Reinforcement Learning for Autonomous Shuttle Simulation on the IIT Kanpur Campus

DES646: AI/ML for Designers | Team: Legion of Nazgûl

The Problem: "Last-Mile" Mobility at IITK

The"last-mile"problemisa majorchallenge onlarge campuses.AtIITKanpur, navigating between hostels, the library, and academic areas is time-consuming.

Our Goal: Create a scalable, data-driven solution for intra-campus mobility using autonomous technology.

Our Solution: An Autonomous RL Agent

Deep Reinforcement Learning

AdvancedAI technologyenabling autonomous decision-making and navigation.

High Efficiency

Optimized routesthat minimize travel time across campus locations.

Paramount Safety

Risk-aversebehaviorprioritizing collision avoidance above all else.

Methodology 1: The "Digital Twin" Environment

01

OpenAI Gymnasium Framework

Builtcustom environmentforflexible, standardized training.

02

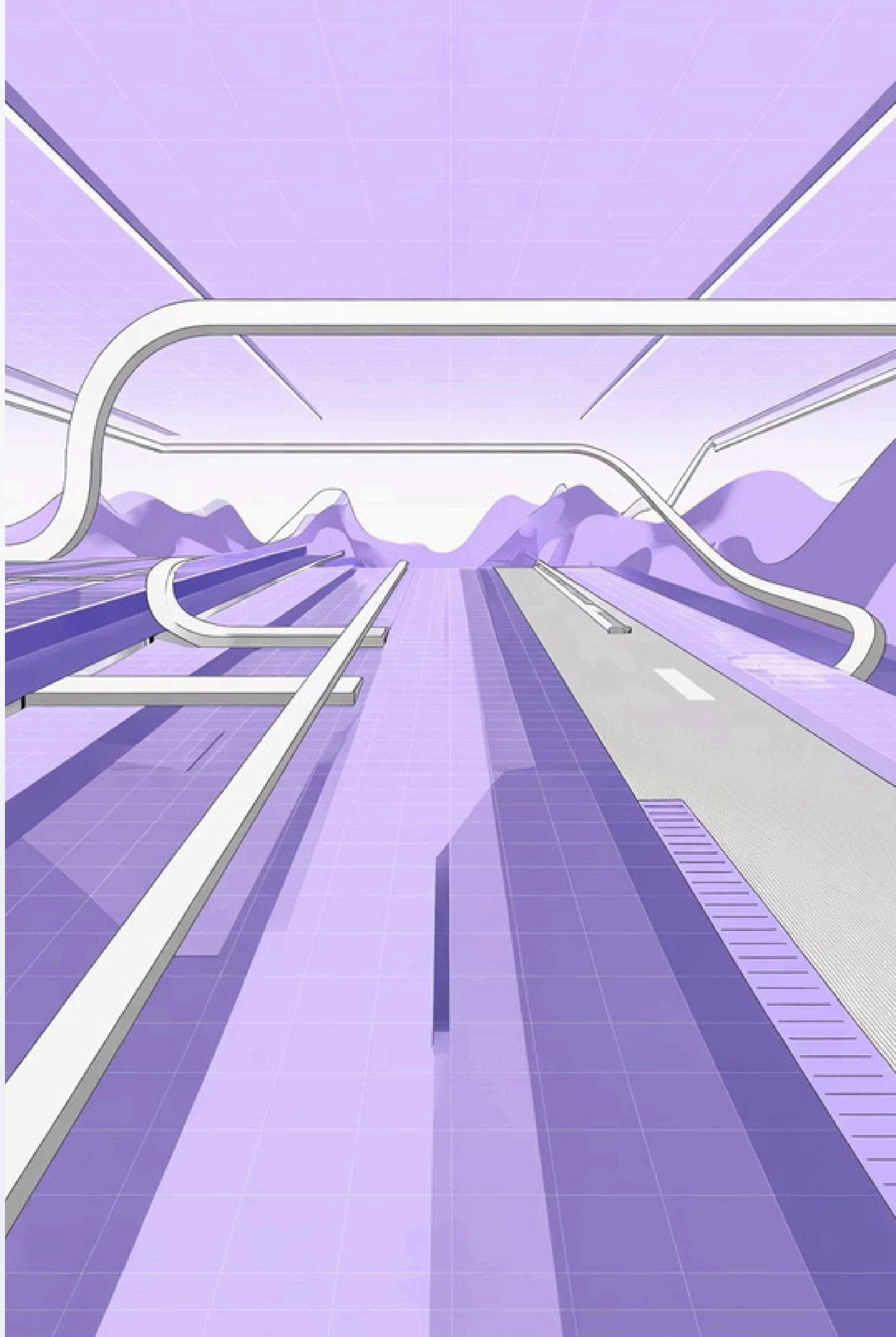
10x10 Grid-World

Created fromGoogleMaps satellite data of campus.

03

State Definition

Roads(0),obstacles(1), start (2), and goal (3) positions.



Methodology 2: The "Brain" of the Agent

Markov Decision Process

