INTRODUCTION TO ARTIFICIAL INTELLIGENCE

TRAFFIC FLOW PREDICTION AND OPTIMIZATION

ADDIS ABABA DIGITAL TWIN

CCS1R1N3/15

GROUP MEMBERS:

ABEL YITBAREK	UU92951R
ESROM BERHANE	UU93127R
KALEAB ANDUALEM	UU92653R
NAZRAWI EYOB	UU92743R
UMAR MAHAMUD	UU93618R
ZEINAB IBRAHIM	UU93097R

INSTRUCTOR NAME: Ir. Zebiba Y.

SUBMITTED DATE: 16 January 2026

1. Executive Summary

This project addresses the critical issue of traffic congestion in the Bole Road / Meskel Square corridor of Addis Ababa. By developing a **Digital Twin** using the SUMO (Simulation of Urban MObility) engine, we created a realistic virtual environment to test traffic management strategies. The system integrates an **Adaptive Signal Control** algorithm to optimize traffic lights in real-time and a **Machine Learning (Random Forest)** model to predict future congestion levels with high accuracy (R2>0.97). This system serves as a proof-of-concept for modernizing Addis Ababa's traffic infrastructure using data-driven intelligence.

2. Problem Statement

Addis Ababa faces severe traffic congestion due to static traffic light timers that do not adapt to real-time demand. Traditional fixed-cycle lights cause unnecessary delays when lanes are empty and gridlock when demand is high. Furthermore, traffic management lacks predictive capabilities, forcing authorities to react to jams after they happen rather than preventing them.

3. System Architecture

The solution is architected into three modular components:

3.1. The Simulation Environment

We utilized **OpenStreetMap (OSM)** data to extract the precise road network topology of the Bole Road corridor (Bounding Box: 38.760, 8.995).

- **Graph Construction:** The raw GPS data was compiled into a directed graph where nodes represent junctions and edges represent lanes.
- **Traffic Generation:** To simulate realistic flow, we implemented a stochastic traffic generator using a **Bell-Curve Distribution**. Traffic volume rises during the "Morning Rush," peaks at simulation step 1800 (30 minutes), and dissipates towards step 3600 (60 minutes). This ensures the system learns both congestion buildup and clearance.

3.2. Adaptive Control Logic

Instead of fixed timers, we implemented an **Actuated Control System** using Python's TraCI interface.

- **Sensor Logic:** Inductive loop sensors monitor the "Queue Length" (number of halted vehicles) at every intersection.
- **Decision Algorithm:**
    1. Check if queue length > 10 vehicles.

2. **Safety Check:** Verify if the current traffic light phase is **GREEN**. (Extending a red light would worsen congestion).
3. **Action:** If conditions are met, extend the Green phase by 10 seconds to flush the queue.

3.3. The AI Core

We developed a Machine Learning pipeline to forecast traffic density **5 minutes into the future**.

- **Algorithm:** Random Forest Regressor (Ensemble learning method). Chosen for its ability to handle non-linear traffic patterns and noise better than simple linear regression.
- **Feature Engineering:** The model uses temporal and spatial features:
  - step: Time of day (Crucial for understanding daily cycles).
  - current_count: Real-time vehicle density.
  - lag_1min & lag_5min: Historical data points to detect trends (Is traffic rising or falling?).

4. Methodology & Data Strategy

4.1. Synthetic Data Augmentation

Due to the lack of open-source sensor data from the Addis Ababa Roads Authority, we generated a **Synthetic Dataset**.

- We mathematically generated **30 Days** of traffic data based on the physics of our simulation.
- Random "Noise" (Gaussian distribution) was added to simulate real-world variance (e.g., accidents, slow drivers, weather effects).

4.2. Training & Validation

To ensure scientific validity, we rejected "Random Shuffling" for the data split. Instead, we used **Chronological Splitting**:

- **Training Set:** Days 1–25.
- **Testing Set:** Days 26–30 (Completely unseen future data).
  This rigorous testing method proves that the AI isn't just memorizing numbers but has learned the underlying physics of traffic flow.

