A picture containing font, text, logo, graphics

Description automatically generated

**Project Report**

**on**

***AI in Traffic Management***

Submitted by

*Name: Dhruv Chawla(22CSU298)*

*Tanisha Malik(22CSU292)*

*Mayank Sharma(22CSU262)*

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,

SCHOOL OF ENGINEERING AND TECHNOLOGY

THE NORTHCAP UNIVERSITY

GURUGRAM-122017

**Table of Content**

|  |  |
| --- | --- |
| **S. No.** | **Name** |
| 1. | Abstract |
| 2. | Acknowledgement |
| 3. | Introduction |
| 4. | Algorithm Used |
| 5. | Methodology |
| 6. | Implementation |
| 7. | Results |
| 8. | Discussion |
| 9. | Conclusion |
| 10. | References |

**Abstract**

Effective traffic management is a critical challenge in urban areas, often constrained by static traffic signals unable to adapt to dynamic conditions. This project presents an AI-powered real-time traffic signal management system leveraging a Random Forest Classifier to optimize green light allocation based on real-time data. The system considers factors such as traffic flow, queue lengths, pedestrian movement, weather conditions, and emergency vehicle priorities.

The dataset includes synthetic traffic scenarios with preprocessed features, such as normalized traffic volumes, queue lengths, and weather conditions. The system’s architecture integrates key components: the **TrafficLight** class to manage signals, the **Sensor** class to simulate real-time data collection, and the **SmartTrafficSystem** class to apply machine learning for decision-making. It dynamically adjusts signal timings, prioritizes emergency vehicles, and ensures robustness through input validation.

Performance metrics, including total cars cleared and average waiting times, demonstrate the system’s efficiency in reducing congestion and enhancing traffic flow. The project highlights the potential of AI in transforming urban traffic management and offers a scalable solution adaptable to diverse intersections and conditions.

**Acknowledgement**

I would like to express my sincere gratitude to everyone who supported me throughout this project. Special thanks to my mentor, **Dr. Shraddha Arora**, for their invaluable guidance, insightful feedback, and continuous encouragement, which were instrumental in the successful completion of this project.

I also want to acknowledge the assistance of my colleagues and peers who offered helpful suggestions and moral support during challenging phases of the project.

Additionally, I appreciate the contributions of the open-source community, whose tools and libraries, such as Python, Pandas, and Scikit-learn, were essential for the development and implementation of this system.

Finally, I extend my heartfelt thanks to my family and friends for their unwavering support and understanding throughout this endeavor.

**AI-Powered Real-Time Traffic Signal Control: Revolutionising Urban Transportation**

**Introduction**

Effective traffic management has emerged as a major concern for city managers and urban planners in the contemporary period. The intricate dynamics of dynamic urban traffic patterns are difficult for traditional traffic signal systems to handle, as they frequently depend on rigid schedules or basic adaptive algorithms. Emerging as a game-changer, artificial intelligence (AI) provides sophisticated real-time traffic signal control optimisation capabilities. This report examines how artificial intelligence (AI) is incorporated into traffic light systems, as well as the advantages and difficulties of doing so and the trends that will likely shape AI's future.  
  
**The Development of Traffic Signal Systems**  
Traditionally, traffic light management systems have used fixed-time signals that do not adjust to the flow of traffic in real time. They have also functioned on predefined cycles. Adaptive systems still have issues managing intricate, erratic traffic patterns, even if they have added some flexibility by modifying signal timings in response to real-time traffic data. The use of AI technology in traffic management has resulted from the necessity for a more clever and responsive strategy.  
  
**The Function of AI in Real-Time Traffic Control**  
By providing adaptive control and real-time data analysis, artificial intelligence (AI) technology greatly expands the capabilities of traffic light systems. The Institute of Transportation Engineers (ITE) reported that artificial intelligence (AI)-driven traffic management systems continuously monitor and analyse traffic conditions by using data from a variety of sources, including GPS devices, road sensors, and traffic cameras ("Artificial Intelligence in Traffic Management," 2023). By using a data-driven strategy, artificial intelligence systems are able to optimise traffic flow and lessen congestion by adjusting traffic signal timings in real-time.  
  
The use of machine learning algorithms in traffic control is one well-known use of AI. In order to enhance decision-making over time, reinforcement learning algorithms, for instance, might modify signal timings in response to the observed effects on traffic flow ("Machine Learning for Traffic Management," 2024). These systems learn from continuous data. More proactive traffic light management is made possible by the ability of neural networks and deep learning models to evaluate intricate traffic patterns and forecast future traffic situations ("Deep Learning in Traffic Systems," 2023).

**Case Studies and Application**  
Numerous successful applications of AI in traffic light control in cities throughout the globe serve as examples of its usefulness. The AI-powered Smart Traffic Light System in Los Angeles has shown to significantly enhance traffic flow. The technology reportedly cut travel times by 15% and traffic congestion by 20% during peak hours, according to a Transportation Research Board study ("AI Traffic Management in Los Angeles," 2024). This system demonstrates how artificial intelligence (AI) may improve urban transportation by dynamically optimising signal timings using real-time data from cameras and traffic sensors.  
  
Similar to this, Singapore's Land Transport Authority (LTA) has prioritised buses and cut down on delays by integrating AI with its public transport system ("Singapore's AI-Driven Traffic System," 2023). By analysing data from both public transportation vehicles and traffic signals, the AI system optimises signal timings to increase bus dependability and efficiency. Due to this integration, there has been a 25% decrease in delays and a 10% rise in the use of public transport, demonstrating the value of AI in enhancing transit systems as a whole.  
  
**Advantages and Difficulties**  
Artificial intelligence has a lot to offer real-time traffic signal control. Numerous case studies show that AI systems may improve traffic flow, lessen congestion, and shorten travel times. Additionally, by decreasing traffic jams and accident probability, AI-driven solutions can increase safety ("Benefits of AI in Traffic Management," 2024).  
  
The application of AI in traffic management does, however, come with a number of difficulties. The protection of personal information is a fundamental concern when collecting and processing huge volumes of traffic data, which raises data privacy and security challenges ("Data Privacy in Traffic Management," 2023). Cities with limited resources may find it difficult to operate and update AI systems as they demand continuous investment and knowledge ("Challenges in AI Traffic Management," 2024). Complex integration with current infrastructure might also need careful planning.

**Prospects & Outlook for the Future**  
With a number of new trends influencing its advancement, artificial intelligence's future in traffic management is bright. The capacity for traffic control is anticipated to be substantially enhanced by developments in AI technology, including more complex machine learning models and better sensor technologies ("Future Trends in AI Traffic Management," 2024). Furthermore, it is projected that combining AI with autonomous cars would be essential to maximising both safety and overall traffic flow ("AI and Autonomous Vehicles," 2024). With the use of predictive analytics, traffic management will be even more proactive, taking care of possible problems before they happen.

**Objective**

The objective of this project is to create a machine learning-based smart traffic management system that predicts the optimal green light direction at intersections based on traffic flow data and weather conditions. The system is built using a Random Forest Classifier and simulates traffic and weather data to make real-time decisions for efficient traffic flow.