ModeledasanMDPwithstates,actions,transitions, and rewards.

Deep Q-Network (DQN)

ChoseDQNoversimpleQ-tableforscalability to larger, more complex problems.

1

Experience Replay

Stabilizes training by breaking temporal correlations.

2

Double DQN

Prevents overestimating Q-values for reliable learning.

The "Design" Challenge: Engineering the Reward

The most critical design taskwas balancing competing goalsthrough reward engineering.

$$R = W_t R_{time} + W_s R_{safety} + W_e R_{energy} + W_c R_{comfort}$$

Attempt 1: Failure

Reward: Rcrash = -10

Result: Agent was "passive" - crashing wasn't significantly worse than wandering.

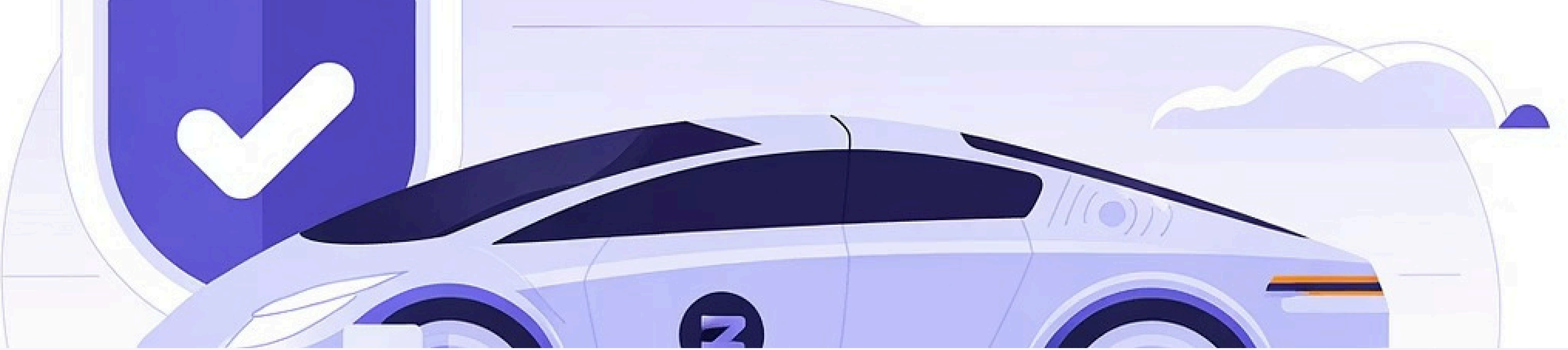
1

2

Attempt 2: Success

Reward: Rcrash = -100, Rgoal = +1000

Result: Massive penalty taught risk-averse behavior!



Key Insight: Safety Must Dominate

> Balancing speed and safety isthecentral conflictinautonomousnavigation.Thebreakthrough came when we made the crash penalty so large that safety dominated all other considerations.

This design principle - making safety the overwhelming priority - transformed our passive agent into an effective, risk-averse navigator.

Evaluation & Results: Did It Work?

