**1. Introduction (2–3 pages)**

* Motivation for autonomous vehicles (AVs)
* Challenges in urban driving environments
* Importance of real-time decision-making
* Why reinforcement learning (RL)?
* Research question and objectives
* Contribution of the paper

**2. Background and Literature Review (6–8 pages)**

**2.1. Overview of Autonomous Vehicles**

* Levels of autonomy (SAE levels)
* Sensor technology: LiDAR, radar, cameras
* Decision-making pipeline in AVs

**2.2. Basics of Reinforcement Learning**

* Key concepts: agent, environment, states, actions, rewards
* Exploration vs. exploitation
* Value-based methods (e.g., Q-learning)
* Policy-based methods (e.g., Policy Gradients)
* Actor-Critic models

**2.3. RL in AVs – State of the Art**

* Existing RL-based driving models (e.g., Deep Q-Networks, PPO, SAC)
* RL in simulated vs. real environments
* Limitations of traditional methods (e.g., rule-based systems)

**3. Problem Formulation (3–4 pages)**

* Defining the decision-making problem
* Modeling urban driving as an RL problem
* State space (sensor inputs)
* Action space (vehicle controls)
* Reward structure (safety, efficiency, comfort)
* Handling uncertainty and stochasticity (e.g., other drivers)

**4. Reinforcement Learning Approaches for AVs (6–8 pages)**

**4.1. Value-Based Methods**

* Q-learning and Deep Q-Networks (DQN)
* Advantages and limitations in AVs

**4.2. Policy Gradient Methods**

* REINFORCE, A2C, PPO
* Suitability for continuous control

**4.3. Actor-Critic Methods**

* DDPG, TD3, SAC
* Use in high-dimensional and continuous environments

**4.4. Imitation Learning and Hybrid Methods**

* Combining supervised learning with RL
* Inverse RL

**5. Real-Time Implementation Challenges (4–5 pages)**

* Real-time sensor fusion
* Computational constraints (latency, power)
* Training vs. inference
* Sim-to-real transfer
* Safe exploration in real-world deployment

**6. Simulation Environments and Benchmarks (2–3 pages)**

* CARLA, SUMO, TORCS, OpenAI Gym
* Custom vs. standardized environments
* Benchmarking protocols

**7. Case Studies and Comparative Analysis (5–6 pages)**

* Comparative evaluation of RL algorithms for AVs
* Metrics: collision rate, route efficiency, response time
* Handling human-driven vehicle behavior (prediction, adaptation)
* Safety and efficiency trade-offs

**8. Future Directions (2–3 pages)**

* Multi-agent RL for traffic scenarios
* Federated or distributed learning
* Explainable RL and transparency
* Integration with V2X communication
* Ethical and legal considerations

**9. Conclusion (1–2 pages)**

* Summary of findings
* Contributions to AV development
* Final thoughts on RL applicability and deployment

**10. References (3–5 pages)**

* Aim for at least 40–60 scholarly references, especially recent (past 5 years)