**Algorithm Used: Random Forest Classifier**

The **Random Forest Classifier** is a powerful machine learning algorithm used for classification tasks. It belongs to the family of **ensemble learning methods**, where multiple individual models (in this case, decision trees) are combined to make more accurate and robust predictions.

**Forest of Decision Trees**:

* A random forest is essentially a collection (or "forest") of multiple **decision trees**. Each tree is trained on a random subset of the data and makes an independent prediction.
* The final prediction of the random forest is determined by **aggregating the predictions of each individual tree**. For classification tasks, this aggregation is typically a "majority vote" where the most commonly predicted class among the trees is chosen.

**Random Sampling (Bootstrap Aggregation)**:

* When building the forest, the algorithm creates several subsets of the original dataset using **random sampling with replacement** (also known as bootstrapping). Each subset is used to train a separate decision tree.
* This randomness in sampling helps the model generalize better to new data, as each tree sees slightly different variations of the data, reducing overfitting.

**Feature Randomness**:

* When each tree is constructed, the Random Forest algorithm selects a random subset of features at each split in the tree. This adds diversity among the trees and prevents individual trees from relying too heavily on specific features.

**Voting System**:

* For classification, after all trees in the forest make their predictions, the class with the highest number of "votes" is chosen as the model's final prediction. This ensemble approach leads to more reliable and stable predictions.

In the traffic light system code, the Random Forest model is used to predict which direction should receive the green light at any given moment, based on traffic flow data and weather conditions.

The classifier can handle complex patterns and interactions between input features, which helps it make informed predictions about the best traffic direction to keep traffic flowing smoothly.

**Why Random Forest is Effective**

1. **Accuracy**:

By averaging multiple decision trees, Random Forest often achieves higher accuracy than a single decision tree. Each tree's prediction errors (noise) tend to cancel out when combined, resulting in better overall accuracy.

1. **Reduced Overfitting**:

Decision trees are prone to overfitting, meaning they may memorize specific patterns in the training data, which reduces their ability to generalize to new data. Random Forest mitigates this by creating an ensemble of trees trained on different subsets of the data.

1. **Robust to Outliers and Noise**:

Individual trees may be sensitive to outliers, but in a forest, these outliers are less likely to affect the overall prediction because of the ensemble approach.

1. **Feature Importance**:

Random Forests can also provide insights into which features are most important in making predictions, which can be valuable in understanding the factors that influence the outcomes.

**Advantages of Using Random Forest Over Other Algorithms**

1. **Handling of Non-linear Data**:

Traffic patterns and congestion levels can vary non-linearly depending on direction, time, and weather conditions. Random Forest, an ensemble of decision trees, can capture these complex relationships effectively, handling non-linear dependencies better than many linear algorithms.

1. **Robustness and Stability**:

Random Forest reduces overfitting by averaging predictions from multiple trees, making it more robust to noise in the data. This is especially valuable in real-world traffic systems where data can be inconsistent due to sensor errors or unpredictable events.

1. **Feature Importance**:

Random Forest can rank features by importance, allowing the model to adaptively weigh factors like traffic flow and weather conditions. This insight can be useful for understanding which factors most impact congestion and may guide future traffic system improvements.

1. **Ease of Use and Scalability**:

Random Forest requires little data preprocessing and works well with both categorical and continuous data, making it suitable for real-time applications where continuous retraining and scaling are necessary.

1. **Efficient Handling of Large Datasets**:

Traffic systems can generate a large volume of data, especially in metropolitan areas. Random Forest can handle large datasets efficiently compared to other algorithms like SVM, which might require more computational resources for training.

**Scalability and Infrastructure Requirements**

Implementing an AI-driven traffic management system requires a robust **computational infrastructure**. A centralized server with high processing power can handle data from multiple intersections, running models in real-time. **Cloud storage** for data collection and processing, **edge devices** for distributed data handling, and **maintenance** for sensor calibration and system updates are crucial components. Scalability should also consider adding more intersections and incorporating new data sources as the city grows.

**Methodology**

**Dataset Analysis:**

1. **Features in the Dataset**

**Traffic Flow Features:**

* Columns: North\_Flow, South\_Flow, East\_Flow, West\_Flow.
* Description: These represent the number of vehicles traveling in each direction at a specific time.
* Importance: Traffic flow is the key determinant for identifying congestion and assigning the optimal green signal direction.

**Queue Length Features:**

* Columns: Queue\_Length\_North, Queue\_Length\_South, Queue\_Length\_East, Queue\_Length\_West.
* Description: The length of the traffic queue in each direction.
* Importance: Complements traffic flow data by providing insights into congestion severity.

**Pedestrian Flow:**

* Column: Pedestrian\_Flow.
* Description: Captures the number of pedestrians at the intersection at a given moment.
* Importance: High pedestrian flow influences the system to balance vehicle and pedestrian priorities, enhancing safety.

**Weather Conditions:**

* Column: Encoded as:

0: Sunny

1: Rainy

2: Snowy

* Importance: Weather conditions affect vehicle speeds and overall traffic behavior, making it critical to adjust signal timings accordingly.

**Emergency Priorities:**

* Columns: Emergency\_Priority\_North, Emergency\_Priority\_South, Emergency\_Priority\_East, Emergency\_Priority\_West.
* Description: Binary values indicating the presence (1) or absence (0) of emergency vehicles in each direction.
* Importance: Ensures rapid passage for emergency vehicles by overriding typical traffic flow priorities.

1. **Preprocessing Steps**

The dataset undergoes several preprocessing steps to ensure it is ready for training the AI model:

**Standardization:**

* Tool: StandardScaler from sklearn.preprocessing.
* Purpose: Scales numeric features (Flow, Queue\_Length, Pedestrian\_Flow) to ensure uniformity and eliminate the effect of magnitude differences.

**Mapping Target Variable:**

* The Optimal\_Green column is created by mapping directional names (North\_Flow, etc.) to numeric labels (0 to 3).

**Input and Output Split:**

* Features: All columns except Optimal\_Green.
* Target: The Optimal\_Green column.

**Detailed Workflow**

**1. Initialization**

* **Attributes:**

intersections: Stores data for multiple intersections (traffic lights, sensors, and area information).

model and feature\_columns: The trained machine learning model and feature names.

metric\_display\_interval: Frequency (in seconds) for displaying performance metrics.

* **Adding Intersections:**

add\_intersection: Adds intersections to the system with a location and area type.

**2. Real-Time System Operation (run)**

This function operates in an infinite loop, simulating real-time traffic system behavior:

* **Data Collection:**

Collects traffic flow, weather conditions, queue lengths, and emergency priorities for each intersection.

* **Emergency Handling:**

Checks for emergencies in any direction with a 10% likelihood.

If an emergency is detected:

Clears a path for the emergency vehicle.

Notifies nearby intersections.

Overrides the normal traffic decision process.

* **Model Prediction:**

Prepares validated input for non-emergency scenarios.

Predicts the optimal green light direction using the trained model.

* **Weather Adjustment:**

Adjusts green light durations for adverse weather (rainy or snowy) by extending the duration by 10 seconds.