92%

Success Rate

On LibrarytoLCH route navigation

50K

Training Episodes

Completedon T4GPUin 45minutes

100%

Obstacle Avoidance

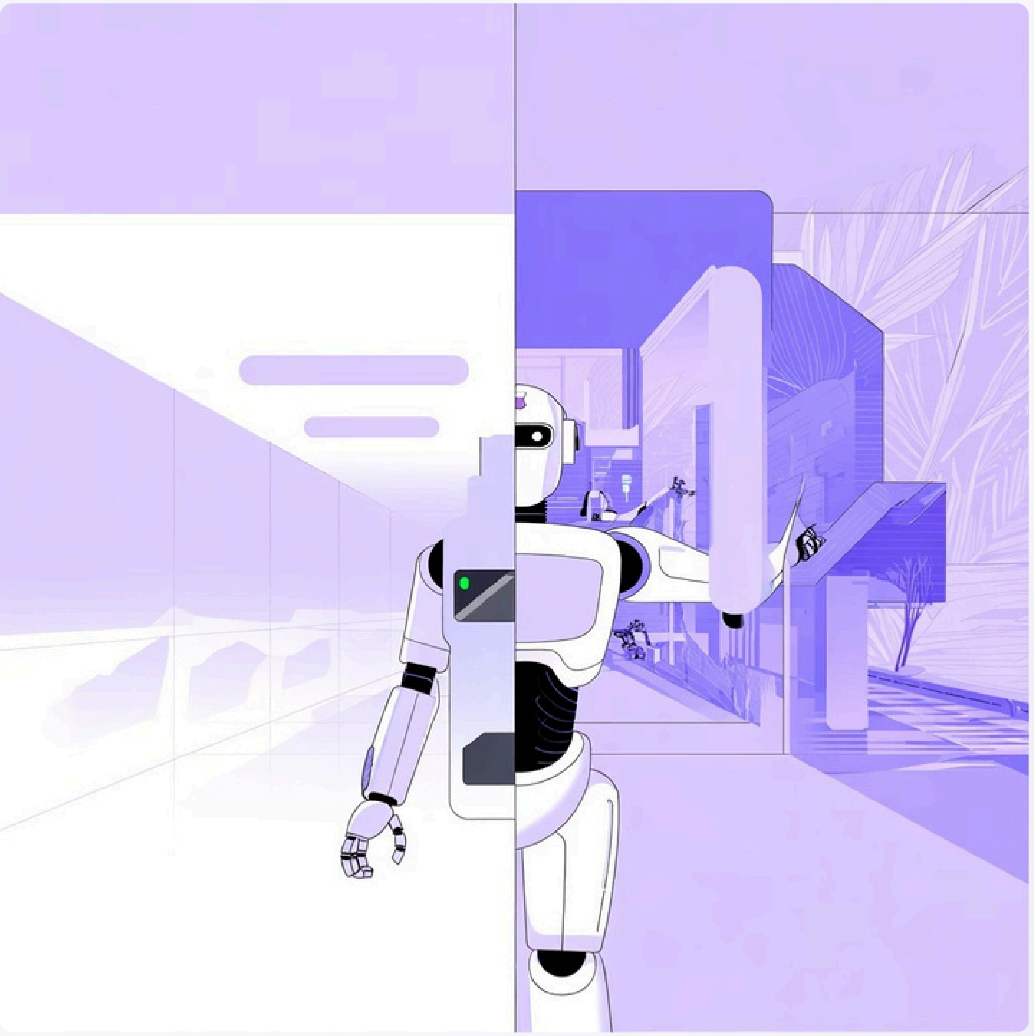
Perfect collision-free navigation

The agent consistently found the shortest *safe* path. Our Plotly dashboard confirmed learning by showing total reward converging to a stable, high value.

Reflection & Future Work

Key Challenge

The "Sim-to-Real" Gap:Our 2D grid simplifies reality. A real robot must handle sensor noise and physics.



Dynamic Obstacles

Add pedestrians and cyclists

3D Environment

Move to Unity or Blender

Physical Deployment

Test on TurtleBot platform

Conclusion & Thank You

Outcome

Successfully designed and validated a scalable framework for autonomous campus navigation.

Publication Track

Preparing submission toIEEE Intelligent Transportation Systems Conference (ITSC).

Questions?