Project Report

Title :Smart City Traffic Pattern Forecasting using Machine Learning

---

Submitted By

Name: Sanket Nagnath Sutar
Domain: Data Science and Machine Learning
Internship Program: Upskill Campus
Industry Partner: UniConverge Technologies Pvt. Ltd.
Mode: Remote Internship

---

1. Introduction

With rapid urbanization, traffic congestion has become a major challenge for smart cities. Efficient traffic management requires accurate prediction of traffic patterns to optimize infrastructure planning and reduce congestion. This project focuses on forecasting traffic volume at different junctions using machine learning techniques to support smart city initiatives.

---

2. Company Overview

Upskill Campus, in collaboration with **UniConverge Technologies Pvt. Ltd.**, provides industry-oriented internship programs focused on developing practical skills in emerging technologies like Data Science and Machine Learning. The internship emphasizes hands-on project experience aligned with real-world problems.

---

3. Problem Statement

The objective of this project is to:

- Analyze traffic patterns at multiple city junctions.
- Predict future traffic volume using historical data.
- Provide insights that can help government authorities plan better traffic management strategies, especially during peak hours and special occasions.

---

4. Dataset Description

The dataset contains traffic information collected from four different junctions in a city.

Key Attributes

- **DateTime:** Timestamp of traffic observation
- **Junction:** Junction number (1 to 4)
- **Vehicles:** Number of vehicles observed
- **ID:** Unique record identifier

The dataset includes separate training and testing files.

---

5. Methodology

5.1 Data Preprocessing

- Removed inconsistencies and handled missing values.
- Converted DateTime into meaningful features:
    - Hour
    - Day
    - Weekday
    - Weekend indicator

5.2 Feature Engineering

- Extracted temporal features influencing traffic flow.
- Prepared cleaned datasets for training and testing.

5.3 Model Training

- Trained separate machine learning models for each junction.
- Used supervised learning techniques for traffic prediction.
- Saved trained models for future inference.

5.4 Prediction & Evaluation

- Compared actual traffic volume with predicted values.
- Analyzed model performance visually.

---

6. Tools and Technologies Used

- **Programming Language:** Python
- **Libraries:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn
- **Model Storage:** Joblib
- **Dashboard:** Streamlit
- **Version Control:** Git & GitHub

---

7. Results and Analysis

- Models successfully captured traffic trends across different junctions.
- Peak traffic hours and weekday patterns were clearly identified.
- Predictions closely matched actual traffic values, indicating good model performance.

---

8. Dashboard Development

An interactive **Streamlit dashboard** was developed to visualize:

- Traffic trends over time
- Hourly and weekday traffic patterns
- Actual vs Predicted traffic volume
- Key performance indicators (KPIs)

Live Dashboard Link

👉 https://smartcity-traffic-forecasting-c6igvxwplcx7h9k6gh7z8e.streamlit.app/

---

9. GitHub Repository

All project files including source code, models, and reports are maintained on GitHub.

📎 **GitHub Link:**
👉 https://github.com/SanketSutar075/smartcity-traffic-forecasting

---

10. Challenges Faced

- Feature mismatch between training and prediction stages.
- Handling time-based data efficiently.
- Managing large model files during deployment.

Solutions Implemented

- Ensured consistent feature engineering pipeline.
- Verified feature alignment before prediction.
- Used structured folder organization for deployment.

---

11. Key Learnings

- Importance of feature engineering in machine learning.
- Real-world data requires extensive preprocessing.
- Visualization plays a critical role in communicating insights.
- End-to-end ML pipeline development is essential for production-ready solutions.

---

12. Conclusion

The Smart City Traffic Pattern Forecasting project successfully demonstrates the application of machine learning techniques to solve a real-world urban problem. The project provides meaningful insights into traffic behavior and can assist authorities in making data-driven decisions for traffic management.

---

13. Future Scope

- Incorporating weather and event-based data.
- Using advanced models like LSTM for time-series forecasting.
- Deploying the solution with real-time data integration.

---

14. Declaration

I hereby declare that this project is an original work carried out by me during my internship under the guidance of Upskill Campus and UniConverge Technologies Pvt. Ltd.

Submitted By

Sanket Nagnath Sutar
Data Science and Machine Learning Intern