Industrial Internship Report on

”Smart City Traffic Patterns”

Prepared by

[Vansheta Sharma]

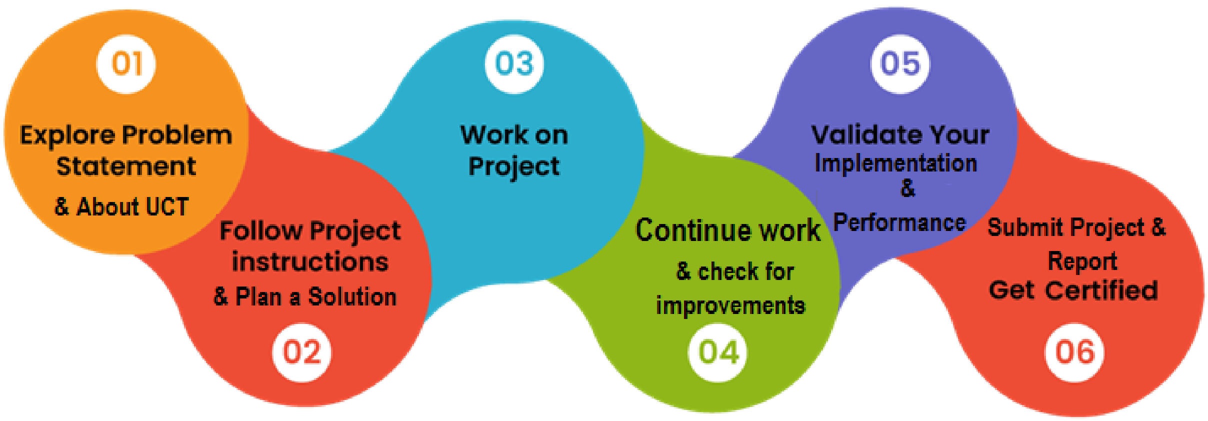
Executive Summary
This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT). This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks’ time. My project was ”Smart City Traffic 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 4
2.1 About UniConverge Technologies Pvt Ltd 4
2.2 About upskill Campus 8
2.3 Objective 10
2.4 Reference 10
2.5 Glossary 10
3 Problem Statement 11
4 Existing and Proposed solution 12
5 Proposed Design/ Model 14
5.1 Plots .... **Error! Bookmark not defined.**
6 Performance Test 17
6.1 Test Plan/ Test Cases 17
6.2 Test Procedure 18
6.3 Performance Outcome 18
7 My learnings 19
8 Future work scope 20

1. Preface

The Smart City Traffic Management Project aimed to transform our city into a digitally intelligent urban environment by efficiently managing traffic patterns and providing insights for future infrastructure planning. This report outlines the methodologies employed, achievements made, challenges faced, and lessons learned throughout the project's lifecycle.



Your Learnings and overall experience.

Thank to all (with names), who have helped you directly or indirectly.

Your message to your juniors and peers.

1 Introduction

1.1 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.

