Energy Conservation in Autonomous Agents using Reinforcement Learning

Volodymyr Beimuk

Supervisor: Dmytro Kuzmenko, Sr. lecturer,

Dep. Of Multimedia Systems

Main Objectives

- Energy-Efficient Autonomous Racing
- Reward Function Optimization with Fuel Penalties
- Multi-Objective RL Strategy Learning (Speed vs. Fuel Efficiency)

Reinforcement Learning (RL)

- Trial and error
- Interaction with environment
- Rewards or penalties

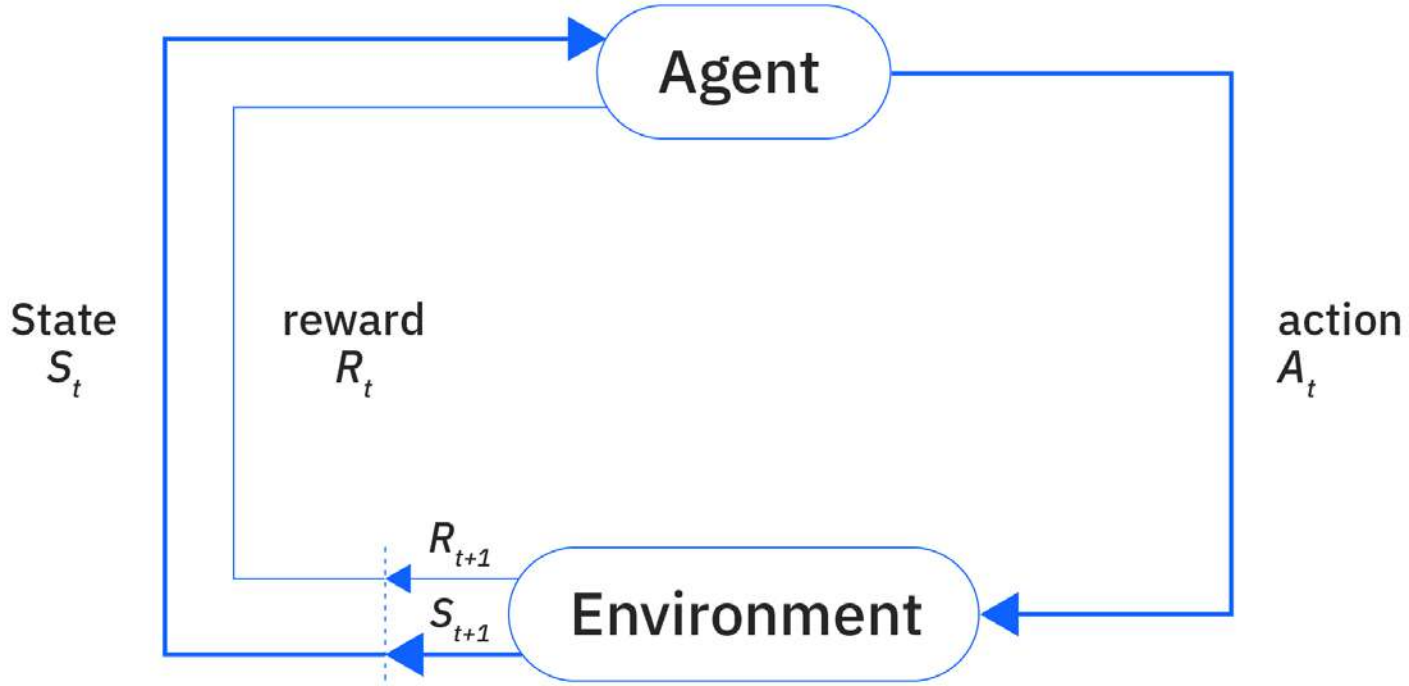


Illustration of a Reinforcement Learning algorithm [4]

