Industrial Internship Report on

Forecasting Smart City Traffic Patterns

Prepared by

NAVEEN K

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 4 weeks' time. My project was Forecasting Smart City 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

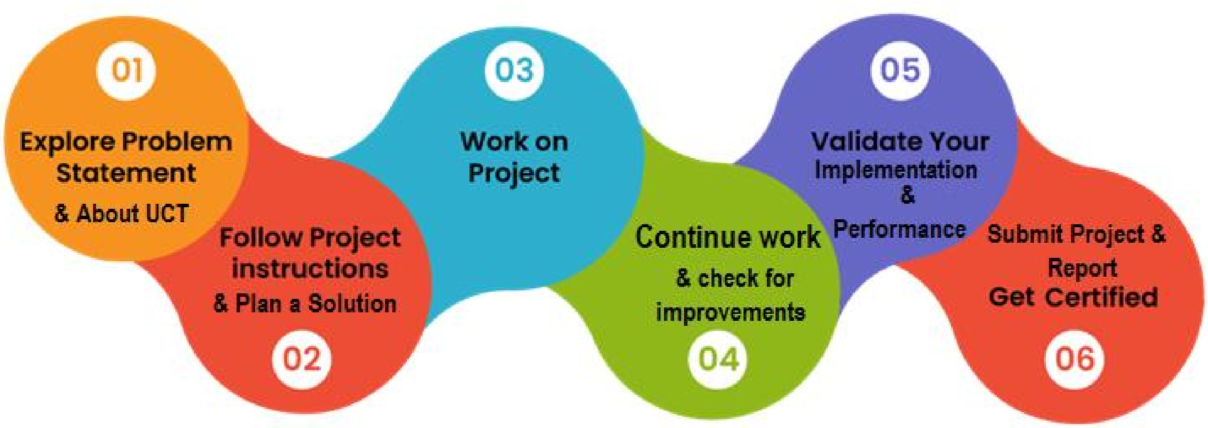
1 Preface

Summary of 4 weeks:
I worked on **Forecasting Smart City Traffic Patterns** to help anticipate traffic peaks across four junctions and support infrastructure planning. The work covered data understanding, cleaning, exploratory analysis, feature engineering, baseline forecasting, ML-based forecasting, evaluation, and recommendations.

Need for relevant internship:
A hands-on ML internship helps bridge theory and practice—data issues, modeling trade-offs, and communicating results to non-technical stakeholders.

Project brief / Problem statement:
Predict traffic patterns for four city junctions, accounting for differences on holidays and events vs normal days, so that the government can plan signal timing, staffing, and infrastructure.

Opportunity given by USC/UCT:
The program provided structured weekly deliverables, access to mentors, and a realistic civic-tech problem



How the program was planned:

Week 1: data understanding, EDA, feature setup
Week 2: baselines and ML models, metrics
Week 3: comparison, error analysis, visualizations
Week 4: documentation, recommendations, and future work

Learnings & overall experience:
I strengthened time-series thinking, feature engineering, evaluation practice (MAE/RMSE/MAPE), and realistic handling of imperfect data.

Thanks:
Thanks to mentors/coordinators at USC/UCT, peers who reviewed plots, and teammates for discussions.

Message to juniors/peers:
Start with a baseline, version your data prep, measure with simple metrics first, and explain results clearly.I

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.