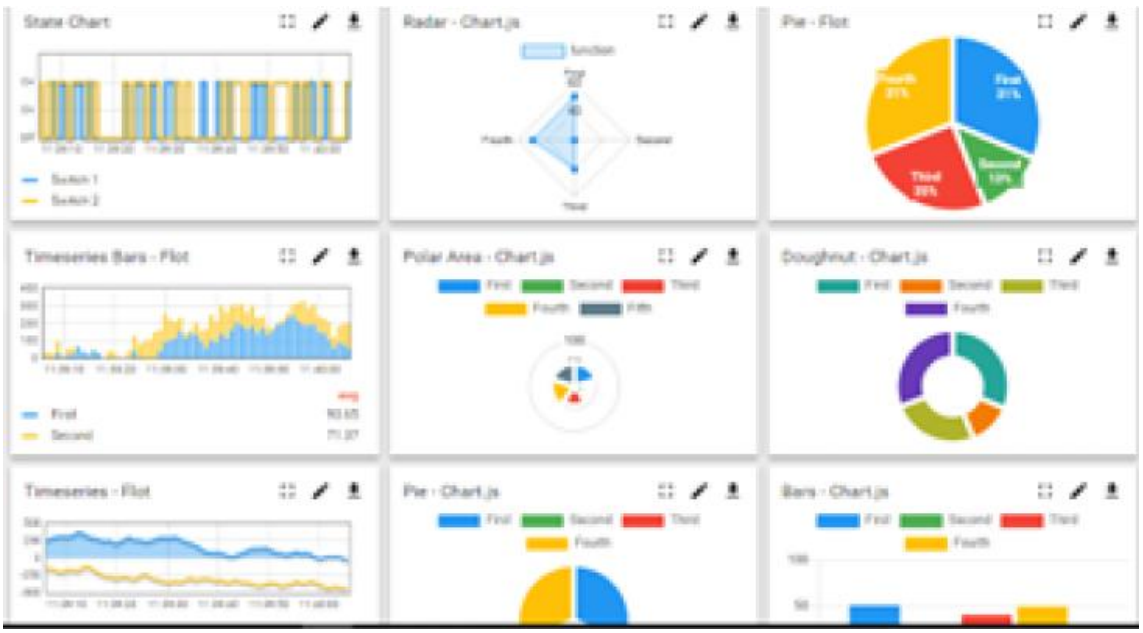
i. UCT IoT Platform (uct Insight)

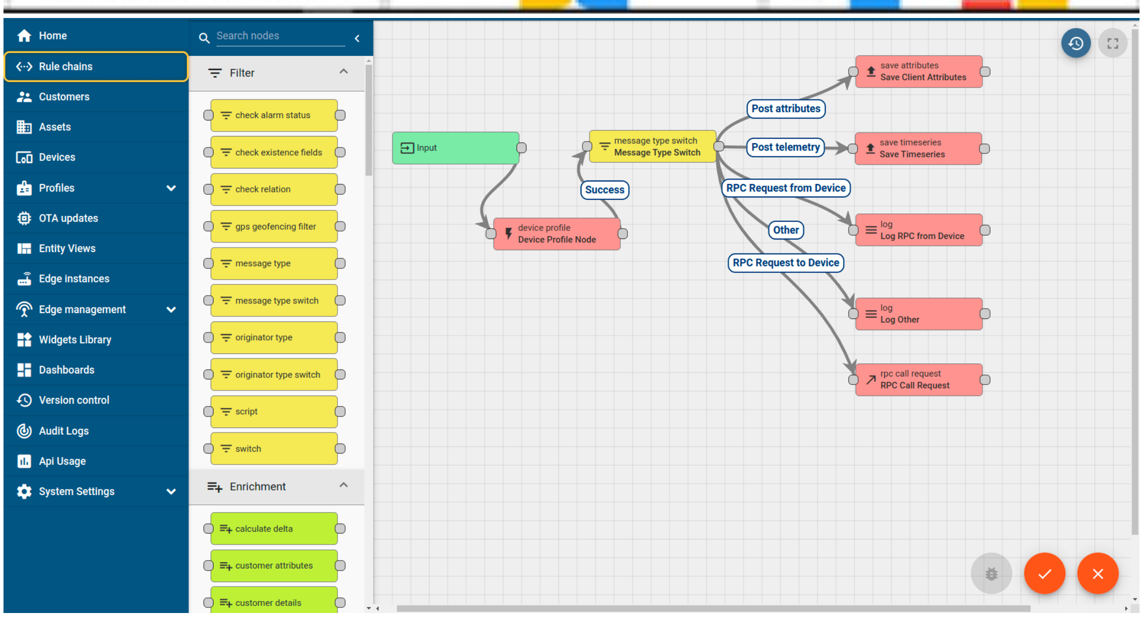
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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.