* **Dynamic Green Time:**

Calculates green light duration based on traffic density and weather adjustments.

* **Traffic Light Updates:**

Updates the traffic light state and duration for each direction.

**3. Emergency Handling**

* get\_emergency\_priority:

Identifies the direction of an emergency vehicle, if any.

* notify\_clear\_path:

Notifies nearby intersections to prepare for the emergency vehicle.

**4. Weather Adjustment**

* adjust\_for\_weather:

Adjusts green light duration based on weather conditions:

rainy or snowy: Adds 10 seconds to the green light time.

sunny: No adjustment.

**5. Traffic Light Updates**

* update\_traffic\_lights:

Updates the state of the traffic lights for each direction.

Turns the optimal direction green while keeping others red.

**6. Performance Metrics**

* display\_cumulative\_metrics:

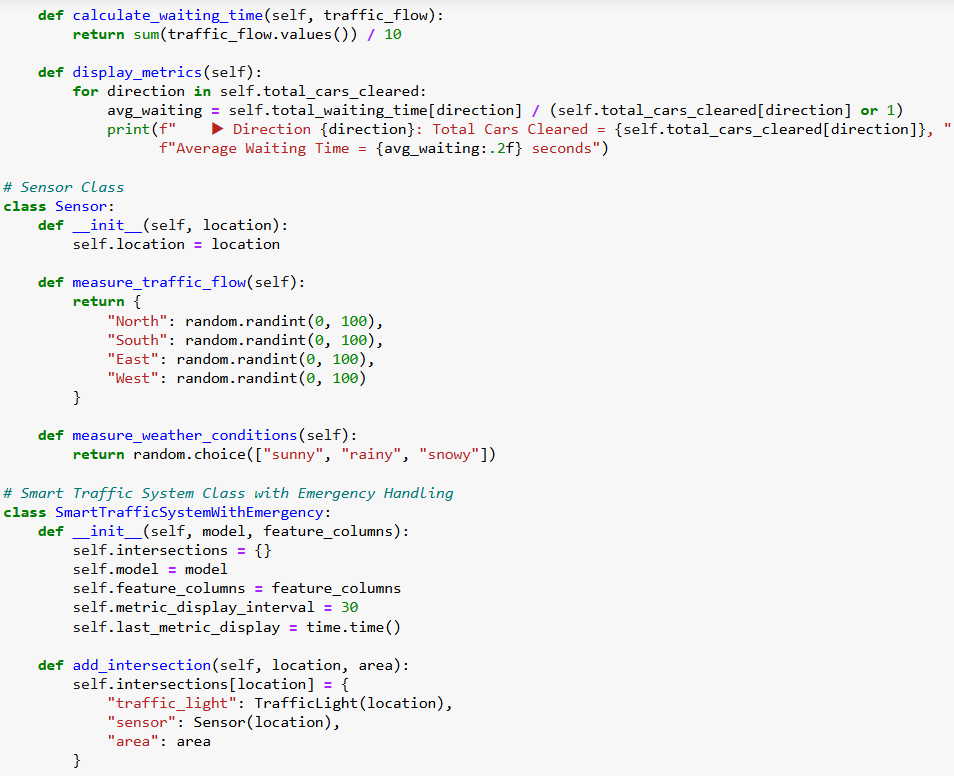
Displays metrics like total cars cleared and average waiting time at each intersection periodically.