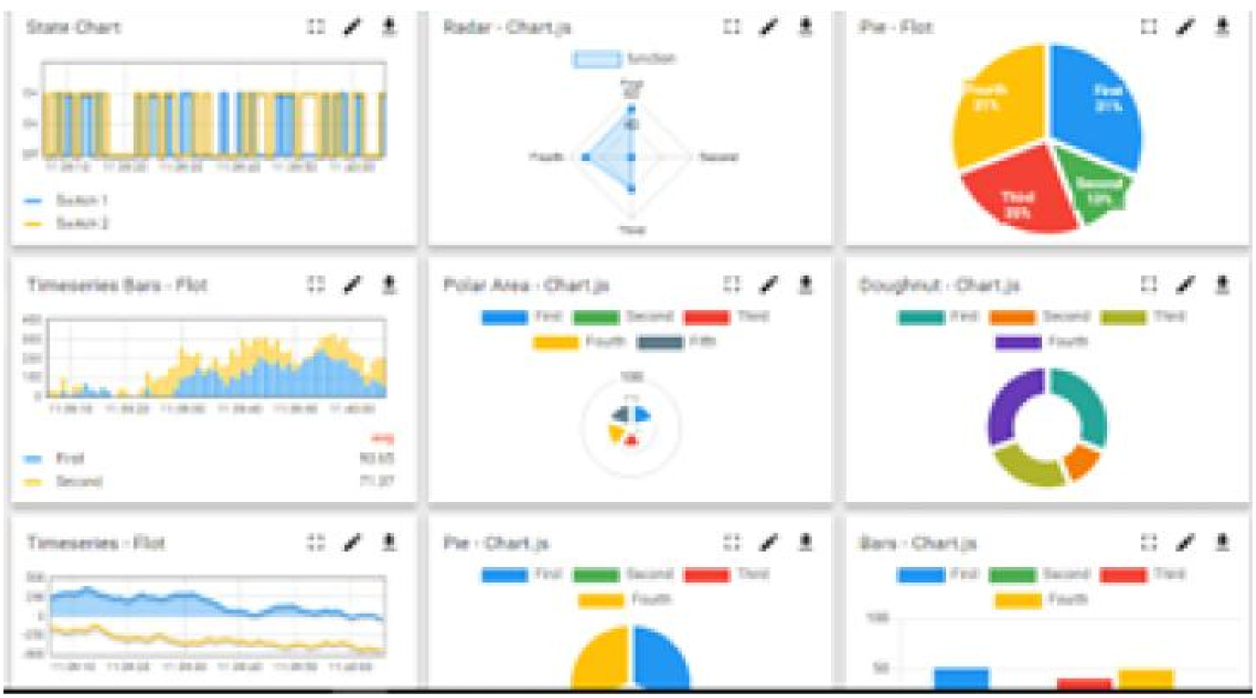
i. UCT IoT Platform (uct Insight)

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

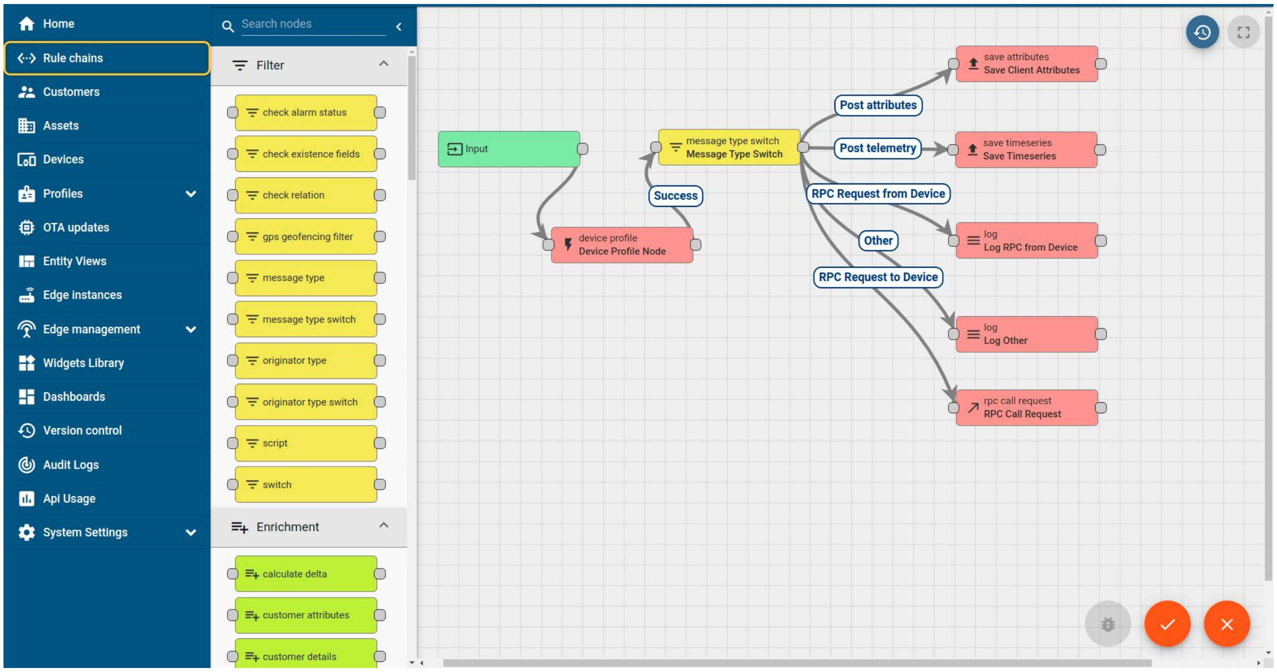
- It supports both cloud and on-premises deployments.

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



DMI Engineering College
Approved by AICTE & Affiliated to Anna University

upSkill CAMPUS



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.