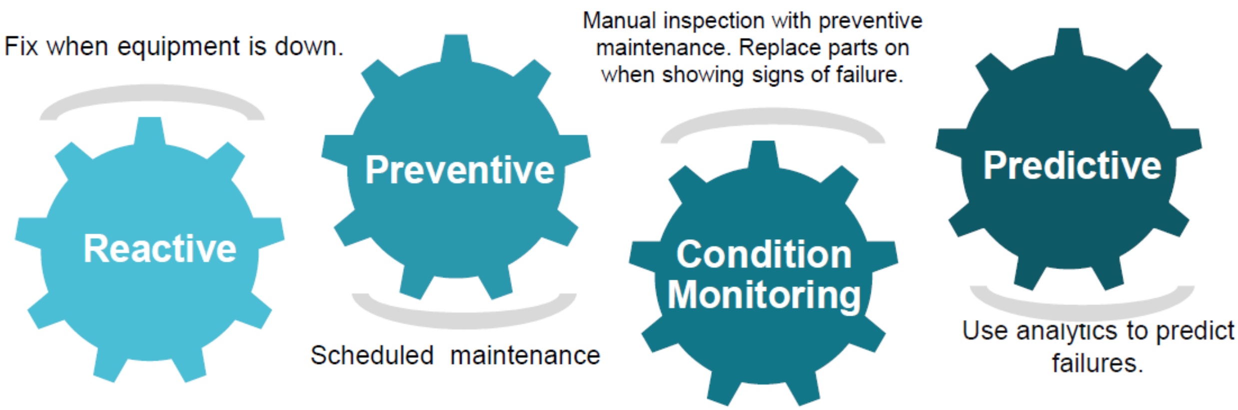
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

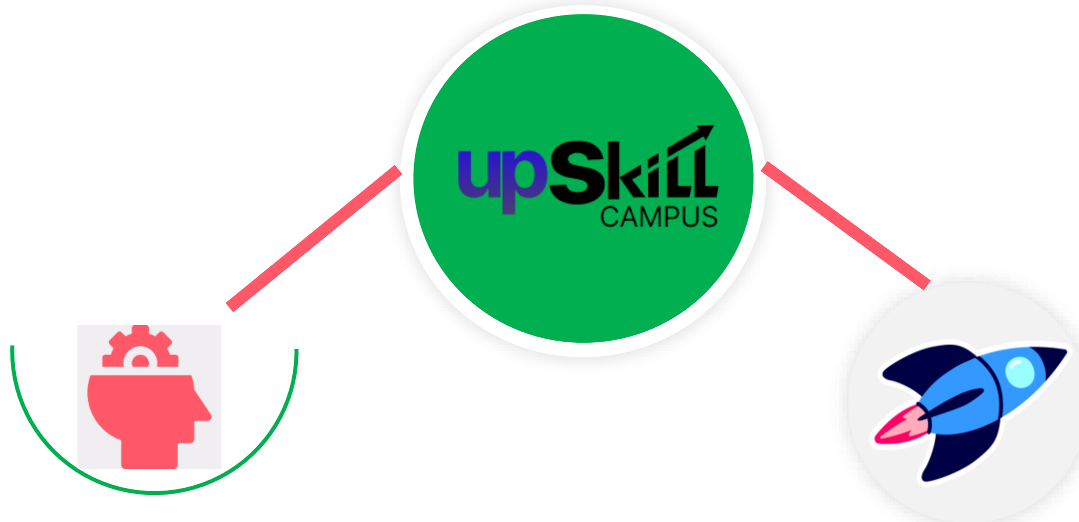
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



1.2 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.

