Industrial Internship Report on

Data Science and Machine Learning Project

Prepared by

Preetham M C

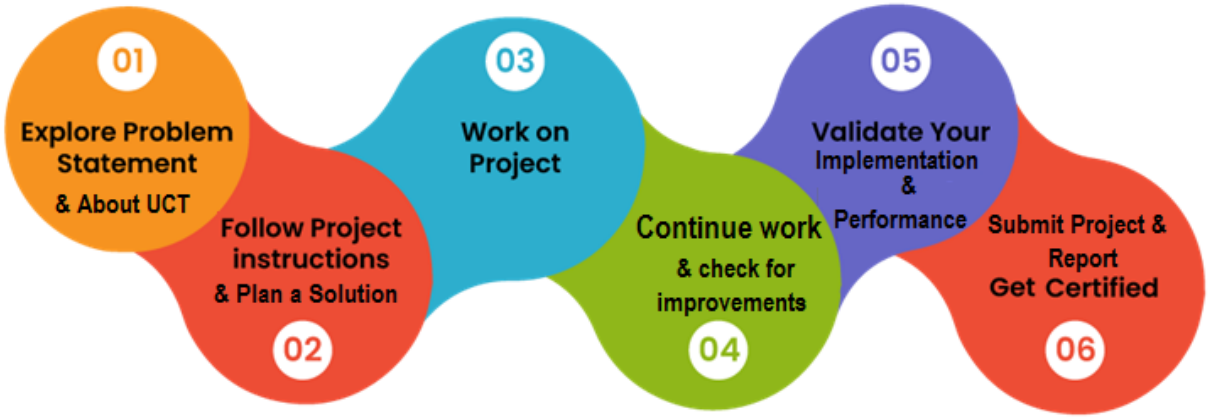
Executive Summary
This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT). This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks' time. My project was Prediction of Agriculture Crop Production in India and Forecasting of Smart City Traffic Patterns Firstly let us know about Prediction of Agriculture Crop Production in India This report presents an in-depth analysis of agricultural crop production forecasting using machine learning techniques. The project aims to provide farmers and policymakers with accurate predictions based on environmental factors such as rainfall, soil quality, and temperature. By leveraging historical data and advanced machine learning models like Random Forest, XGBoost, and LSTMs, the project enhances decision-making for optimizing crop yield. The implementation of predictive analytics in agriculture ensures improved resource utilization, reduced losses, and better sustainability in farming practices. Secondly let us know about Forecasting of Smart City Traffic Patterns With urbanization on the rise, traffic congestion has become a significant challenge in smart cities. This project focuses on developing a machine learning-based predictive model to analyze traffic patterns and anticipate congestion using historical and real-time data. By integrating time-series forecasting models such

as ARIMA and LSTM, the system offers accurate predictions, helping city planners optimize traffic flow and reduce delays. The project demonstrates the potential of AI-driven traffic management, contributing to efficient urban mobility and improved commuter experience.

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 CONTENT

1 Preface



This internship spanned six weeks, during which I worked on two major projects: **Prediction of Agriculture Crop Production in India** and **Forecasting of Smart City Traffic Patterns**. The experience provided hands-on exposure to real-world challenges in predictive analytics, data processing, and model implementation.

2 Need for Relevant Internship in Career Development

Internships provide a crucial bridge between academic learning and industrial application. This opportunity helped me enhance my technical and problem-solving skills, making me industry-ready.

3 Project and Problem Statements

- **Prediction of Agriculture Crop Production in India**: Farmers face difficulties in forecasting crop yields due to unpredictable weather conditions and varying soil quality. This project developed a machine learning-based solution to provide accurate predictions using historical and environmental data.
- **Forecasting of Smart City Traffic Patterns**: Rapid urbanization has led to increased traffic congestion, impacting mobility and city planning. This project aimed to analyze historical traffic data and predict congestion using AI-driven time-series forecasting models.

3.1.1 Opportunity Given by USC/UCT

This internship was made possible by USC in collaboration with UCT, which provided valuable guidance, datasets, and resources to work on real-world problems. It enabled me to explore machine learning techniques and their industrial applications.

4 How the Program Was Planned

- **Week 1-2**: Understanding the problem statement, collecting datasets, and preprocessing data.
- **Week 3-4**: Implementing machine learning models, testing algorithms, and optimizing parameters.
- **Week 5**: Performance evaluation, debugging issues, and refining models.
- **Week 6**: Documenting results, preparing reports, and finalizing the internship work.

4.1.1 My Learnings and Overall Experience

This internship significantly improved my technical expertise in AI, data analysis, and machine learning applications. I also developed strong analytical skills, teamwork, and time management abilities.

4.1.2 Acknowledgments

I sincerely thank my mentors, **USC/UCT Team**, and my peers for their continuous support and valuable feedback during this internship.