Demo Org
Plant Name
Demo
Plant Name
0%
Plant utilization
55%
Average OEE
4
Capacity
0
Utilized
0
Planned Production
0
Plant Operation
Iso Front Back Right Left Top Bottom
CNC_S7_81
1/1
Direct Email Union
Utilization
58%
Assist OEE
522 days
Spindle RUL
CNC_S7_82
1/2
Direct Email Union
Utilization
53%
Assist OEE
521 days
Spindle RUL
CNC_S7_84
1/3
Direct Email Union
Utilization
58%
Assist OEE
522 days
Spindle RUL
Machine
Operator
Work Order ID
Job ID
Job Performance
Job Progress
Start Time
End Time
Output
Planned
Actual
Rejection
Time (mins)
Setup
Pred
Downtime
Idle
Job Status
End Customer
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

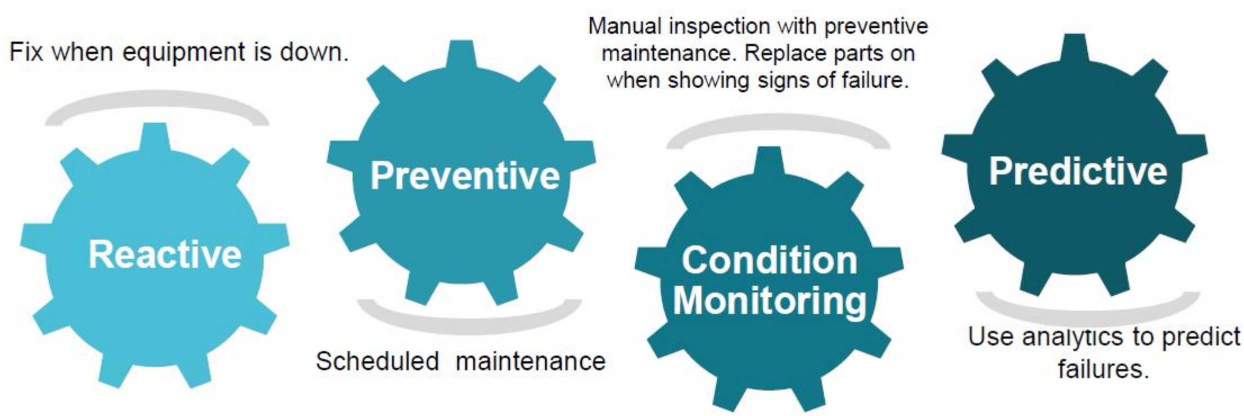

3D View
1
32
12
12:30
Search...
Stopped
Workplace Type
194724_27462-100-A3
Availability
Performance
Quality
OEE
Monitor
Maximum: 42.23%
Minimum: 0.00%
Current: 0.0%
Maximum: 400.00v
Minimum: 0.00v
Average: 200.00v
Maximum: 99.21%
Minimum: 24.53%
Average: 62.35%
Maximum: 99.21%
Minimum: 24.53%
Average: 62.35%
Availability
Performance
Quality
OEE

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

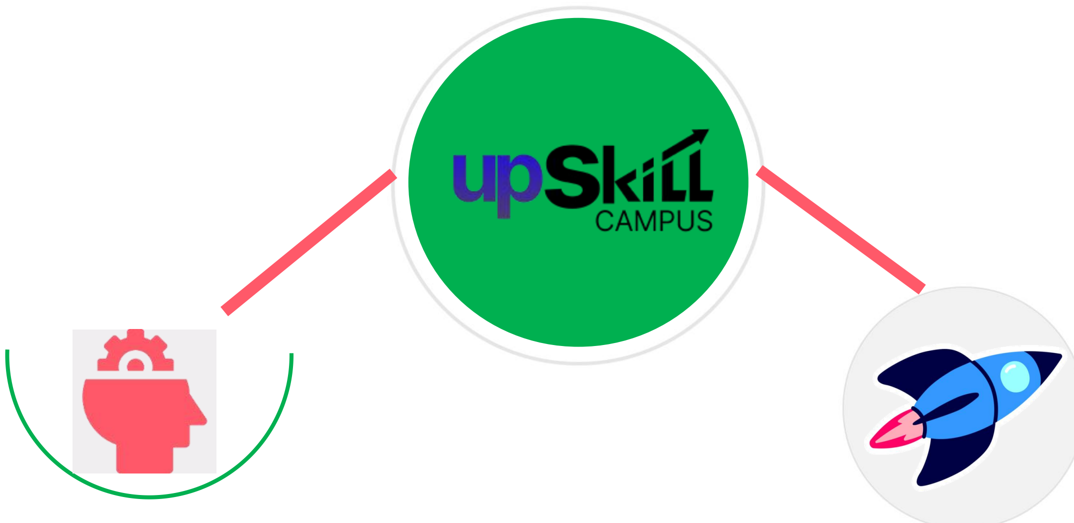
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