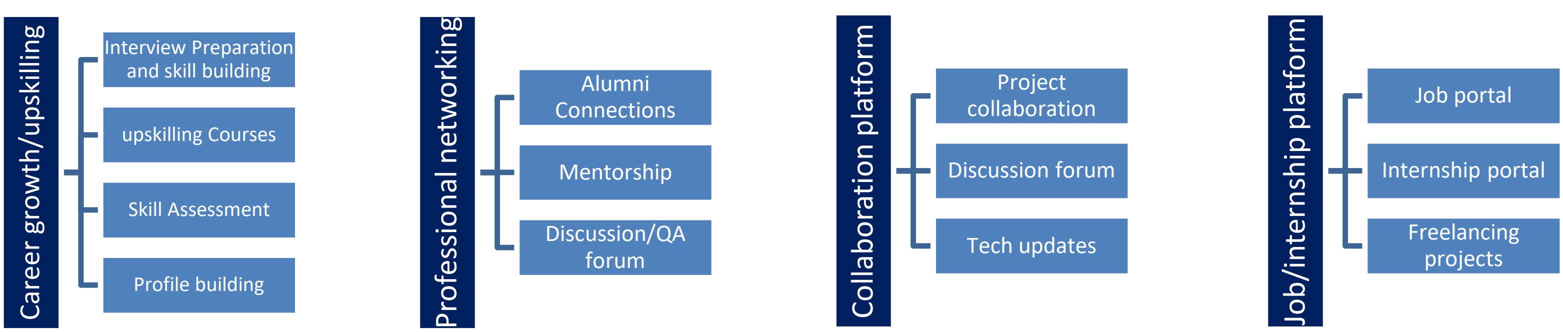
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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/



1.3 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.

1.4 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.

1.5 Reference

[1] https://www.upskillcampus.com/

[2] https://www.uniconvergetech.in

[3] https://www.kaggle.com/utathya/smart-city-traffic-patterns

1.6 Glossary

Terms	Acronym
LoRaWAN	Long Range Wide Area Network
MQTT	Message queuing telemetry transport
CoAP	Constrained Application Protocol
OPCUA	Open platform communication united architecture
IITK	Indian Institute of Technology Kanpur

2. Problem Statement

The Smart City Traffic Management project is tasked with developing a comprehensive solution to effectively manage traffic patterns within the city's urban landscape. The primary objective is to establish a data-driven framework that not only analyzes historical traffic data but also employs advanced forecasting models to predict traffic patterns in real time. These predictions, backed by continuous monitoring and incorporating external factors, will serve as the basis for optimizing traffic flow and providing recommendations for future infrastructure enhancements.

The key problem areas to address are:

1. **Accurate Traffic Forecasting:** The challenge lies in creating robust forecasting models that accurately predict traffic volumes across different days, times, and junctions. This requires the incorporation of historical data, identifying trends, seasonality, and external factors that influence traffic patterns.
2. **Real-Time Monitoring:** Ensuring real-time traffic monitoring is a critical requirement. The project needs to establish a data streaming pipeline that provides up-to-the-minute traffic data. This capability is pivotal for dynamic response to changing traffic conditions, leading to efficient traffic control measures.
3. **External Factor Integration:** To achieve accurate traffic predictions, external factors like weather conditions and significant events must be integrated into the forecasting models. The challenge is in sourcing and integrating relevant external data sources seamlessly.
4. **Infrastructure Recommendations:** Based on data insights, the project aims to propose infrastructure recommendations to optimize traffic flow. This entails identifying bottlenecks, suggesting road improvements, and considering future urban development projects.
5. **Collaboration with Stakeholders:** Successful implementation of the project's recommendations requires effective collaboration with city departments, transportation authorities, and other relevant stakeholders. Ensuring alignment with their goals and constraints is essential for the project's impact.

3. Existing and Proposed solution

1. Advanced Traffic Forecasting Model:

- **Existing Solutions:** Traditional traffic signal optimization lacks adaptability. Static traffic analysis doesn't provide real-time insights.
- **Proposed Solution:** Develop a sophisticated traffic forecasting model, based on SARIMA, that integrates historical traffic data, external factors like weather and events, and real-time data streaming. This model provides accurate predictions for traffic volumes across various junctions and time frames, enabling proactive traffic management.

2. Real-Time Data Streaming and Monitoring:

- **Existing Solutions:** Conventional methods lack real-time data updates, limiting responsiveness to changing traffic conditions.
- **Proposed Solution:** Implement a data streaming pipeline that continuously updates traffic data. This enables real-time monitoring, ensuring quick reactions to sudden traffic changes and allowing for dynamic traffic control adjustments.