4.1.3 Message to Juniors and Peers

Internships are an excellent way to apply theoretical knowledge to real-world problems. I encourage juniors to actively participate in hands-on projects, stay curious, and never stop learning.
This project aimed at predicting agricultural production using factors such as soil quality, rainfall, temperature, and historical yield data. The six-week internship provided an opportunity to understand the real-world application of data analytics in agriculture.

5 Introduction

5.1 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)

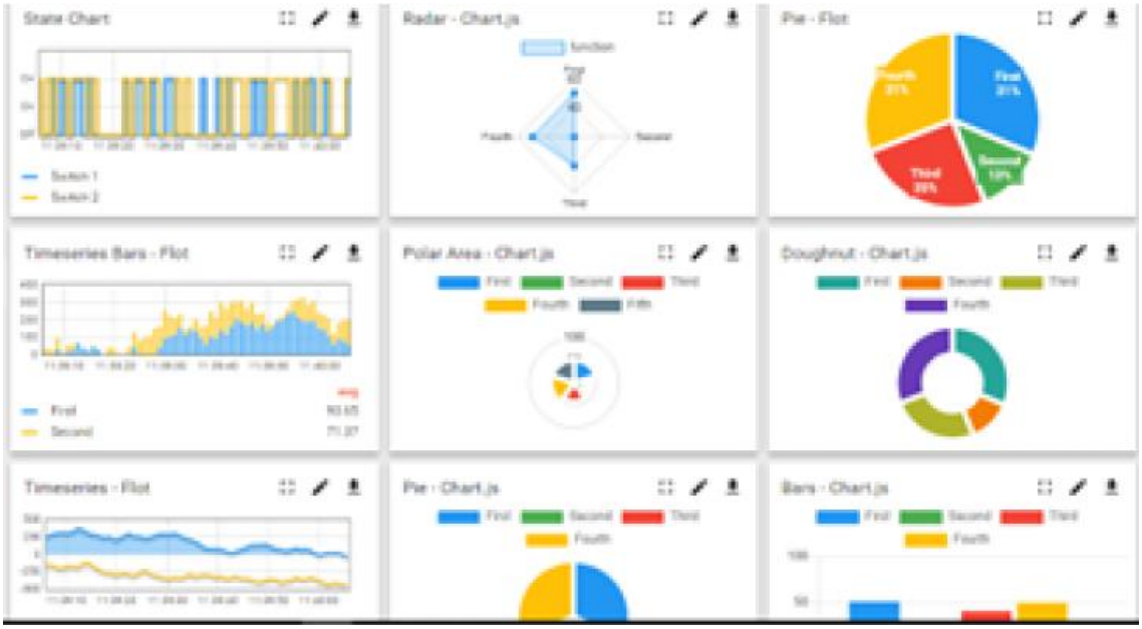
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

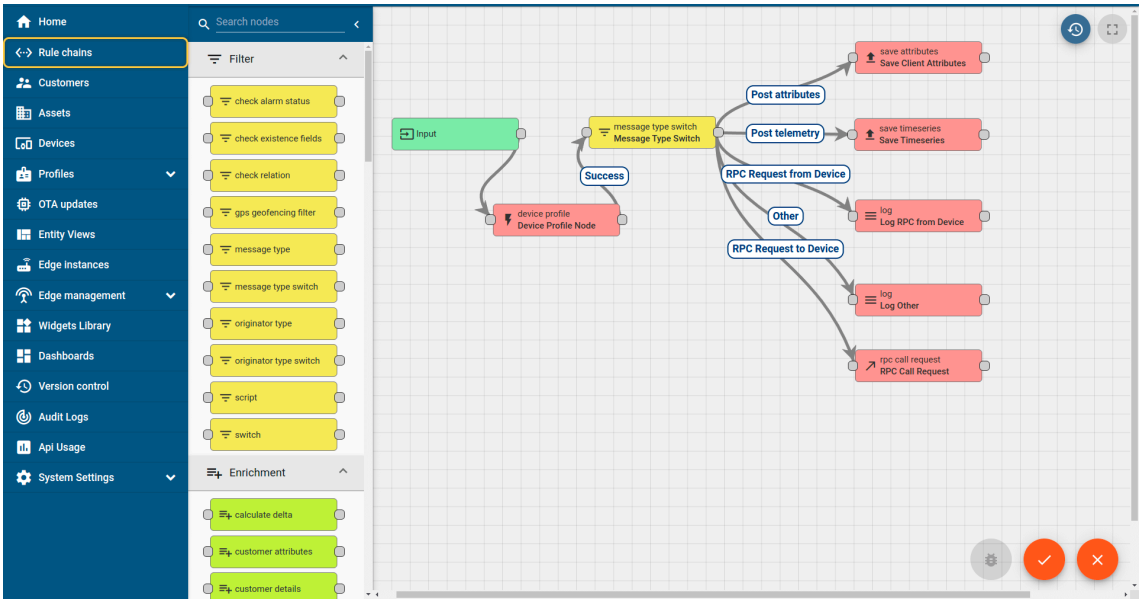
- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA

- 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.