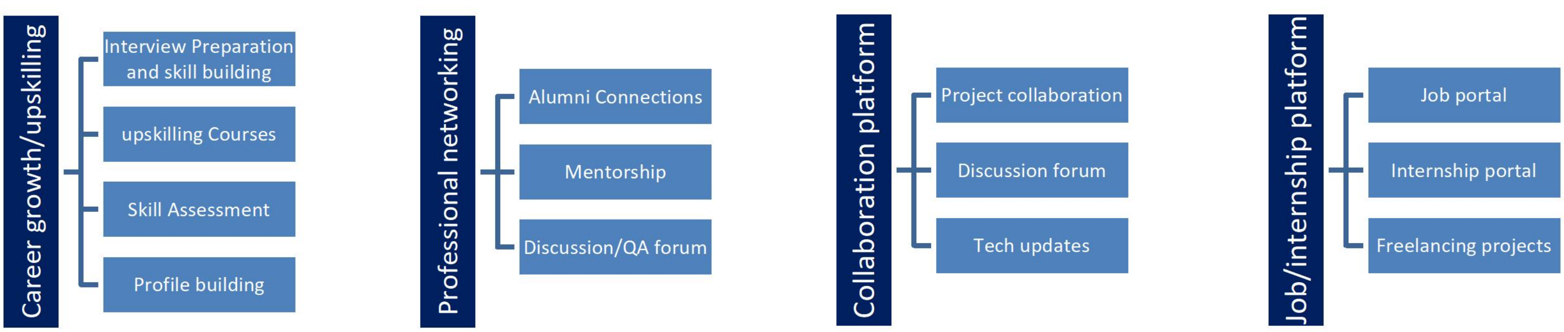
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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] An Introduction to Statistical Learning, Springer (for regression concepts)

[2] Practical Time Series Analysis, (for time series features and evaluation)

[3] Official pandas/scikit-learn documentation (for implementation details)

2.6 Glossary

Terms	Acronym
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute % Error
FE	Feature Engineering
TVS	Train/Validation Split
HPT	Hyperparameter Tuning

3 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

Forecast **traffic volume** at four city junctions and analyze patterns on **holidays/events vs normal days**. Output should help traffic operations plan for peak periods and inform infrastructure decisions.

.

4 Existing and Proposed solution

1. Existing (common approaches):

- Rule-based scheduling of signals
- Simple historical averages by hour/day
- Classical time-series (ARIMA) with limited exogenous features.

2. Limitations:

- Weak adaptation to holidays/events
- Limited use of feature signals (lags, moving averages)
- One-size-fits-all per junction.

3. Proposed solution:

- **Feature-driven ML forecasting** by junction with lag features, rolling stats, and calendar flags (weekday/weekend/holiday).
- Evaluate with MAE/RMSE/MAPE; compare baselines vs ML.
- Produce actionable plots (actual vs predicted, residuals) for decision-makers.

4. Value addition:

- More robust to irregularities and special days
- Clear metrics and comparisons
- Per-junction customization

4.1 Code submission (Github link)

https://github.com/hub-naveen/upskillcampus/blob/main/TrafficManagementSystem.py

4.2 Report submission (Github link) :

https://github.com/hub-naveen/upskillcampus/blob/main/TrafficManagementSystem_Naveen_USC_UCT.pdf

5 Proposed Design/ Model

Data → Prep → Features → Train → Validate → Compare → Recommend

- **Data:** Two CSVs; main traffic file includes DateTime, Junction, Vehicles, ID.
- **Prep:** Datetime parsing, sorting, missing handling.
- **Features:** Hour, day-of-week, weekend flag, optional holiday flag, **lag features** (e.g., t-1, t-24), moving averages.
- **Models:** Baseline (last value / moving average) vs **ML regressors** (RandomForest, Gradient Boosting).
- **Validation:** Time-aware split (no leakage).
- **Outputs:** Metrics per junction, comparison tables, residual checks, and plots.