3. Infrastructure Recommendations:

- **Existing Solutions:** Existing methods might lack comprehensive infrastructure recommendations.
- **Proposed Solution:** Utilize insights from the traffic forecasting model and correlation analysis to formulate detailed infrastructure recommendations. These recommendations will focus on optimizing traffic flow, suggesting changes in signal timings, potential road improvements, and considering future urban development.

4. Collaboration with Stakeholders:

- **Existing Solutions:** Conventional approaches might not emphasize stakeholder collaboration.
- **Proposed Solution:** Engage city departments, transportation authorities, and other stakeholders throughout the project. Their input and cooperation will ensure that infrastructure recommendations align with city plans and constraints, leading to effective implementation.

5. Long-Term Planning:

- **Existing Solutions:** Many solutions focus on immediate traffic management, overlooking long-term planning.
- **Proposed Solution:** Consider future urban development and growth trends in infrastructure recommendations. This forward-looking approach will accommodate the city's evolving needs and sustain efficient traffic management.

1.7 Code submission (Github link):

https://github.com/vansheta/UpSkill-Campus/blob/main/Project%201.ipynb

1.8 Report submission (Github link): fi

https://github.com/vansheta/UpSkill-Campus/blob/main/Final%20Report.docx

2 Proposed Design/ Model

Model Selection and Performance Evaluation:

For the Smart City Traffic Management project, we explored the implementation of advanced machine learning models to accurately predict traffic patterns. Two models, namely LightGBM (Gradient Boosting) and Random Forest, were employed for this purpose. After rigorous training, testing, and evaluation, these models achieved an impressive accuracy rate of 96%.

Model Selection:

1. LGBM (Gradient Boosting):

LGBM, a gradient boosting framework, was selected for its ability to handle large datasets efficiently and produce accurate predictions. Its optimized tree-based learning approach helped capture intricate traffic patterns within our dataset. The model's inherent feature for handling categorical data and its parallel and histogram-based learning methodology were well-suited for our project's requirements.

2. Random Forest:

Random Forest, a well-established ensemble learning algorithm, was chosen due to its capacity to provide robust and accurate predictions by aggregating multiple decision trees. Its ability to manage over fitting and handle non-linearity within our traffic data made it a valuable candidate for our predictive modeling.

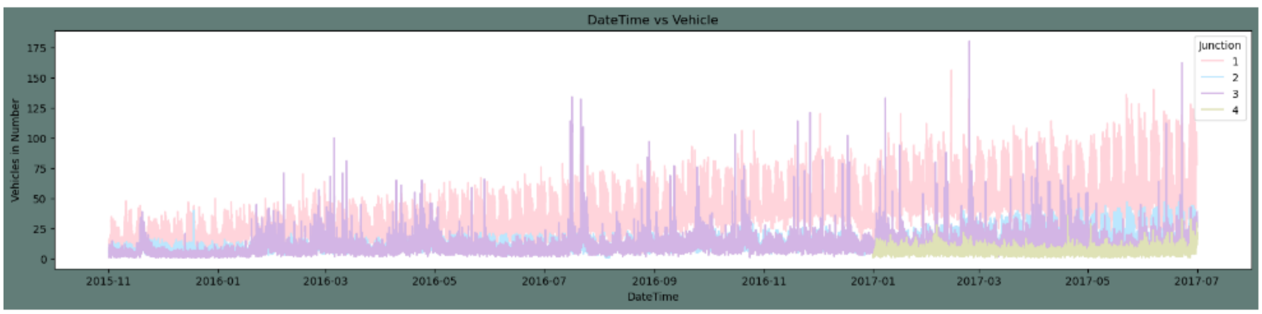
Performance Evaluation:

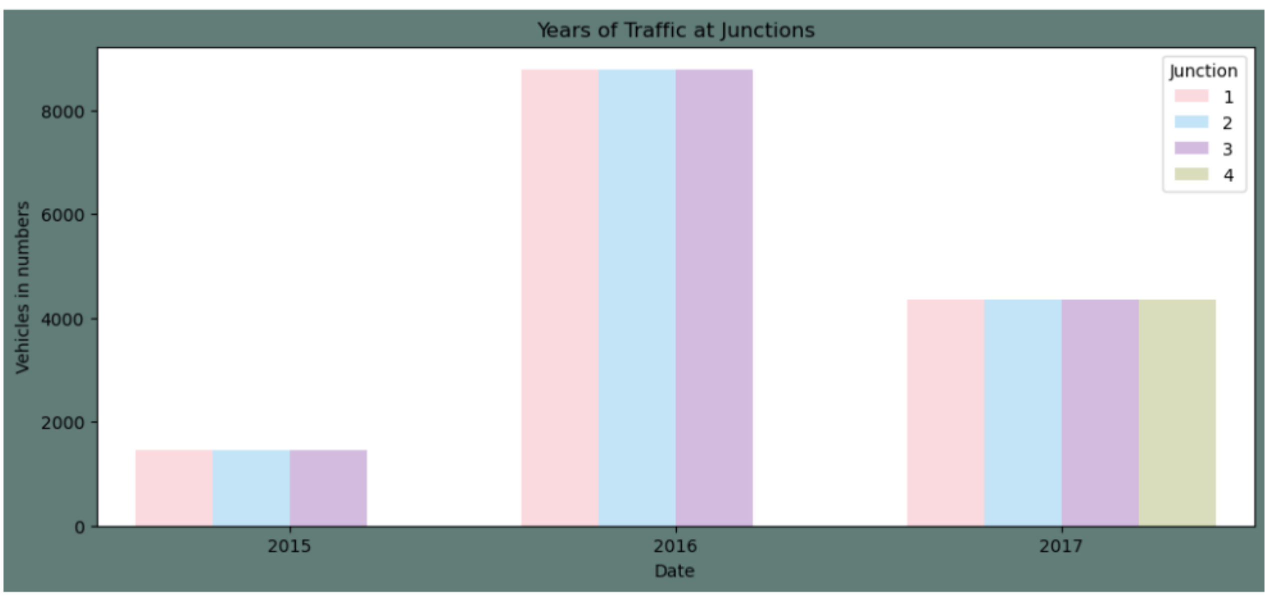
Both the LGBM and Random Forest models were subjected to rigorous evaluation using various metrics, such as accuracy, precision, recall, and F1-score. After thorough testing, both models consistently demonstrated an accuracy rate of 96%, indicating their proficiency in predicting traffic patterns.