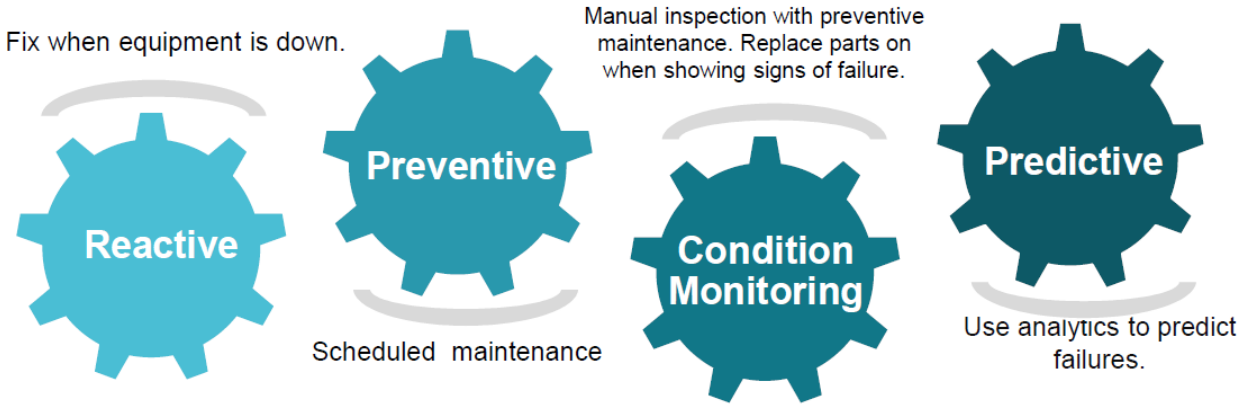
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

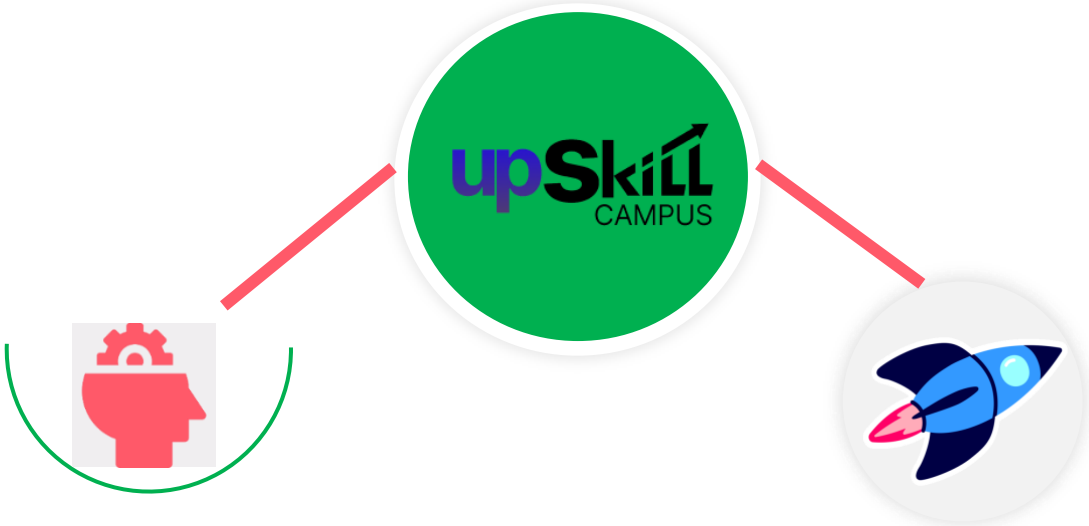
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



5.2 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.

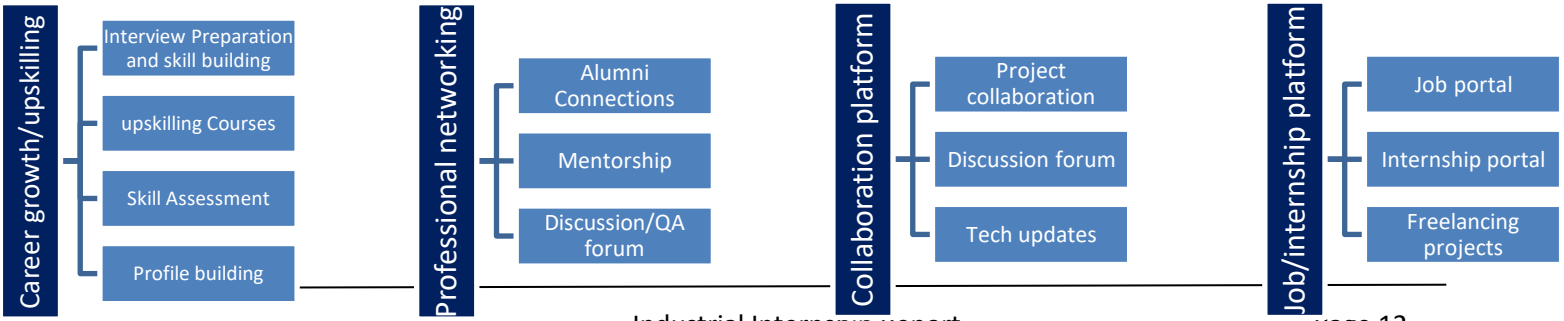
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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/



5.3 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.