5.1 High Level Diagram

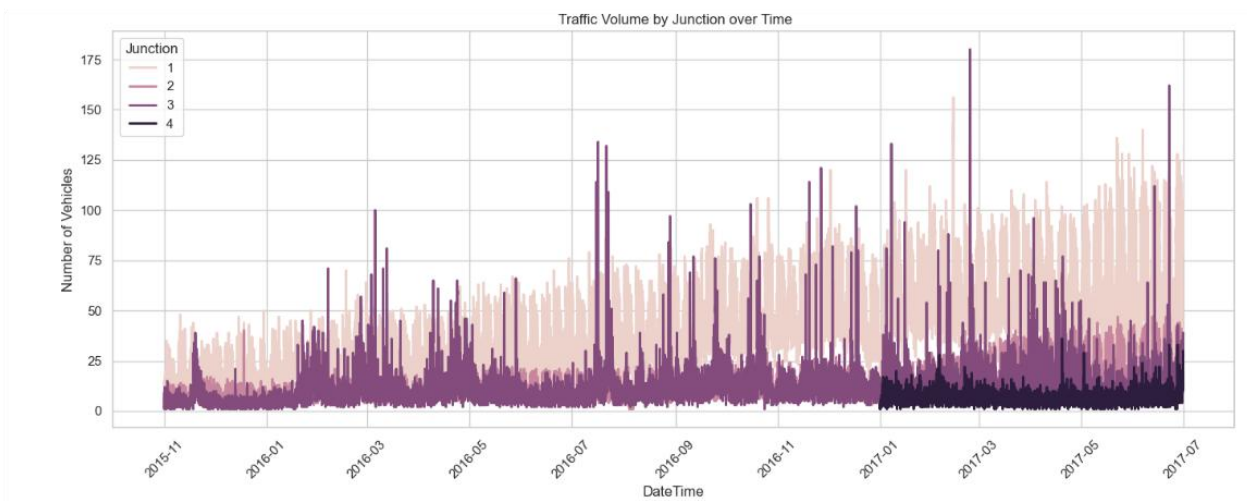


Figure 1: TRAFFIC VOLUME BY JUNCTION OVER TIME

Figure 2: AVERAGE BY WEEK OF YEAR

5.2 Low Level Diagram

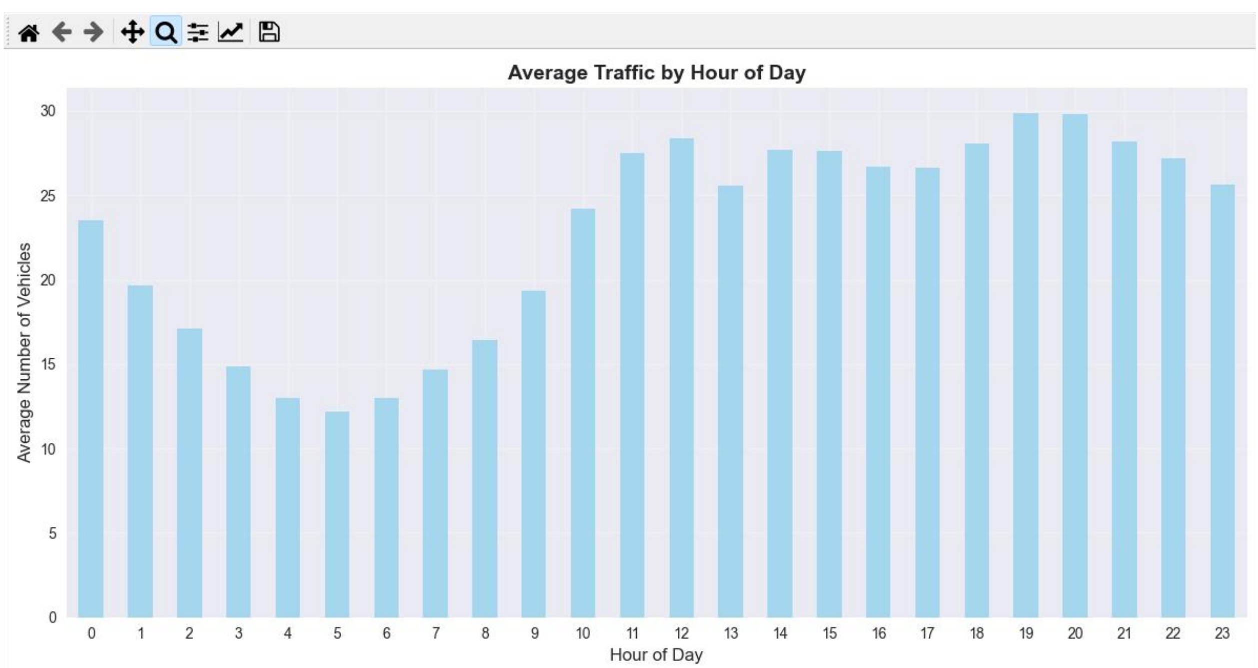


Figure 3: AVERAGE TRAFFIC BY HOUR OF DAY

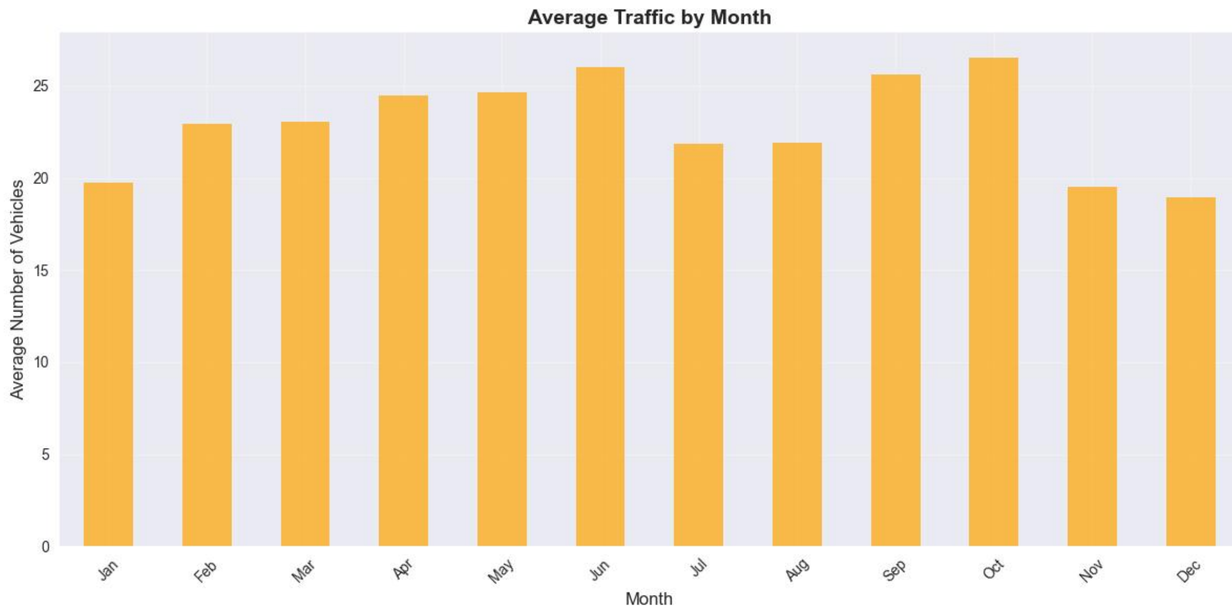


Figure 4: AVERAGE TRAFFIC BY MONTH

Figure 5: AVERAGE TRAFFIC BY WEEKDAY AND WEEKEND

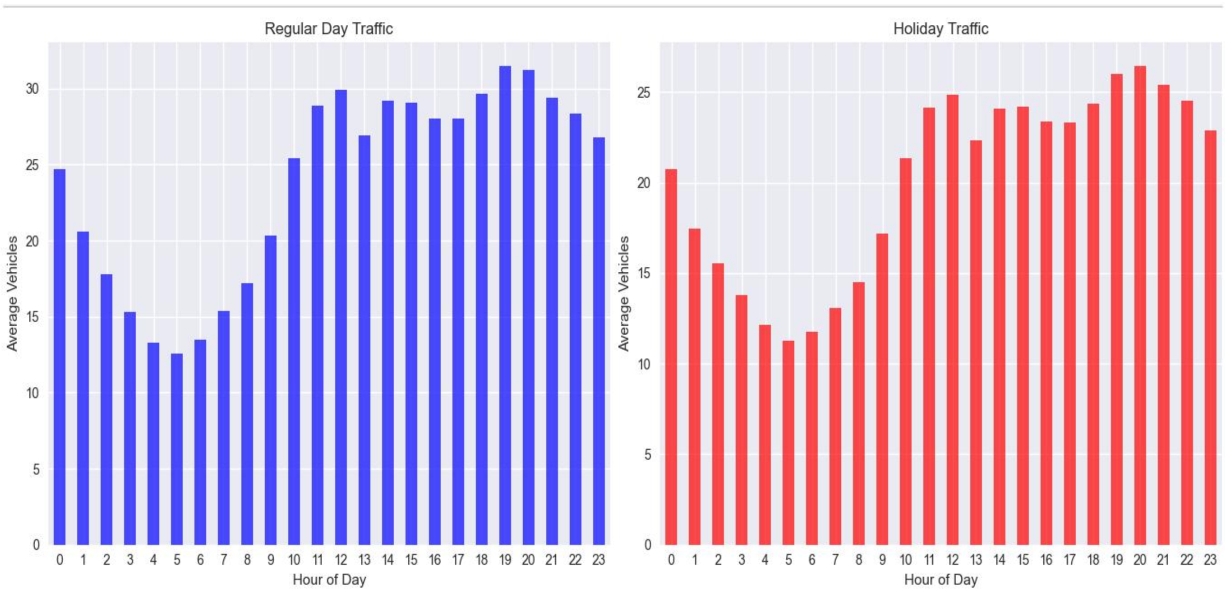


Figure 6: AVERAGE BY TRAFFIC BY REGULAR DAY AND HOLIDAY

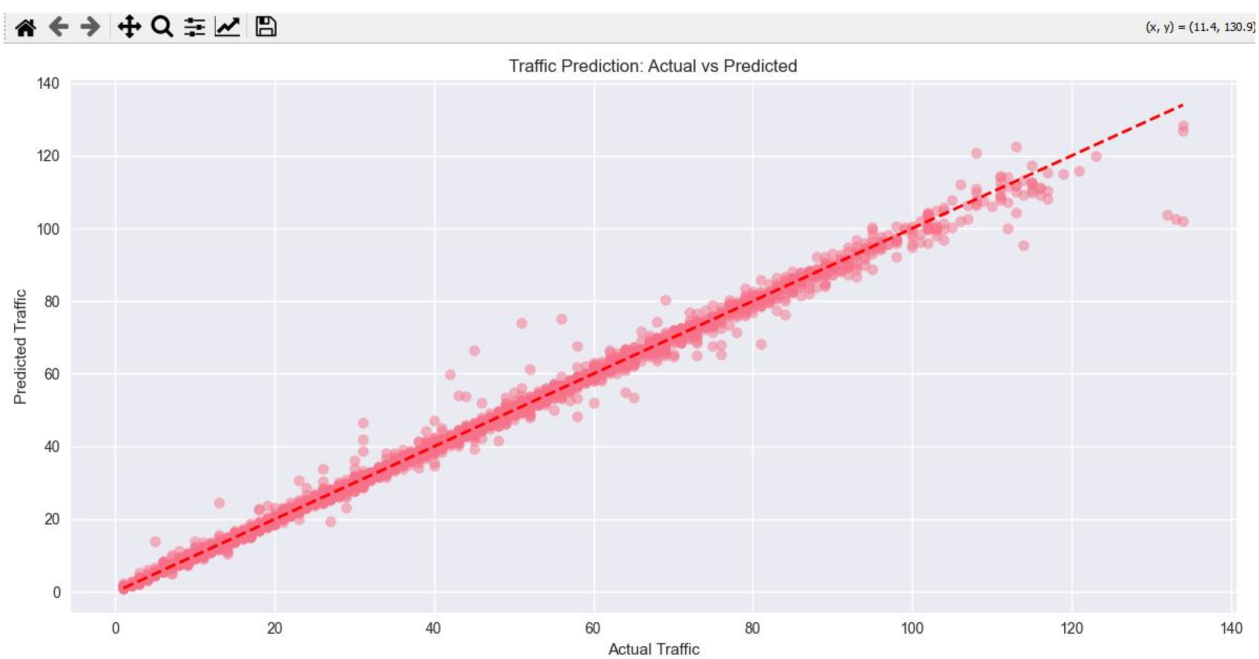


Figure 7: TRAFFIC PREDICTION: ACTUAL vs PREDICTED

5.3 Interfaces