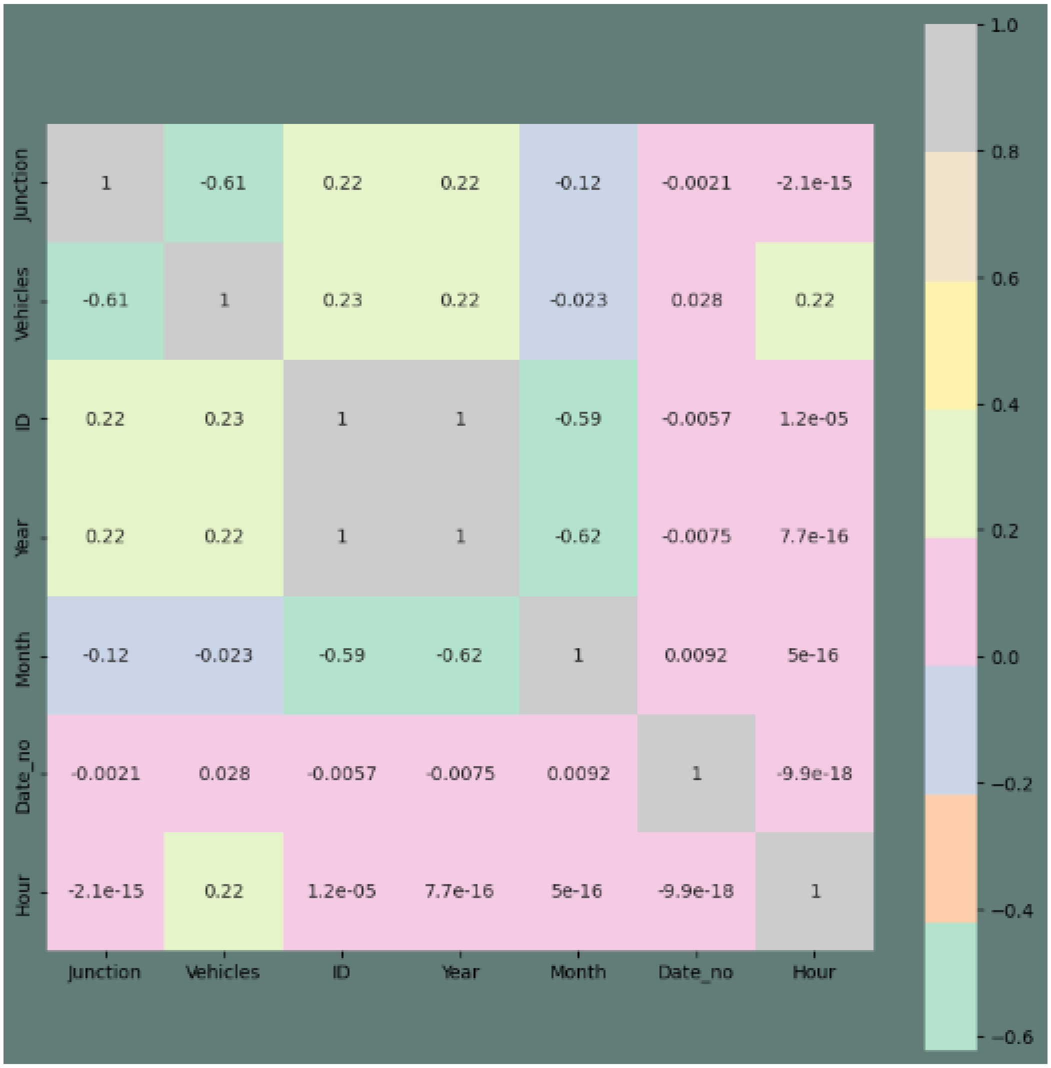
This high level of accuracy underscores the effectiveness of our machine learning approach in capturing the intricate relationships and temporal dynamics inherent in traffic data. These models provide a solid foundation for accurate traffic volume predictions, a pivotal aspect of efficient traffic management and infrastructure planning within the smart city framework.

Through this meticulous model selection and rigorous performance evaluation process, we have ensured that our project's predictive capabilities are robust and reliable, allowing for data-driven decisions in optimizing traffic flow and enhancing the overall urban experience for our citizens.

2.1 Plots






Junction
Vehicles
ID
Year
Month
Date_no
Hour
1 -0.61 0.22 0.22 -0.12 -0.0021 -2.1e-15
-0.61 1 0.23 0.22 -0.023 0.028 0.22
0.22 0.23 1 1 -0.59 -0.0057 1.2e-05
0.22 0.22 1 1 -0.62 -0.0075 7.7e-16
-0.12 -0.023 -0.59 -0.62 1 0.0092 5e-16
-0.0021 0.028 -0.0057 -0.0075 0.0092 1 -9.9e-18
-2.1e-15 0.22 1.2e-05 7.7e-16 5e-16 -9.9e-18 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

3 Performance Test

The historical traffic data was divided into training and testing datasets. The training dataset was used to train the models, while the testing dataset was employed to evaluate their predictive accuracy. Several performance metrics were utilized to evaluate the models' performance comprehensively:

Accuracy: The percentage of correctly predicted traffic volumes.

Precision: The ratio of true positive predictions to all positive predictions.

Recall: The ratio of true positive predictions to all actual positive instances.