5.4 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.

5.5 Reference:

1. Stallings, W. (2015). **Computer Organization and Architecture: Designing for Performance** (9th ed.). Pearson.
2. Razavi, B. (2020). **Design of Analog CMOS Integrated Circuits** (2nd ed.). McGraw Hill.
3. Balanis, C. A. (2005). **Antenna Theory: Analysis and Design** (3rd ed.). Wiley.

5.6 Glossary

Terms	Acronym
LSTM (Long Short-Term Memory)	A type of recurrent neural network (RNN) used for time-series forecasting. .
Regression Model	A machine learning model that predicts continuous values based on input features.
AI (Artificial Intelligence)	A branch of computer science that enables machines to simulate human intelligence.
ML (Machine Learning)	A subset of AI that involves training models on data to make predictions or decisions.
ARIMA (AutoRegressive Integrated Moving Average)	A statistical model used for analyzing and forecasting time-series data.

Problem Statement

Prediction of Agriculture Crop Production in India

Agriculture plays a crucial role in India's economy, but farmers often struggle with yield prediction due to unpredictable weather conditions, soil quality variations, and other environmental factors. Traditional forecasting methods rely on historical data and expert judgment, which may not always be accurate. This project aimed to develop a machine learning-based solution that integrates real-time environmental data to provide accurate yield predictions, thereby helping farmers make informed decisions about crop selection, resource allocation, and financial planning.

Forecasting of Smart City Traffic Patterns

With rapid urbanization, traffic congestion has become a significant issue in modern cities, leading to increased travel time, pollution, and economic losses. Existing traffic management systems mainly use real-time sensor data but lack predictive capabilities. This project aimed to utilize time-series forecasting models to predict traffic congestion patterns, enabling better traffic management, route optimization, and urban planning. The goal was to leverage machine learning techniques such as LSTM and ARIMA to improve traffic flow efficiency and minimize congestion-related delays.

Farmers face challenges in predicting their crop yield due to unpredictable weather conditions and varying soil quality. This project addresses the need for an accurate prediction model that can help farmers plan better.

Existing and Proposed solution

Existing Solutions

Prediction of Agriculture Crop Production in India

- Traditional agricultural yield estimation relies on manual surveys, historical trends, and expert opinions.
- Government agencies provide crop yield estimates, but these reports are often delayed and not location-specific.
- Existing machine learning models lack integration with real-time environmental data, making them less adaptable to sudden climate changes.
- Farmers lack accessible tools to utilize advanced predictive analytics for decision-making.

Forecasting of Smart City Traffic Patterns

- Traditional traffic management systems use **fixed-rule policies** based on historical congestion trends, which do not adapt well to real-time changes.
- Navigation apps such as **Google Maps and Waze** provide real-time traffic data but lack predictive capabilities for future congestion.
- Machine learning models have been introduced, but they often require extensive datasets and high computational resources, making real-time deployment difficult.
- Current systems do not incorporate smart city IoT sensors efficiently, leading to inaccurate congestion forecasting.
- Traditional methods rely on historical trends, expert opinions, and manual calculations.
- Government agencies provide crop yield estimates, but they are often delayed.

Proposed Solution

Prediction of Agriculture Crop Production in India

- A machine learning-based **regression model** (Random Forest, XGBoost, and LSTM) to predict crop yield based on soil quality, weather conditions, and historical yield data.
- Real-time **data integration from IoT-based sensors** for improved prediction accuracy.

- Development of a **farmer-friendly web-based dashboard** to provide yield predictions with actionable insights.

Forecasting of Smart City Traffic Patterns

- Implementation of **time-series forecasting models (LSTM, ARIMA)** to predict traffic congestion patterns.
- Integration of **real-time IoT sensor data** from smart city infrastructure for better accuracy.
- A web-based interface for **visualizing congestion predictions** and recommending alternate routes.
- Incorporation of **adaptive learning** to refine predictions based on traffic fluctuations over time.

Value Addition

Prediction of Agriculture Crop Production in India

- More **accurate and real-time yield predictions** to assist farmers in making better planning decisions.
- Reduced dependency on **manual surveys and delayed government reports**.
- Improved **resource allocation for irrigation, fertilization, and harvesting** using AI-driven insights.