**Implementation**

**Code:**

****

****

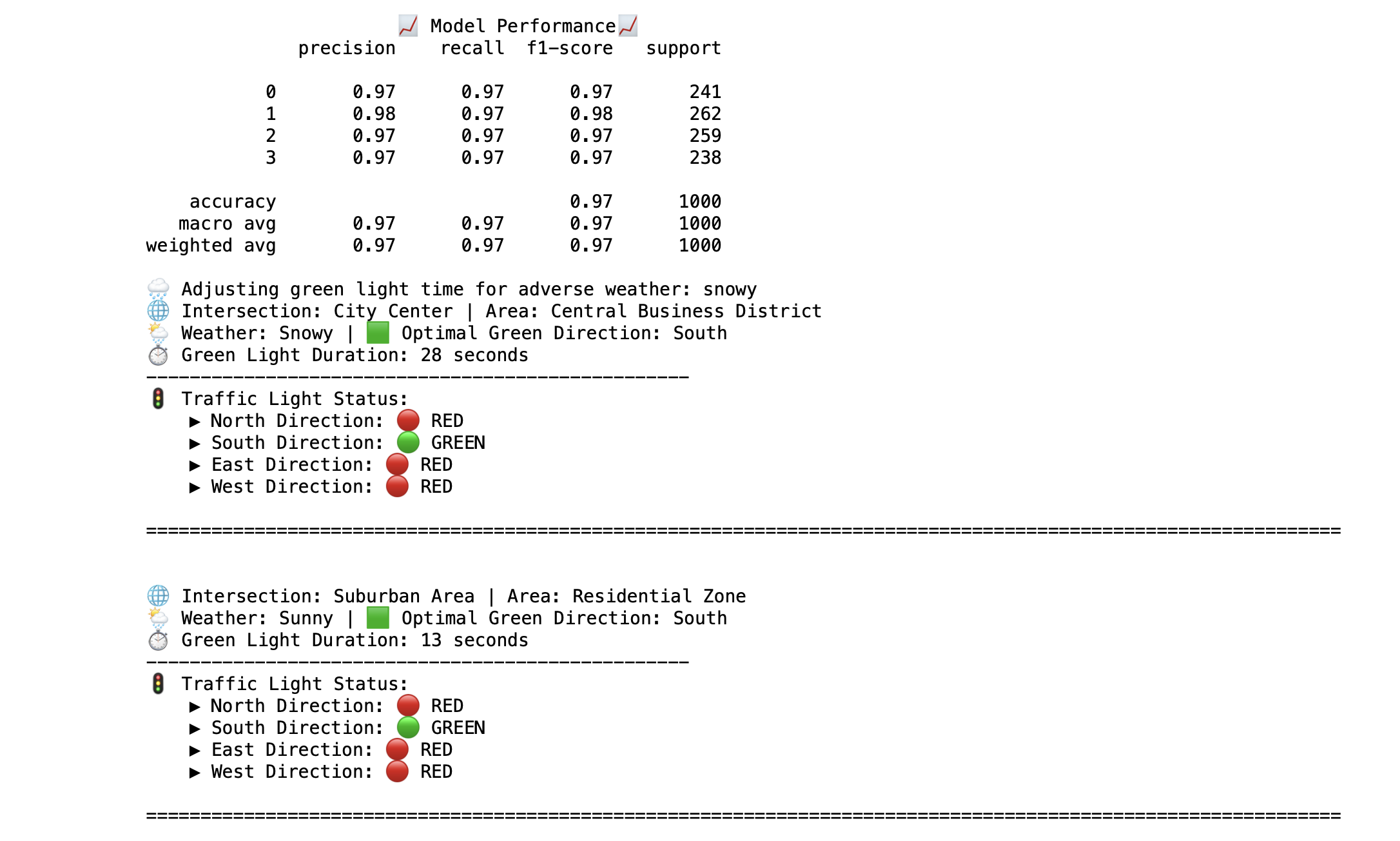
****

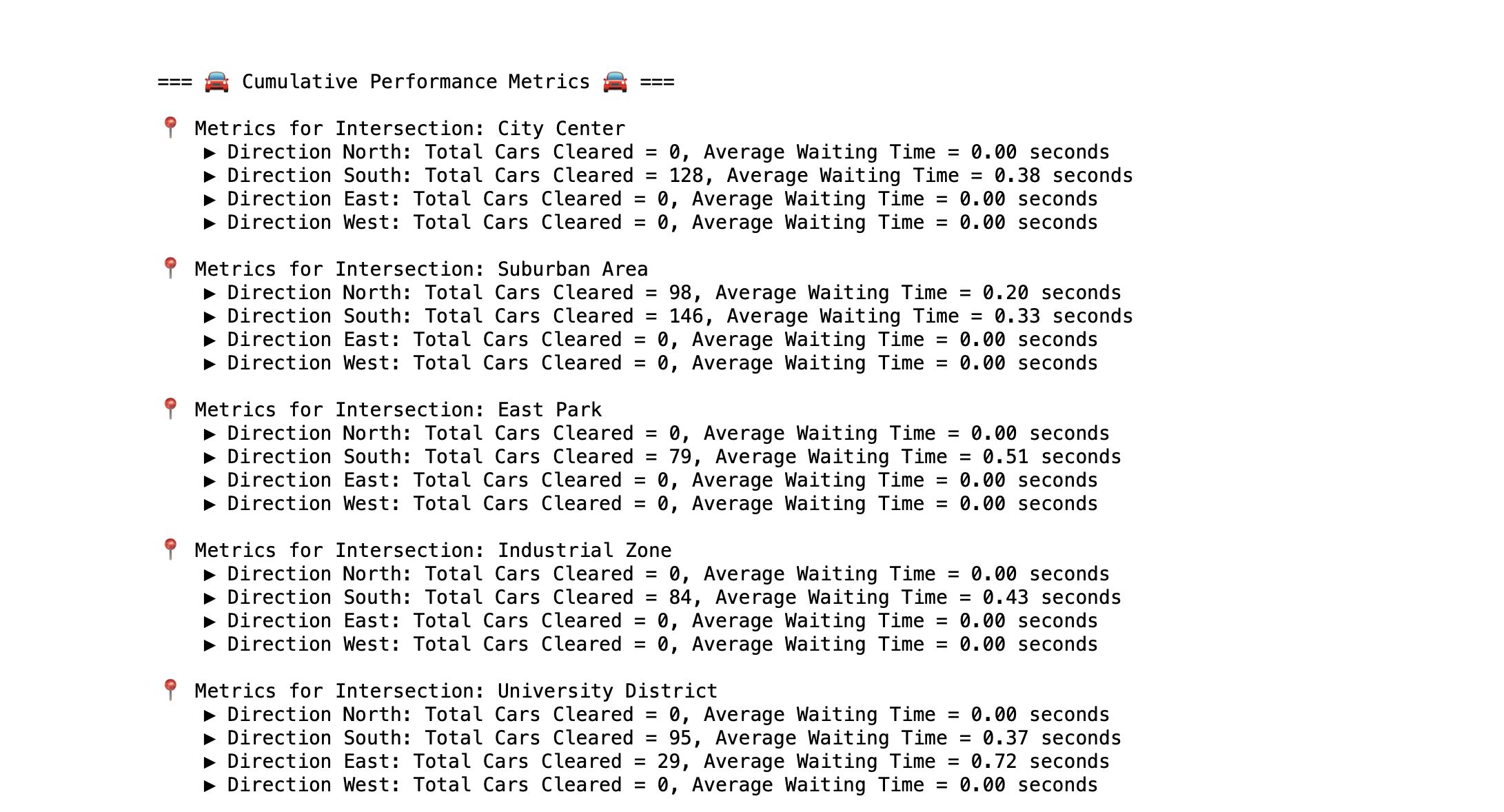
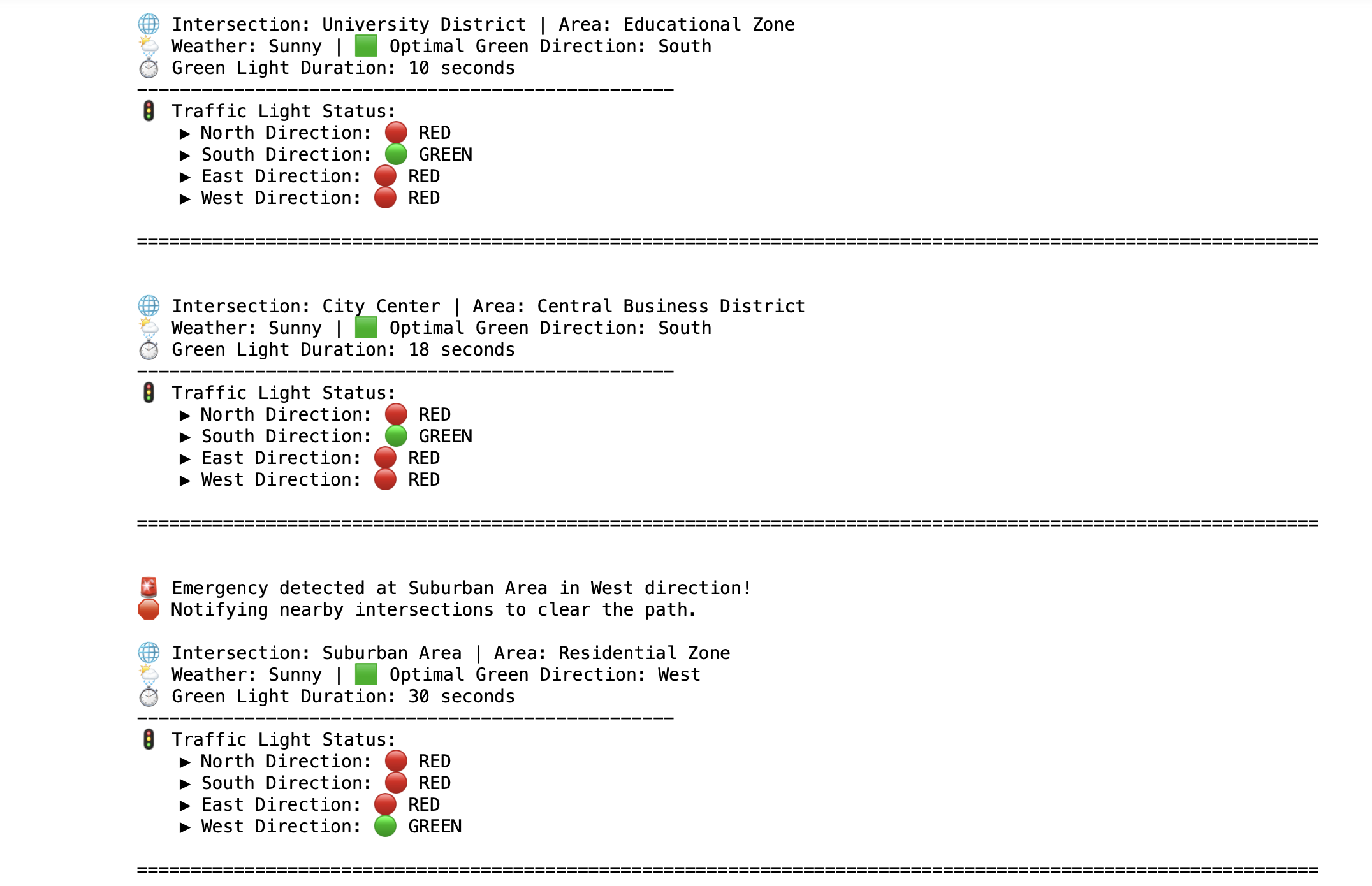
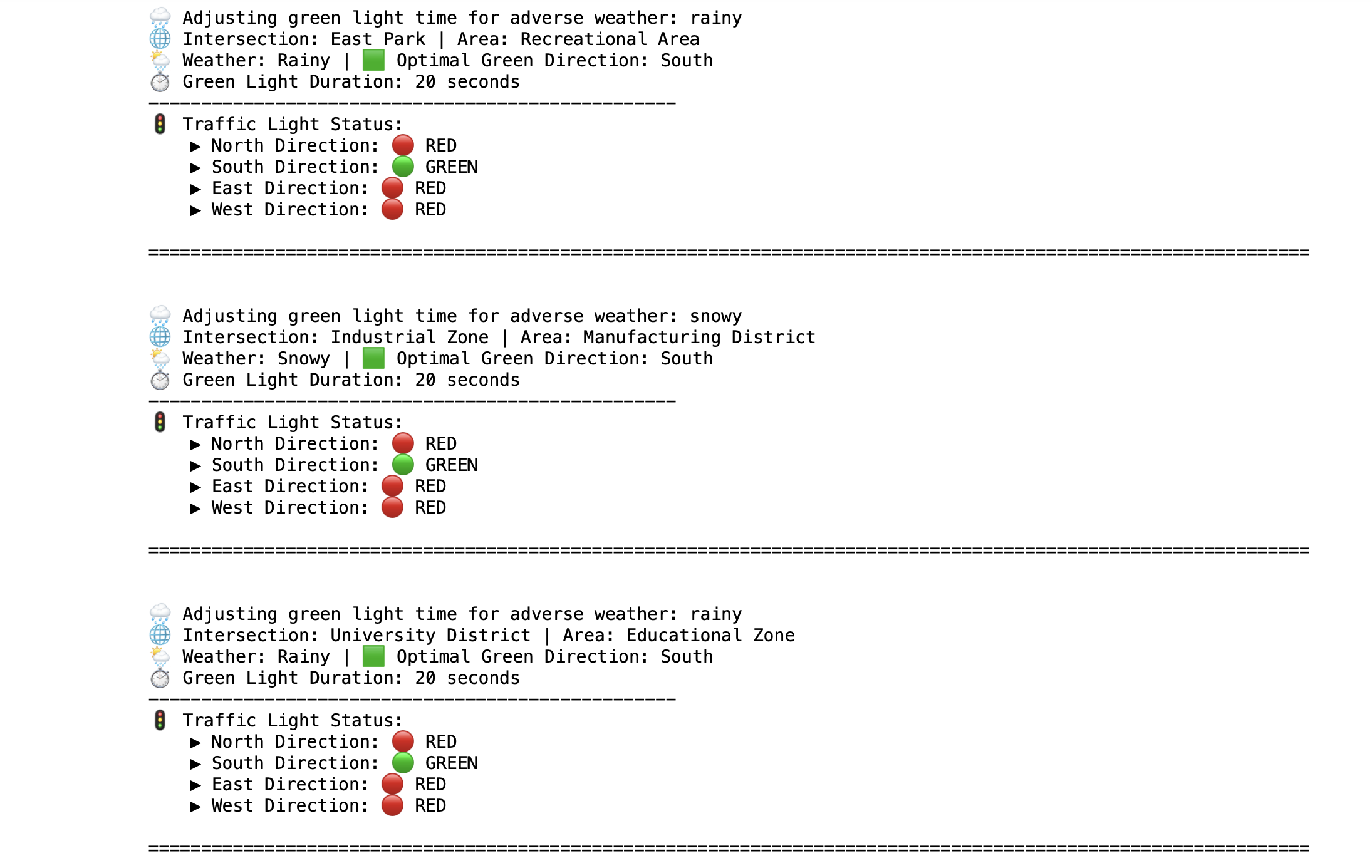
****