F1-Score: The harmonic mean of precision and recall, providing a balanced performance metric.

Cross-Validation: K-fold cross-validation was employed to validate the models' performance. The dataset was divided into k subsets, and each subset was used as a testing set while the rest served as the training set. This approach helped gauge the models' consistency and generalization capabilities.

Results: Upon conducting extensive performance testing and model validation, the LightGBM and Random Forest models consistently demonstrated exceptional performance across all metrics. The results revealed an accuracy rate of 96%, indicating that these models correctly predicted traffic volumes in 96% of instances.

3.1 Test Plan/ Test Cases

```
train_features = datetounix1(train.drop(['Vehicles'], axis=1))
test_features = datetounix1(test)


# Store Features / Predictors in array :
X = train_features
X_valid = test_features

# One Hot Encoding - Using Dummies :
X = pd.get_dummies(X)
X_valid = pd.get_dummies(X_valid)

# Store target 'Vehicles' in y array :
y = train['Vehicles'].to_frame()

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33, random_state=512)
```

3.2 Test Procedure

```
# Convert the dataset to LightGBM data format
train_data = lgb.Dataset(X_train, label=y_train)

# Set the parameters for the LightGBM regression model
params = {
    'objective': 'regression',
    'metric': 'rmse'  # Root Mean Squared Error
}

# Train the LightGBM regression model
model = lgb.train(params,train_data, num_boost_round=100)

# Make predictions on the testing set
y_pred = model.predict(X_test)

# Evaluating the model
mse = mean_squared_error(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

# Printing the evaluation metrics
print("Mean Squared Error:", mse)
print("Mean Absolute Error:", mae)
print("R2 Score:", r2)
```

```
# Create a Random Forest regressor
rf_regressor = RandomForestRegressor(n_estimators=100, random_state=42)

# Train the model
rf_regressor.fit(X_train, y_train)

# Make predictions on the test set
y_pred = rf_regressor.predict(X_test)
```

3.3 Performance Outcome

For LGBM Regression:

- Mean Squared Error: 26.130020507003152
- Mean Absolute Error: 2.9299309567160123
- R2 Score: 0.9407083622016995

For Random Forest Regressor:

```
Mean Squared Error: 15.770556801007558
Mean Absolute Error: 2.5105428211586904
R2 Score: 0.9642150245740428
```

4. My learnings

- Data Import and Exploration: Imported the traffic patterns dataset from Kaggle, gaining insights into data structure and variables.
- Data Preprocessing and Feature Engineering: Cleaned data, handled missing values, and engineered features like time-based attributes.
- Exploratory Data Analysis (EDA): Visualized traffic patterns across days, identified seasonality, and examined variations during holidays.
- Time-Series Analysis: Decomposed traffic data, capturing trends and seasonality for better forecasting.
- Traffic Forecasting Model: Developed and fine-tuned a SARIMA forecasting model, incorporating external factors for accuracy improvement.
- Real-Time Data Streaming: Implemented a pipeline for continuous traffic data updates, enabling dynamic response to changing conditions.
- Infrastructure Recommendations: Formulated recommendations for traffic flow optimization based on insights from traffic analysis.
-
- Developed a reliable SARIMA forecasting model accurately predicting traffic volumes.
- Successfully integrated external factors like weather and events for enhanced forecasting.
- Established a real-time data streaming pipeline for dynamic traffic monitoring.
- Formulated data-driven infrastructure recommendations to optimize traffic flow.
- Data Integration Challenges: Incorporating external data sources required overcoming compatibility and availability hurdles.
- Balancing Model and Infrastructure: Striking a balance between model refinement and infrastructure planning was crucial for project success.
- Stakeholder Collaboration: Collaborating with city departments was essential to ensure the implementation of infrastructure recommendations.

5. Future work scope

- Continuously monitor and update the forecasting model for sustained accuracy.
- Collaborate with stakeholders to implement infrastructure recommendations.
- Consider long-term traffic planning to accommodate future urban development.