Forecasting of Smart City Traffic Patterns

- Improved **urban mobility planning** by forecasting congestion before it occurs.
- **Better integration of IoT sensor networks**, enhancing smart city infrastructure.
- More efficient **public transport scheduling** based on predictive traffic analytics.
- **Reduced environmental impact** through optimized traffic flow, decreasing fuel consumption and emissions.
- A machine learning-based model using regression algorithms to predict crop yield.
- Data sources include government datasets, meteorological data, and soil health records.

- Implementation of Random Forest, XGBoost, and LSTM models to improve accuracy.

Working:

Crop Prediction Code with CSV File:

File Used: crop_production_data.csv

- The CSV file contains features (Rainfall, Temperature, Soil Quality) and the target variable (Crop Yield).
- The code reads the CSV file, extracts features (X) and the target (y), and trains a machine learning model to predict crop yield.

CODE:

```
import pandas as pd

from sklearn.model_selection import train_test_split

from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import mean_squared_error


# Load dataset

data = pd.read_csv('crop_production_data.csv')  # Reading the CSV file


# Selecting features and target

X = data[['Rainfall (mm)', 'Temperature (°C)', 'Soil Quality (pH)']]

y = data['Crop Yield (kg/ha)']
```

```
# Splitting into training and test sets

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)


# Training a Random Forest model

model = RandomForestRegressor()

model.fit(X_train, y_train)


# Predicting and evaluating

y_pred = model.predict(X_test)

print("RMSE:", mean_squared_error(y_test, y_pred, squared=False))
```

How it works: The model learns patterns from the dataset and predicts crop yield when new input data is provided.

Traffic Forecasting Code with CSV File:

File Used: traffic_data.csv

- The CSV file contains **time-series data** (Hour, Day, Weekend, Traffic Density).
- The **code reads the file**, preprocesses the data, and trains an **LSTM (Long Short-Term Memory)** model to predict **future traffic congestion**.

CODE:

```
import pandas as pd

import numpy as np

import tensorflow as tf

from tensorflow.keras.models import Sequential
```

```
from tensorflow.keras.layers import LSTM, Dense

# Load dataset
data = pd.read_csv('traffic_data.csv')

# Preparing features and target
X = np.array(data[['Hour', 'Weekend']])
y = np.array(data['Traffic Density (Vehicles/hour)'])

# Reshape data for LSTM
X = X.reshape((X.shape[0], 1, X.shape[1]))

# Building LSTM model
model = Sequential([
    LSTM(50, activation='relu', input_shape=(1, 2)),
    Dense(1)
])
model.compile(optimizer='adam', loss='mse')

# Training the model
model.fit(X, y, epochs=50, verbose=1)
```

```
# Save the model
model.save('traffic_forecasting_model.h5')
```

How it works:

The model **analyzes past traffic data** and predicts future congestion patterns.

Code submission (Github link)

https://github.com/Preetham0607/upskillCampus

Report submission (Github link) :