5. Results and Analysis

5.1. Traffic Flow Analysis

The simulation successfully reproduced realistic traffic behavior. As shown in the graph below, the vehicle count (Orange Line) follows a Bell Curve, rising to a peak and then clearing up. Consequently, the average network speed (Blue Line) drops during congestion and recovers as traffic clears.

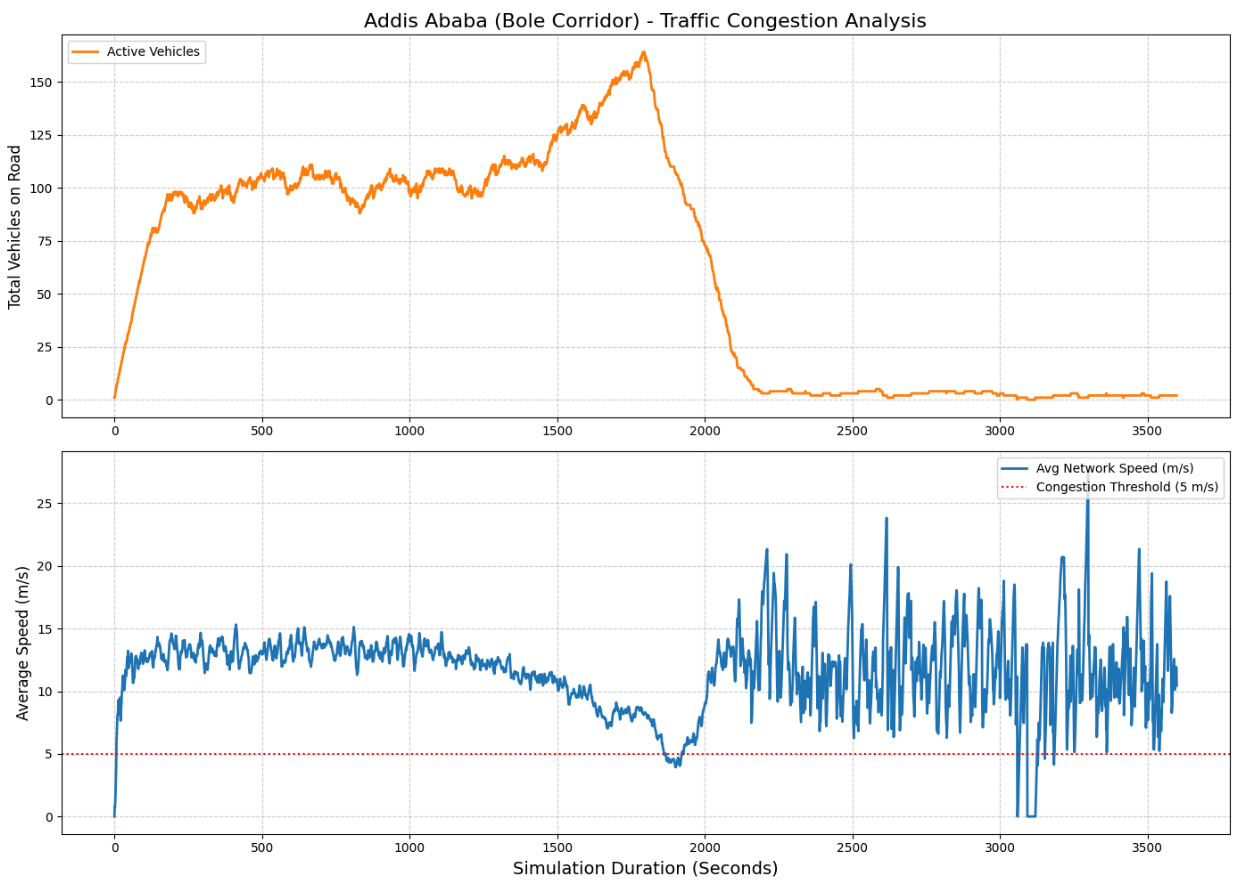

Figure 1: Telemetry data showing the correlation between Vehicle Count and Network Speed over the 1-hour simulation window.