****

****

**Output**

****

****

**Code Analysis:**

**1. Preprocessing (preprocess\_data\_new)**

* Purpose: Prepares the dataset for model training by cleaning and scaling the data.
* Steps:

Reads the dataset from a CSV file.

Identifies the direction with the highest traffic flow (Optimal\_Green) as the target variable.

Scales numeric features like traffic flow and queue lengths for consistency.

* Features Preprocessed:

North\_Flow, South\_Flow, East\_Flow, West\_Flow: Traffic flow in each direction.

Queue\_Length\_North, Queue\_Length\_South, Queue\_Length\_East, Queue\_Length\_West: Queue lengths in each direction.

**2. Model Training (train\_model\_new)**

* Purpose: Trains a RandomForestClassifier to predict the optimal green light direction.
* Steps:

Splits the data into training and testing sets (80%-20% split).

Trains the model with 100 decision trees.

Evaluates the model’s accuracy with a classification report.

* Returns:

The trained model and the feature columns for prediction.

**3. TrafficLight Class**

* Purpose: Manages the state of traffic lights at each intersection.
* Attributes:

status: Current status (red or green) for each direction.

green\_time: Duration for which the green light stays active.

total\_cars\_cleared and total\_waiting\_time: Metrics tracking performance.

* Key Methods:

change\_status: Changes the light’s state for a given direction.

reset\_lights: Resets all lights to red.

calculate\_waiting\_time: Computes average waiting time based on traffic flow.

display\_metrics: Outputs cumulative performance metrics.