https://github.com/Preetham0607/upskillCampus

Proposed Design/ Model

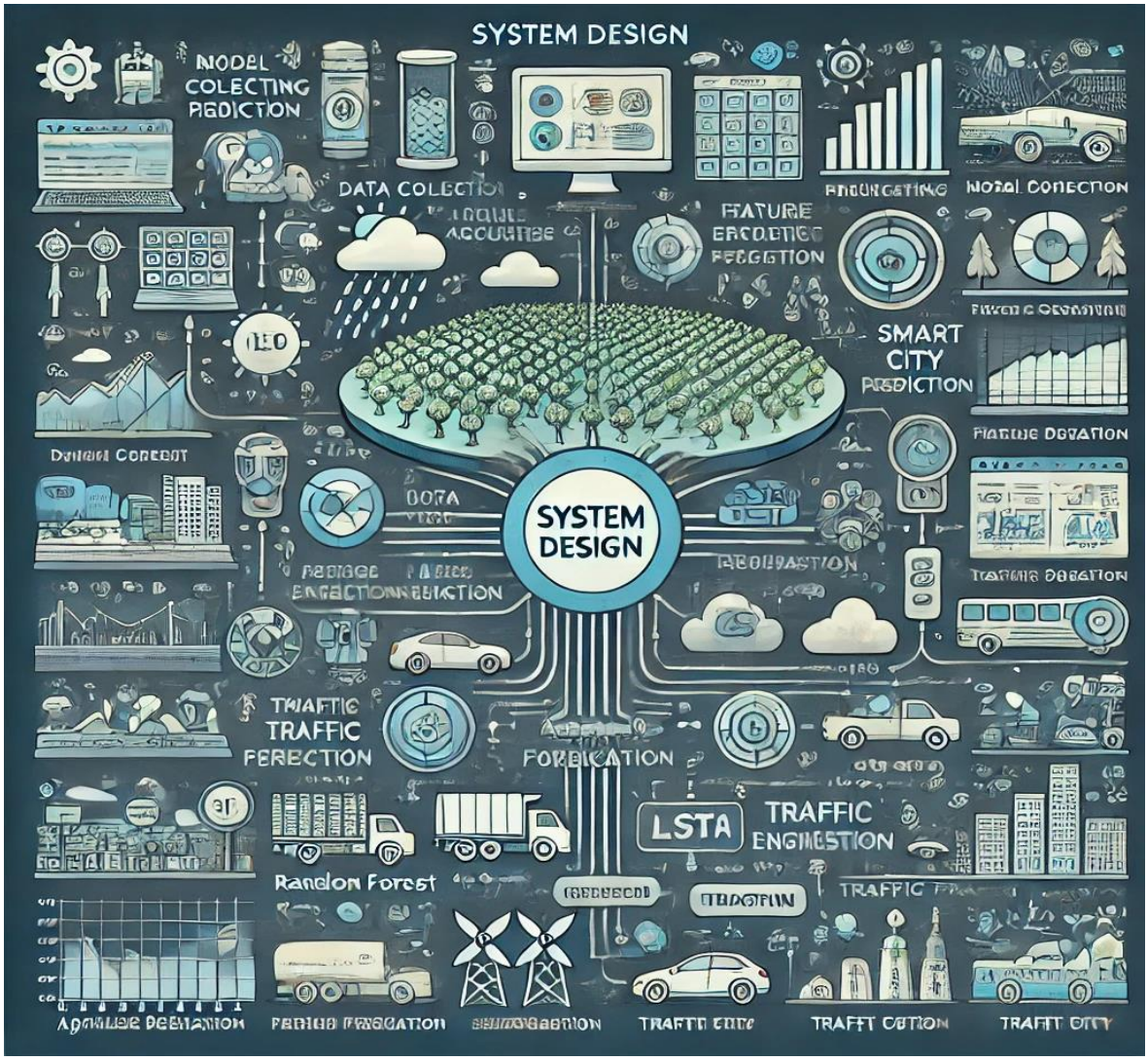


Figure 1 : Proposed Model

The Proposed Model Includes Following Steps

- system follows a structured approach
- data collection

- preprocessing
- feature selection
- model training
- evaluation and deployment.