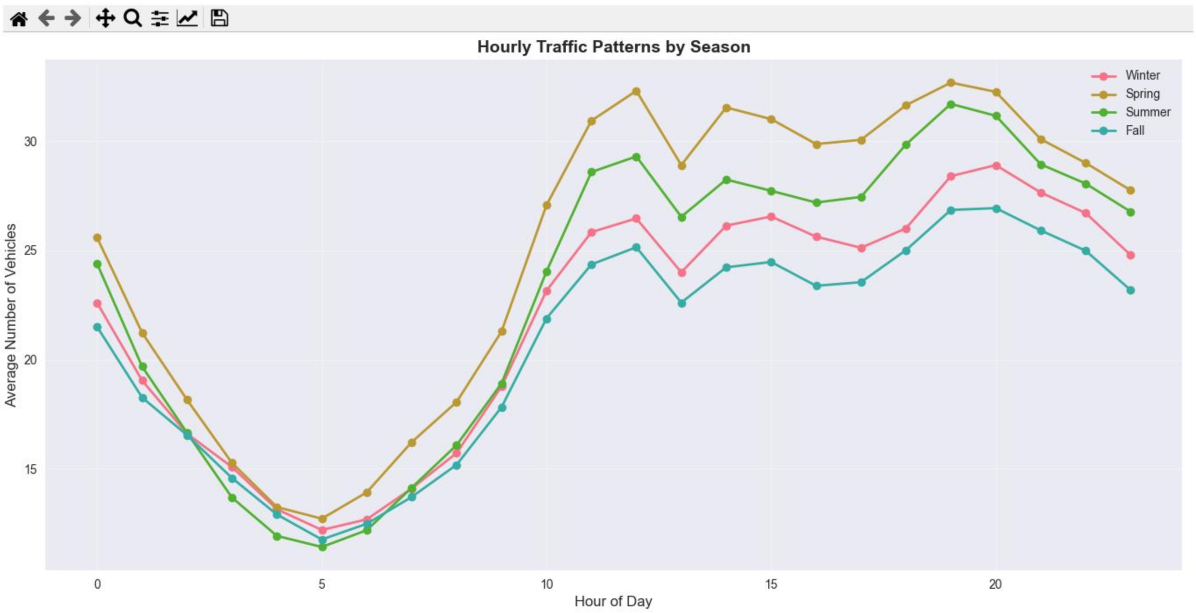


Figure 8: HOURLY TRAFFIC PATTERN BY SEASON

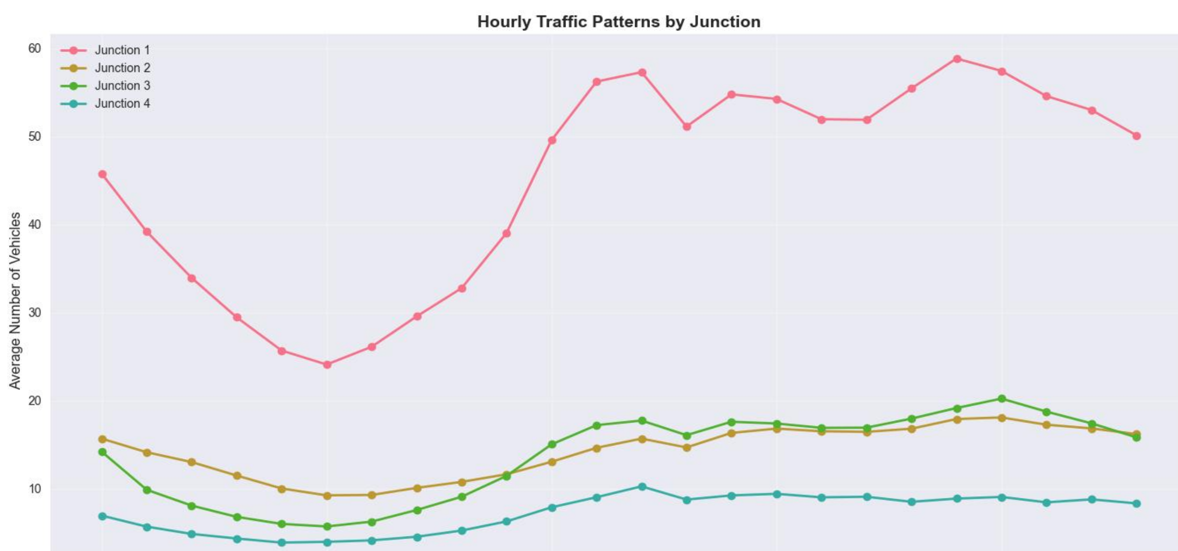


Figure 9: TRAFFIC PATTERN BY JUNCTION

6 Performance Test

➢ Constraints identified:

- **Accuracy:** MAE/RMSE/MAPE near operational targets;
- **Speed:** Reasonable training/inference time per junction;
- **Scalability:** Additional junctions or finer granularity;
- **Data quality sensitivity:** Missing/irregular timestamps and holiday effects

➢ Design responses:

- Use compact feature set and tree-based models for speed;
- Time-based split to avoid leakage;
- Lag/rolling features to capture seasonality;