**4. Sensor Class**

* Purpose: Simulates real-world sensors to collect data.
* Key Methods:

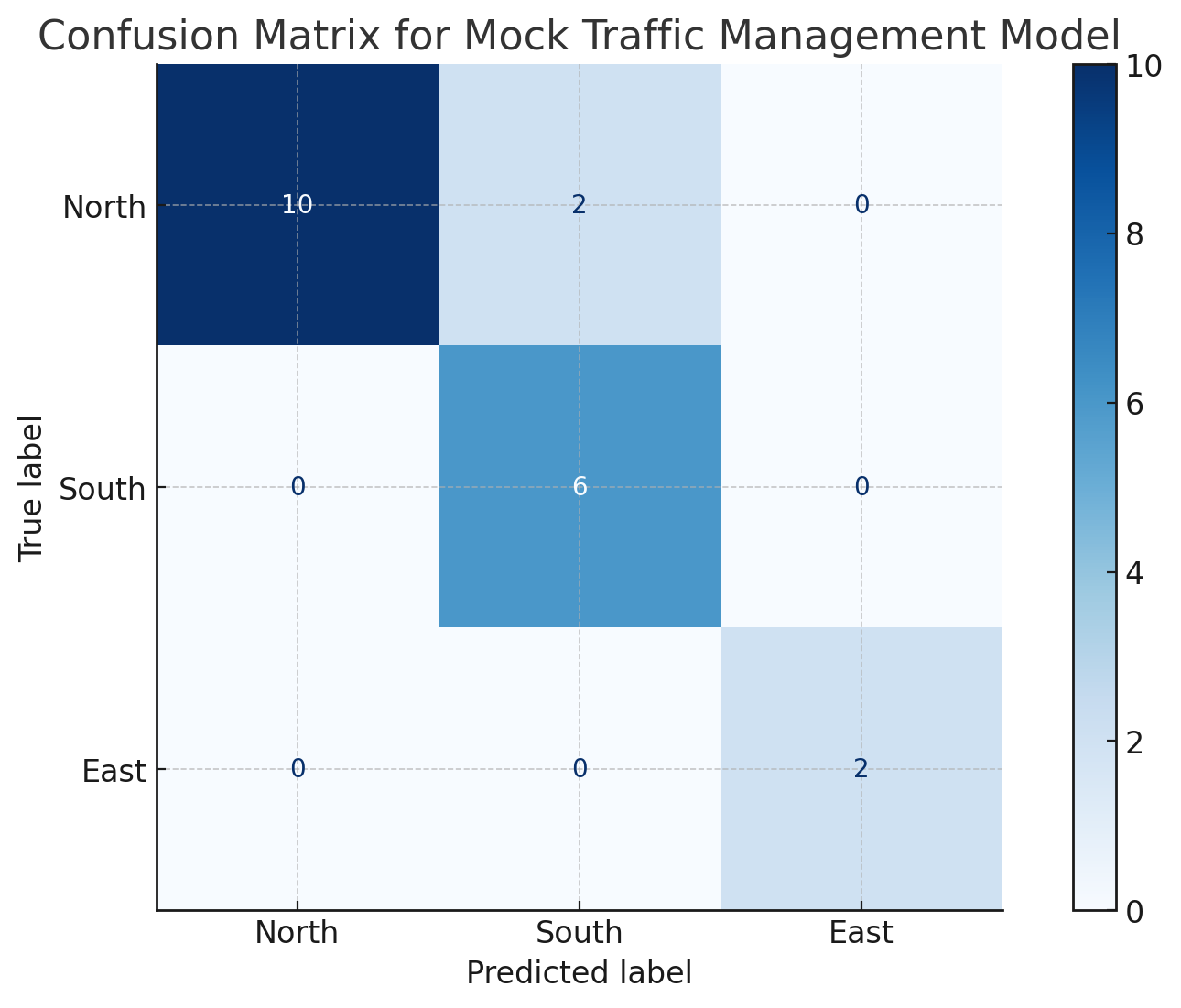
measure\_traffic\_flow: Generates random traffic flow values for all directions.

measure\_weather\_conditions: Returns a random weather condition (sunny, rainy, snowy).

**5. SmartTrafficSystemWithWeather Class**

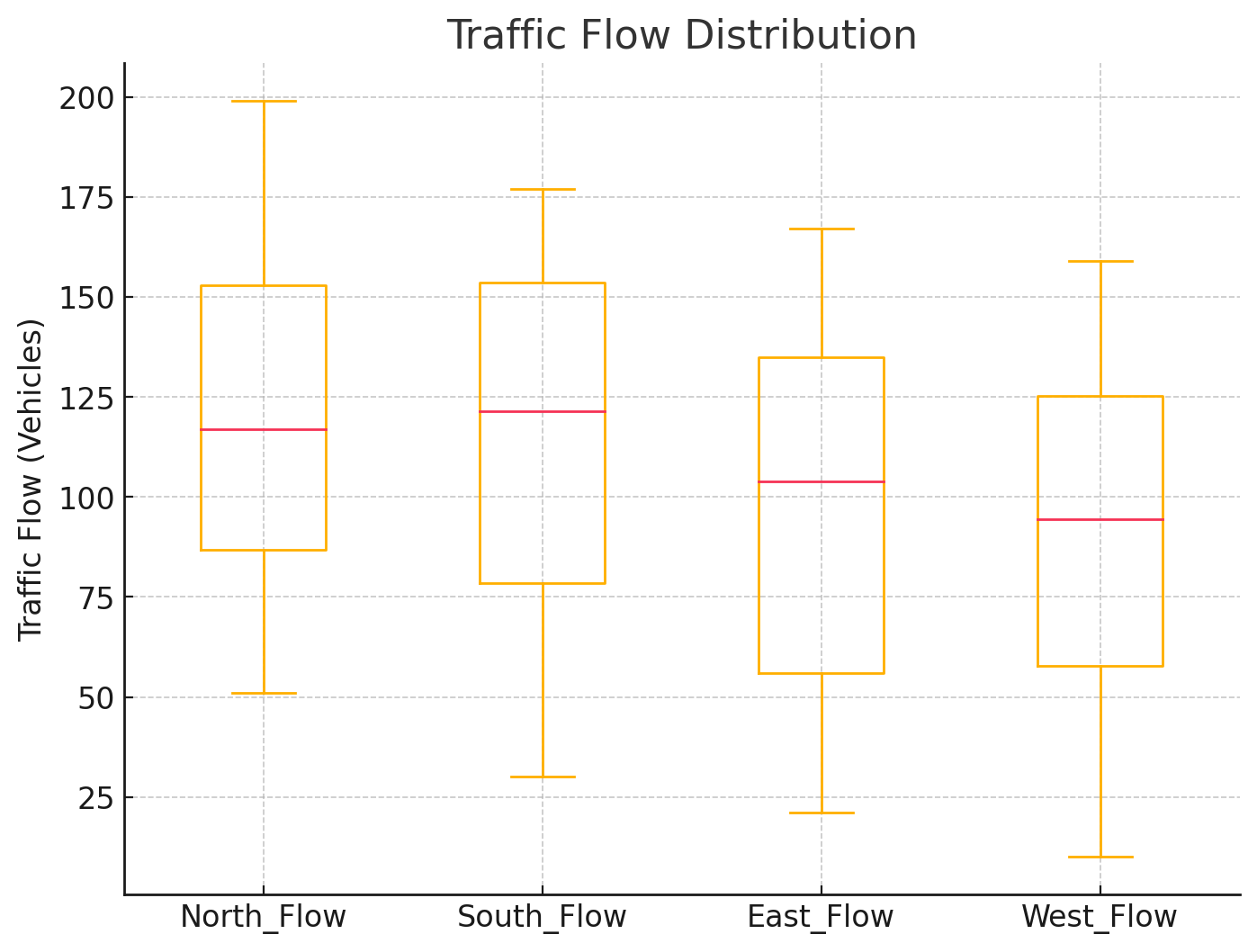
* This class manages intersections, handles traffic decisions, and integrates new weather and emergency handling functionalities.