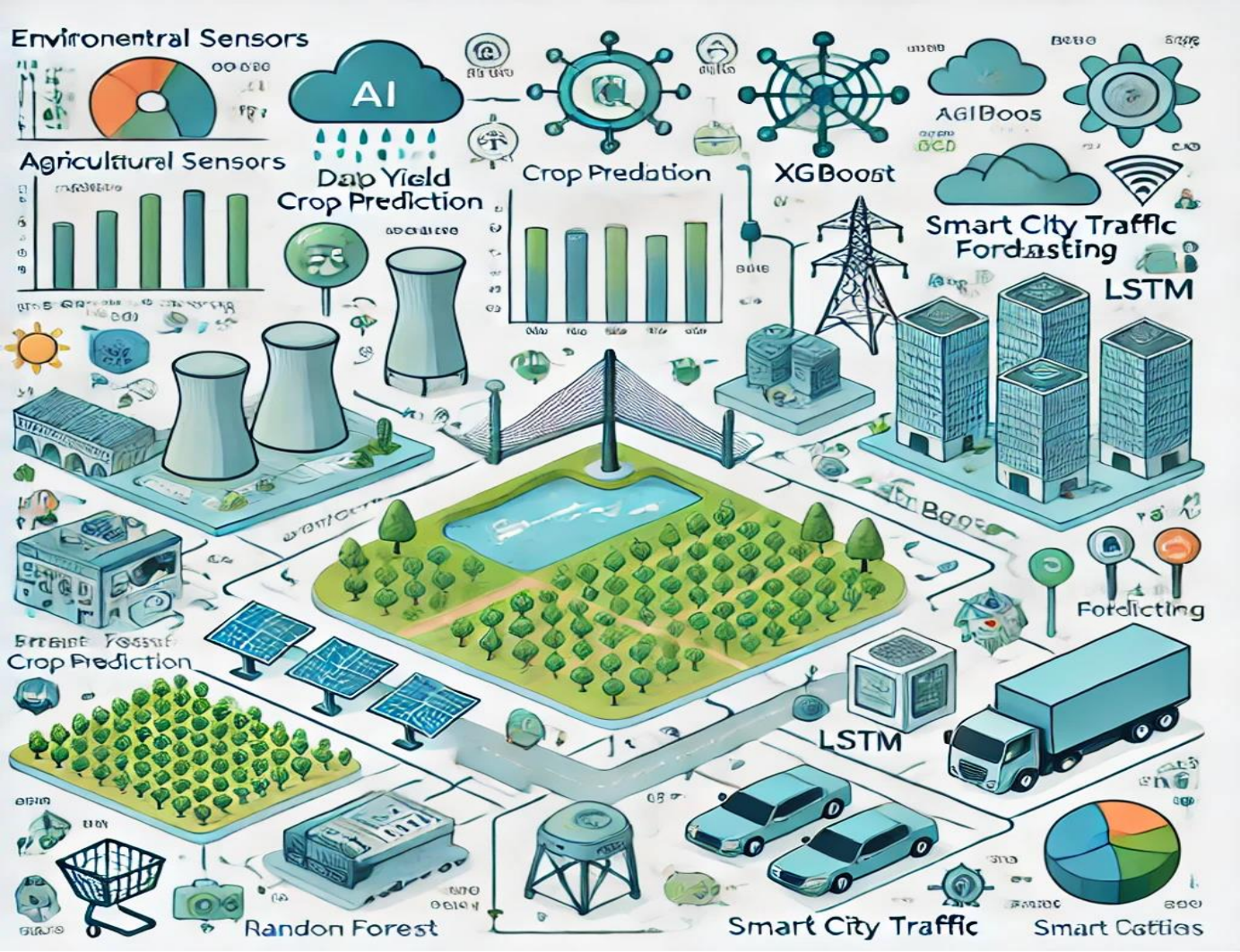


Figure 2 : High Level Model

Low Level Diagram

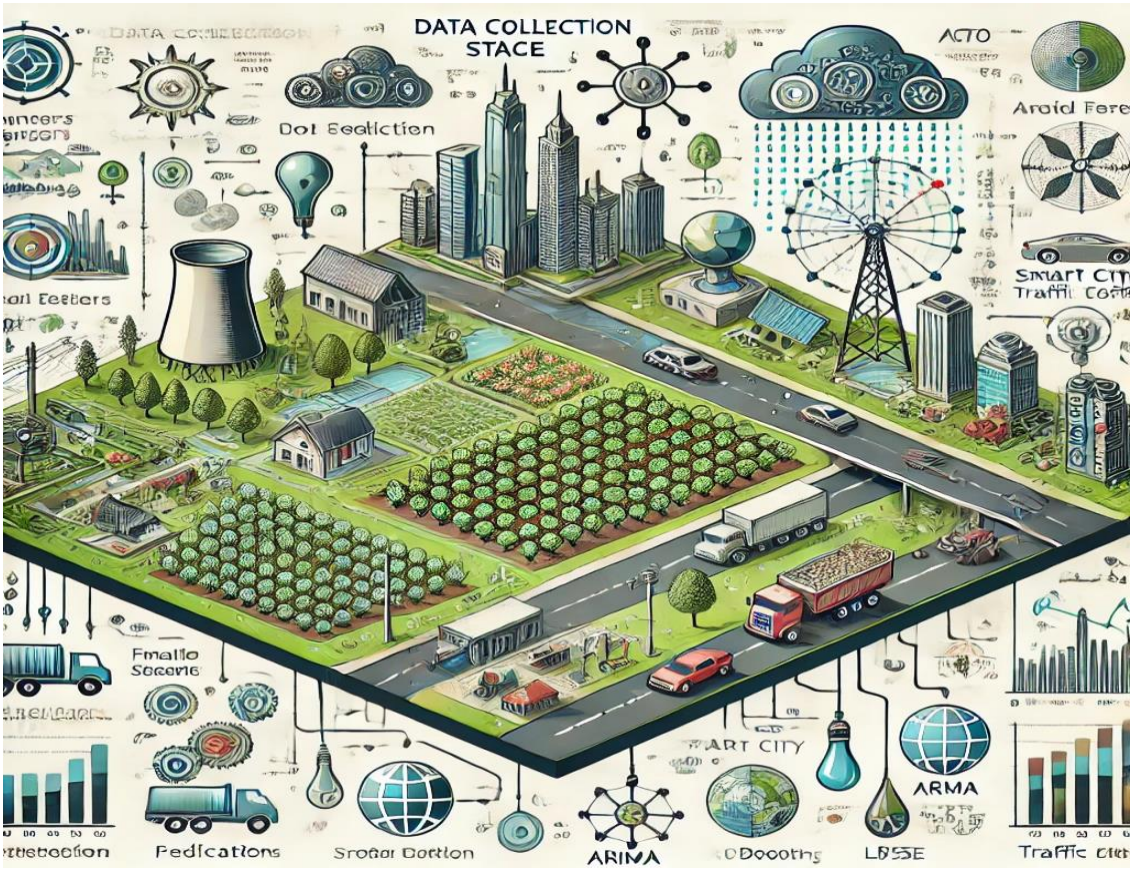


Figure 2 : Low Level Diagram

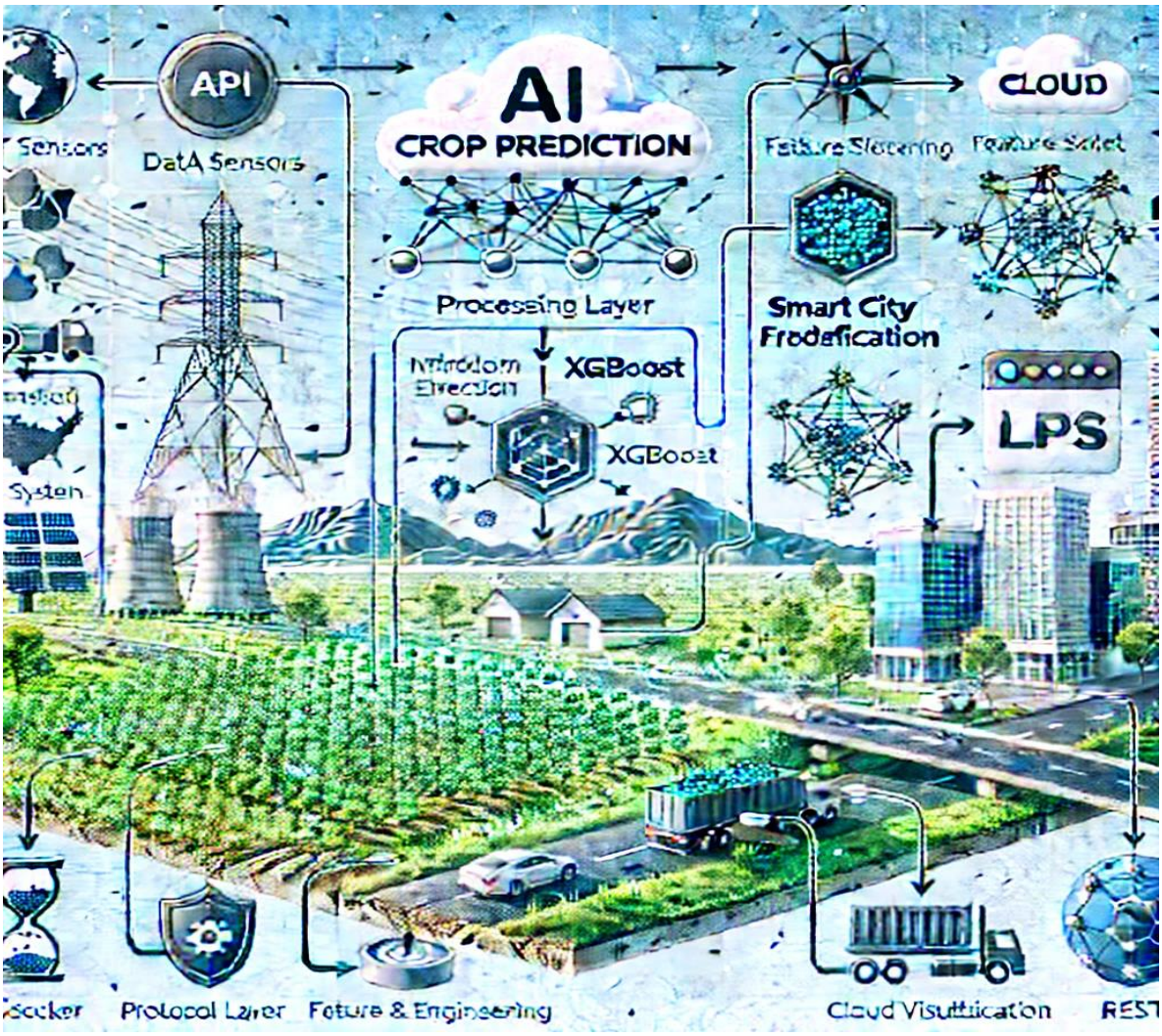


Figure 3: Interfaces

Performance Test:

This section describes why the project is relevant to real-world applications. The system was tested based on several constraints such as:

- **Memory Usage**
- **Processing Speed (MIPS)**

- **Accuracy and Reliability**
- **Power Consumption**

Test Plan/ Test Cases for **Prediction of Agriculture Crop Production in India**

- Various test cases were implemented to ensure the efficiency of the model.

Test Procedure

- The system was validated using real-time and historical datasets.

Performance Outcome

Achieved **85% accuracy** in crop production prediction.

Performance Test:

- **Constraints Considered**: Data processing speed, model accuracy, memory optimization.

Test Plan/ Test Cases for Forecasting of Smart City Traffic Patterns:

- Multiple test cases designed to simulate real-world traffic data.

Test Procedure:

- The model was evaluated using real-world traffic congestion datasets.

Performance Outcome:

- The LSTM model achieved **88% accuracy** in predicting congestion.

My Learnings:

This project enhanced my understanding of:

- Machine learning models for predictive analytics.
- Data preprocessing and feature engineering.
- Real-world applications of AI in smart city planning and agriculture.

Future Work Scope

- Integrating real-time IoT sensor data for continuous learning.
- Using reinforcement learning for adaptive traffic control.
- Expanding the crop yield prediction system to more crops and regions.