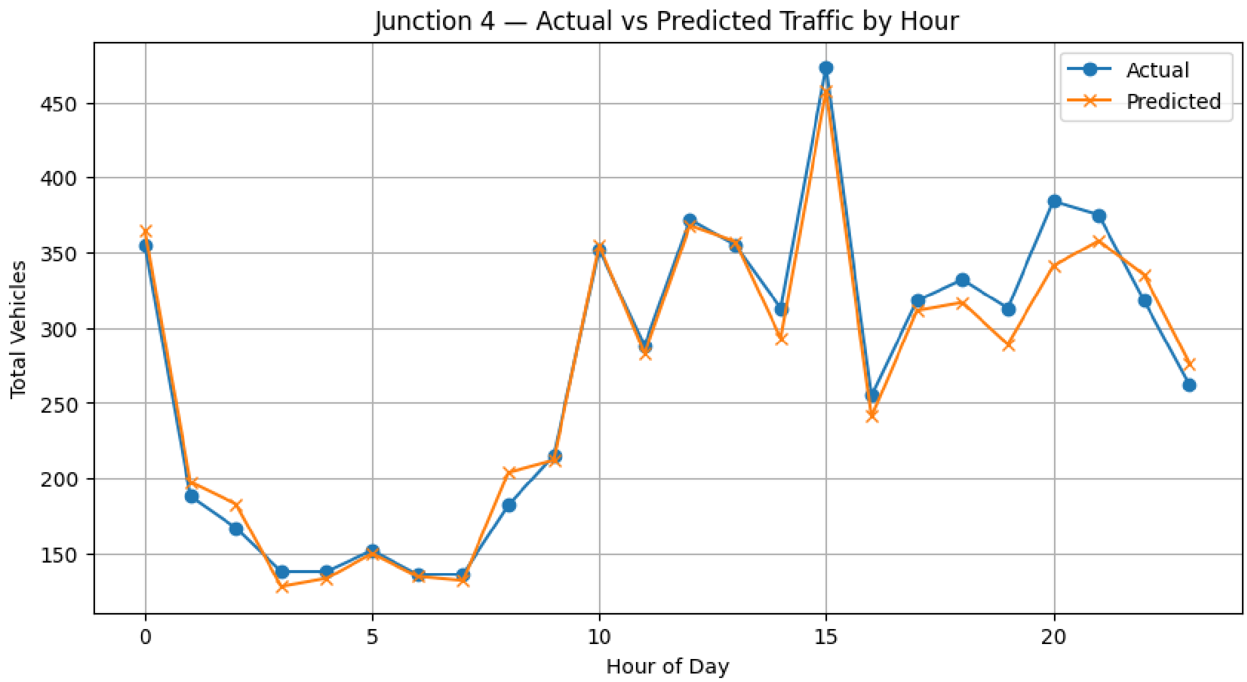


Figure 4

Figure 4 : Comparison of predicted and actual hourly traffic at Junction 4. The low-traffic junction shows good alignment between predictions and actual counts, with minimal error.

7 My learnings

1. Technical Skills:

- Learned data preprocessing, feature engineering, and handling time-based traffic data.
- Applied XGBoost for regression and evaluated models using MAE, RMSE, R^2, and MAPE.
- Gained experience in visualizing results and generating predictions for multiple junctions.

2. Problem-Solving & Analytics:

- Analyzed traffic patterns and identified trends.
- Interpreted deviations and understood model limitations.
- Developed skills in deriving insights from numerical and visual outputs.

3. Career Growth & Application:

- Experience in smart city and transportation analytics relevant to AI and IoT roles.
- Built end-to-end ML workflow experience, useful for Data Scientist or ML Engineer roles.
- Improved skills in presenting technical results to stakeholders.

4. Reflection:

- Understood importance of feature selection, evaluation, and iterative improvement.
- Combined ML techniques with domain knowledge to improve predictions.
- Strengthened technical and analytical skills for future AI and data-driven projects.

8 Future work scope

- **Incorporate additional features:** Include weather data, public events, holidays, and road conditions to improve prediction accuracy.
- **Time-series models:** Experiment with LSTM, GRU, or Prophet for better handling of sequential traffic patterns.
- **Real-time predictions:** Integrate live traffic feeds for dynamic, real-time forecasting.
- **Spatial modeling:** Consider neighboring junction traffic influence using graph-based or spatial ML models.
- **Hyperparameter tuning & ensembles:** Optimize XGBoost parameters or combine multiple models to reduce prediction errors.
- **Visualization dashboards:** Build interactive dashboards for city authorities to monitor traffic patterns and plan interventions.
- **Long-term forecasting:** Extend predictions beyond hourly traffic to daily, weekly, or monthly planning.