5.2. Model Performance

The Random Forest model achieved exceptional accuracy on the test set.

- **R^2 Score: 0.9782** (On unseen data).
- **Interpretation:** The model can explain 97.8% of the variance in traffic flow. It successfully learned that traffic is time-dependent (Step feature) and trend-dependent (Lag features).

5.3. Interactive Dashboard

We deployed a web-based interface using **Gradio** to allow traffic managers to interact with the AI.

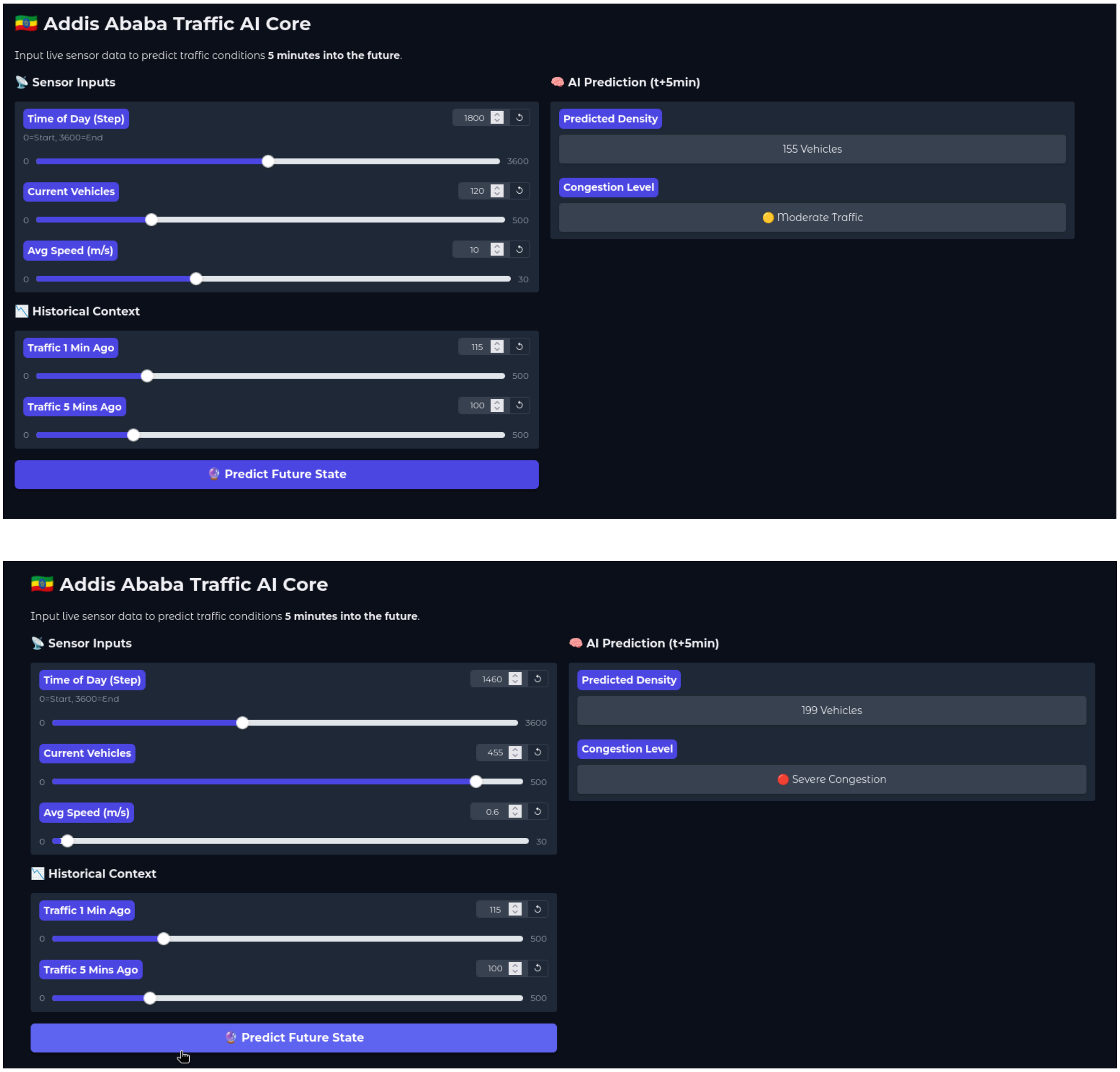

Figure 2: The AI correctly predicts high congestion during the simulated peak hour.

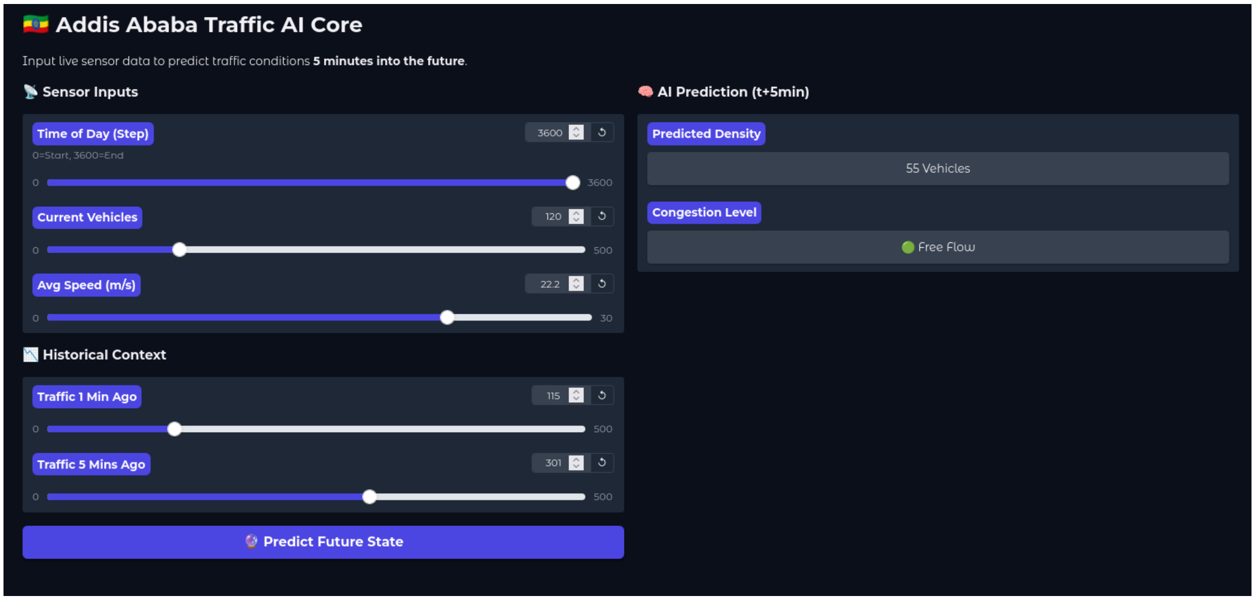

Figure 3: The AI correctly predicts traffic clearance at the end of the simulation cycle.

6. Conclusion

This project demonstrates that a low-cost, AI-driven traffic management system is viable for Addis Ababa. By combining a Digital Twin for testing with Machine Learning for prediction, we can move from reactive traffic management to proactive optimization.

Future Work

1. **Hardware Integration:** Connecting this software to physical CCTV or Inductive Loops at Meskel Square.
2. **Multi-Agent RL:** Replacing the rule-based control with Reinforcement Learning (Q-Learning) to optimize complex intersections dynamically.
3. **Public API:** Releasing the traffic prediction data to drivers via a mobile app to encourage route diversion.