Project Title:

Traffic Flow Optimization Using Reinforcement Learning (RL) in SUMO Simulation

Project Definition

Traffic Flow Optimization using Reinforcement Learning is the development of an intelligent traffic light control system that uses RL agents to adaptively manage signal timings in a simulated environment, aiming to reduce congestion, vehicle delay, and emissions.

Project Objectives

- The project will compare multiple RL algorithms, such as PPO and DQN, to evaluate their performance and reliability.
- The team will fine-tune the reward function to balance efficiency and fairness in traffic control.
- Additional simulations will be conducted with multiple intersections to test scalability.
- The codebase and models will be documented and organized to ensure the pipeline is reproducible.
- Findings will be compiled into a final report and demo video for project presentation.

Performance Metrics

- The average vehicle delay will be measured and targeted to remain under 15 seconds.
- The number of vehicles that pass through the network per hour, or throughput, will be maximized.
- Queue length will be monitored and minimized to reflect improved traffic flow.
- The variance in delay across different types of roads will be calculated to ensure fairness.
- RL-specific metrics such as reward convergence over time will be tracked.
- The speed at which the RL agent can make traffic light decisions will be evaluated for real-time feasibility.
- A preliminary PPO agent showed a 40% reduction in average delay compared to a fixed-time baseline.
- SUMO and its API will be used to gather metrics like vehicle delay, emissions, and queue lengths.
- Visual tools such as TensorBoard and Matplotlib will help interpret learning progress and performance.
- Results will be summarized in a table comparing baseline and RL agent performance.

- The system will be evaluated under different traffic scenarios to ensure robustness.

Ethical Considerations

- Bias toward main roads will be identified and mitigated using fairness-aware reward functions.
- Safety will be considered by simulating edge cases and adding oversight to the RL agent's actions.
- If real-world data is used, it will be anonymized to protect individual privacy.
- The system's design will align with smart city goals and promote ethical AI deployment.
- Ethical risks will be clearly stated and mitigations incorporated directly into the model design.

EXAMPLE SIMULATION

State, Action, and Reward – T-Junction Example

- **State:**
  At a standard T-junction with roads from the north, south, and east, the AI agent will observe the number of stopped vehicles on each approach. Specifically, it monitors the queue lengths at the stop lines of all three incoming roads. This state information reflects real-time congestion and helps the AI assess which direction requires priority.
- **Action:**
  The AI can choose between predefined traffic light phases. For example:
  - **Action 0:** Allow north–south straight movement.
  - **Action 1:** Allow east to north left turn.
  - **Action 2:** Allow east to south right turn or straight.
  - **Action 3:** All red for pedestrian crossing or emergency override.
    These actions are exclusive, and the agent selects one at each time step based on observed traffic.
- **Reward:**
  After each action, the agent receives a reward equal to the decrease in total waiting time across all approaches. If the combined waiting time before the action was 350 seconds and after is 310 seconds, the agent receives a reward of +40. A penalty (negative reward) is given if waiting time increases, discouraging inefficient decisions.