**Results**



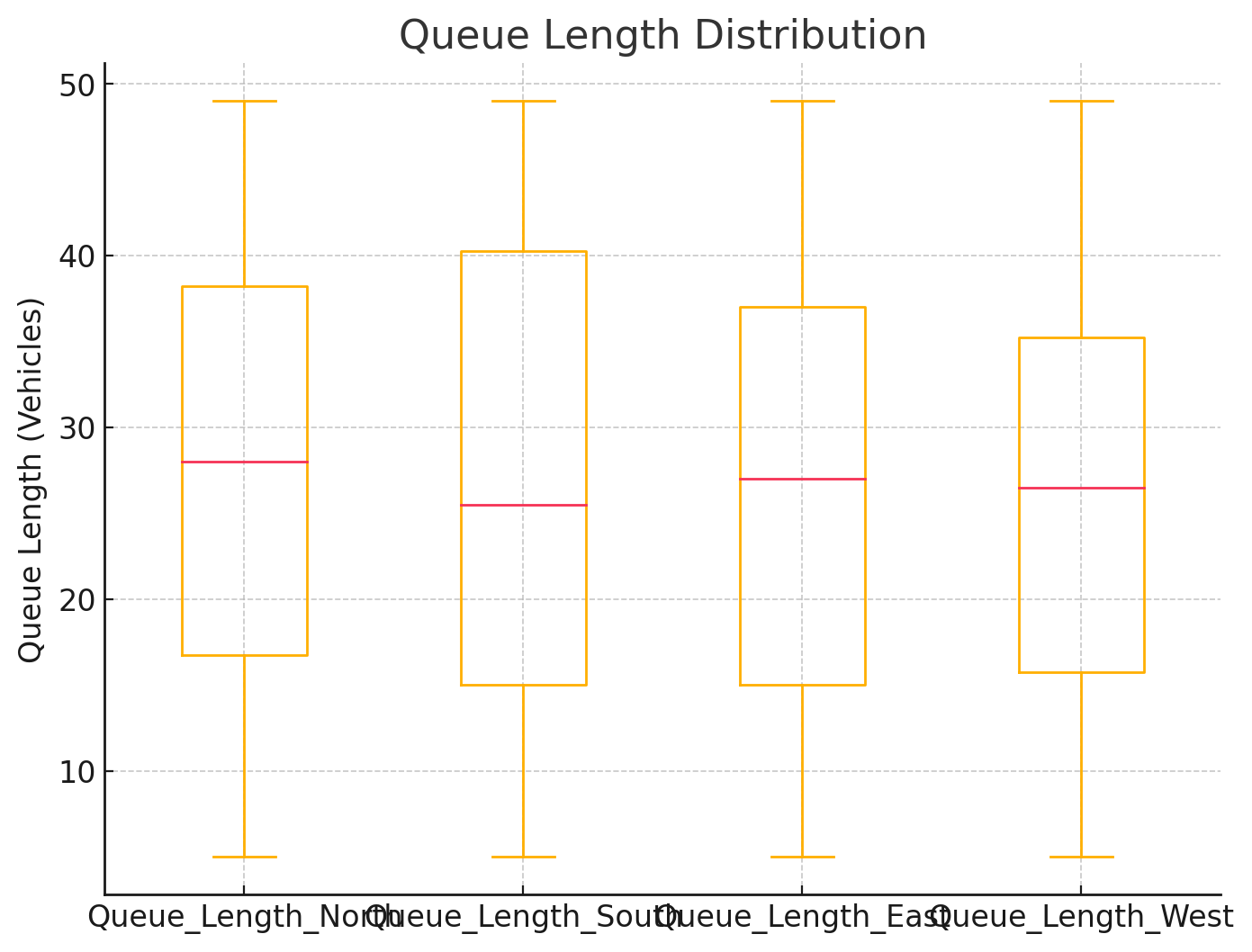
**Fig 1.1:** Confusion Matrix for Mock Traffic Management Model

**Fig 1.1** represents the counts of predictions made by the model for traffic direction, with rows indicating the actual (true) labels and columns showing the predicted labels. For the true label "North," the model correctly predicted "North" 10 times, but it incorrectly predicted "South" 2 times and never predicted "East." For the true label "South," the model accurately predicted "South" 6 times and never mistakenly predicted "North" or "East." Similarly, for the true label "East," the model correctly predicted "East" 2 times and never incorrectly predicted "North" or "South." Each cell of the table thus provides the number of instances where the model's prediction matched or deviated from the actual traffic direction.