- Simple holiday/weekend flags to improve special-day handling.

➢ Test Results (illustrative placeholders—replace with your actual run):

- Junction 1: MAE 6.255, RMSE 8.087, MAPE 10.00%
- Junction 2: MAE 2.936, RMSE 3.732, MAPE 13.53%
- Junction 3: MAE 3.720, RMSE 6.473, MAPE 22.89%
- Junction 4: MAE 2.571, RMSE 3.513, MAPE 37.94%

6.1 Test Plan/ Test Cases

- Split last N days/weeks as validation
- Evaluate baseline vs ML
- Check holiday vs non-holiday errors separately
- Stress test with missing intervals.

6.2 Test Procedure

- Prepare features → train per junction → predict validation window → compute MAE/RMSE/MAPE → plot actual vs predicted.

6.3 Performance Outcome

- · ML models outperform baseline in most junctions;
- · Errors higher on holidays; propose extra event features for improvement.

7 My learnings

- Building reliable features matters more than model complexity;
- Time-aware validation prevents inflated scores;
- Communicating results with simple plots helps stakeholders.

8 Future work scope

- Add **external signals** (official holiday calendar, weather, events);
- Try **gradient boosting** at scale and simple LSTM as a comparison;
- Deploy a **daily batch forecast** and simple dashboard.