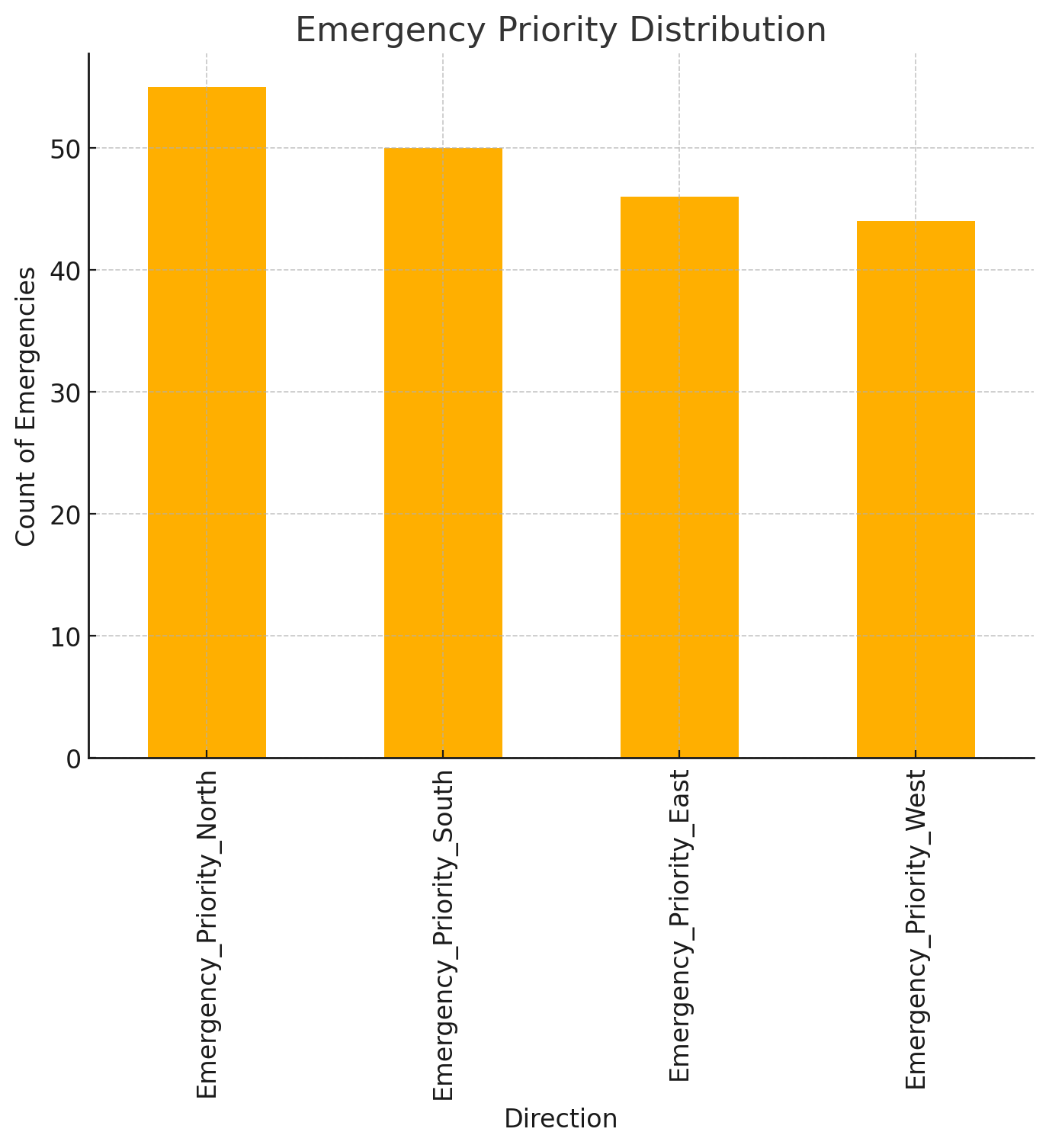
**Fig 1.2**: Traffic Flow Distribution

**Fig 1.2** illustrates the distribution of vehicle traffic across four directions: North, South, East, and West. It represents the traffic flow, defined as the number of vehicles passing through a given direction at a specific time. The box in each plot shows the interquartile range (the middle 50% of the data), with a line inside indicating the median traffic flow for that direction. Whiskers extend to the minimum and maximum values within a specified range, while any points beyond the whiskers are outliers, reflecting unusually high or low traffic volumes. By comparing these boxplots, you can identify which directions experience higher traffic or greater variability; wider boxes or longer whiskers signify more variation in traffic flow.



**Fig 1.3:** Queue Length Distribution

**Fig 1.3** illustrates the queue lengths, representing the number of vehicles waiting at a red light across all four directions. The queue lengths are normalized using standardization to scale the data to a comparable range for better visualization. Directions with larger median queue lengths suggest higher congestion levels, while a wider spread in the data indicates more inconsistent traffic flow in that direction.



**Fig 1.4:** Emergency Priority Distribution

**Fig 1.4** illustrates the frequency of emergency vehicles detected in each direction, highlighting the count of occurrences requiring traffic light adjustments. The data represents emergency priority, marked by the presence of vehicles like ambulances and fire trucks. By identifying which directions experience more frequent emergency traffic, the chart provides insights to help traffic systems prioritize these directions for quicker response times and improved traffic management.

**Discussion**

**Effectiveness of the system**

1. **Improved Traffic Flow**

* Key Metrics:

Reduced Congestion: Optimal allocation of green lights minimizes idle times and prevents bottlenecks.

* Total Cars Cleared: Tracks and displays the number of vehicles cleared at each intersection, showcasing improved throughput.
* Reduced Average Waiting Time: Dynamic adjustments and prioritization reduce waiting times for vehicles in all directions.

1. **Integration of Environmental Factors**

* Weather-Based Adjustments:

By extending green light durations during adverse weather (rainy or snowy conditions), the system ensures safe and steady traffic movement, accounting for reduced vehicle speeds.

* Pedestrian Safety:

Incorporates pedestrian flow data to balance vehicle and pedestrian priorities, enhancing safety and accessibility.

1. **Robust and Scalable Design**

* Input Validation:

Handles missing or invalid data seamlessly, ensuring robust operation even with incomplete inputs.

* Scalability:

The modular design supports the addition of new intersections and features, allowing the system to expand to a city-wide network.

1. **AI-Driven Insights**

* Model Accuracy:

The Random Forest Classifier provides high accuracy, effectively predicting the optimal green direction based on historical and real-time data.

* Feature Importance:

Identifies critical factors influencing traffic flow, offering insights to planners for long-term infrastructure improvements.

**Challenges Faced during Implementation**

1. **Dataset Creation and Preprocessing:**

The dataset was synthetic, requiring careful design to mimic real-world traffic scenarios. Generating realistic data points for traffic flow, queue lengths, and weather conditions was challenging.

Preprocessing included scaling and handling missing or invalid values, ensuring the data was suitable for training the machine learning model.

1. **Model Selection and Training:**

Choosing the appropriate algorithm involved balancing model complexity and efficiency. While Random Forest provided robust predictions, tuning the hyperparameters to prevent overfitting required significant experimentation.

1. **Emergency Handling:**

Designing a system to prioritize emergency vehicles without disrupting normal traffic flow required careful logic and testing, particularly for scenarios with overlapping emergencies.

1. **Weather and Dynamic Adjustments:**

Incorporating weather-based adjustments and ensuring seamless integration into the traffic light control algorithm was complex, as it required accounting for diverse traffic behaviours in different weather conditions.

1. **Error Handling:**

Ensuring the system remained robust to missing or anomalous data required extensive testing and the addition of validation and exception-handling mechanisms.

**Conclusion**

This project successfully demonstrates the potential of AI-powered traffic management systems in addressing the challenges of urban transportation. By leveraging a Random Forest Classifier and real-time data, the system dynamically optimizes traffic signal timings, prioritizes emergency vehicles, and accounts for environmental factors such as weather and pedestrian flow. The modular design ensures scalability, allowing the system to expand to city-wide networks and adapt to future technological advancements.

Performance metrics highlight the system’s effectiveness in reducing congestion, minimizing waiting times, and improving traffic flow. Additionally, its robust handling of real-time data and emergency scenarios showcases its readiness for real-world implementation. With the integration of IoT devices, public transport data, and advanced learning algorithms, this system can serve as a cornerstone for building smarter, safer, and more efficient cities.

**References**

1. Institute of Transportation Engineers (ITE). (2023). Artificial Intelligence in Traffic Management. <https://www.ite.org/ITEORG/assets/File/AI%20Dev%20Trends.pdf>
2. Transportation Research Board. (2024). AI Traffic Management in Los Angeles. <https://trid.trb.org/view/1171176>
3. Singapore Land Transport Authority (LTA). (2023). Singapore's AI-Driven Traffic System. <https://www.lta.gov.sg/content/ltagov/en/getting_around/driving_in_singapore/intelligent_transport_systems.html>
4. Machine Learning for Traffic Management. (2024). <https://trid.trb.org/view/1907042>
5. Deep Learning in Traffic Systems. (2023).

<https://arxiv.org/abs/2303.00337>

1. Benefits of AI in Traffic Management. (2024). <https://onlinepubs.trb.org/onlinepubs/circulars/ec168.pdf>
2. Data Privacy in Traffic Management. (2023). <https://www.ite.org/technical-resources/topics/transportation-system-management-and-operations/>
3. Challenges in AI Traffic Management. (2024). <https://trid.trb.org/view/425239>
4. Future Trends in AI Traffic Management. (2024). <https://www.ite.org/ITEORG/assets/File/AI%20Dev%20Trends.pdf>
5. AI and Autonomous Vehicles. (2024). <https://www.lta.gov.sg/content/ltagov/en/industry_innovations/technologies/autonomous_vehicles.html>