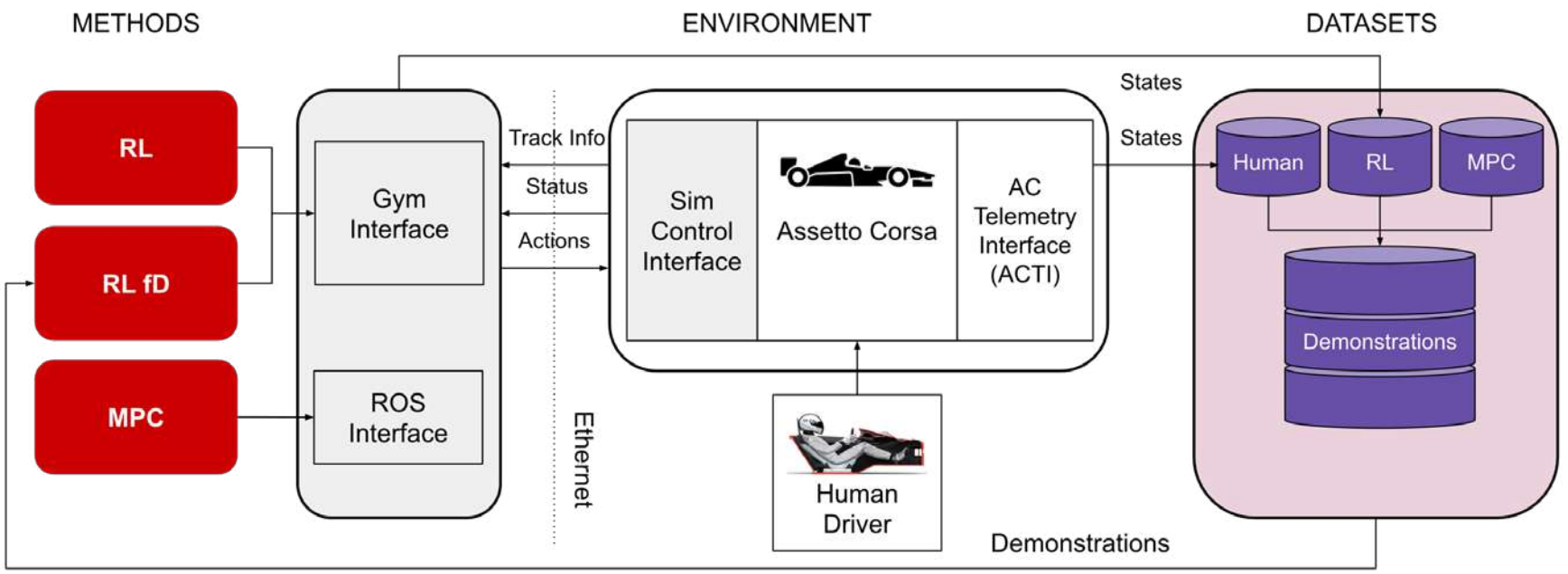
Benchmark

Realistic racing simulator

Extendable interface

Reward $r = v \cdot (1 - a \cdot d)$

- v: speed
- d: deviation from optimal path (L2 distance)
- a: tunable coefficient



Overview of the AssettoCorsaGym repository [2]

Modifications – Proposed Method

1) Efficiency from fuel preservation standpoint

- New reward function: $r = v \cdot (1 - a \cdot d) - b \cdot f$ where
  - f: immediate fuel change since previous simulation step
  - b: controls the strength of fuel penalty

2) Lap time performance and fuel efficiency trade-off analysis

- Logging fuel usage
  - Per each lap
  - Overall best per lap

Methods

Tracks:

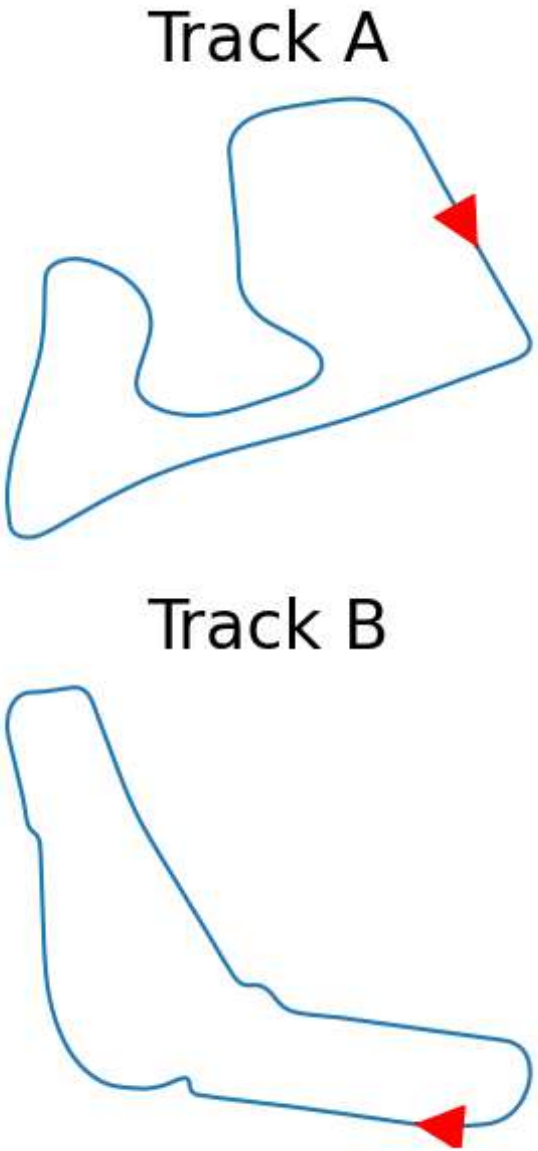
- Track A: short, balanced (4.3 km)
- Track B: long, technical (5.8 km)

Cars:

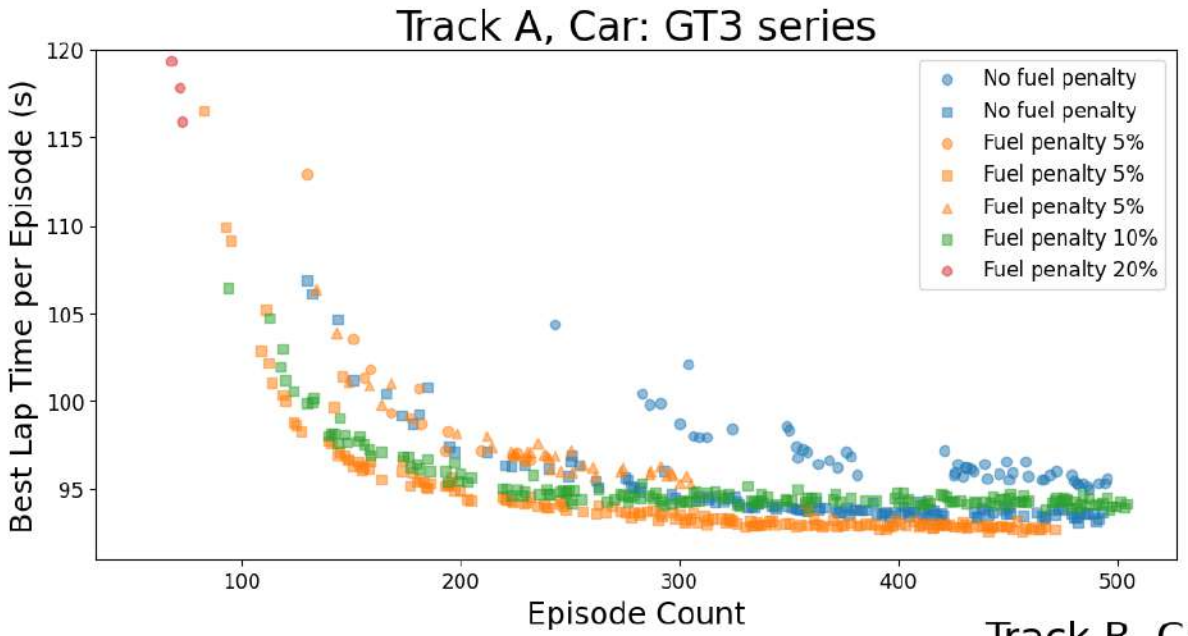
- Formula 3 series (lightweight, easy to control)
- GT3 series (heavy, harder to control)

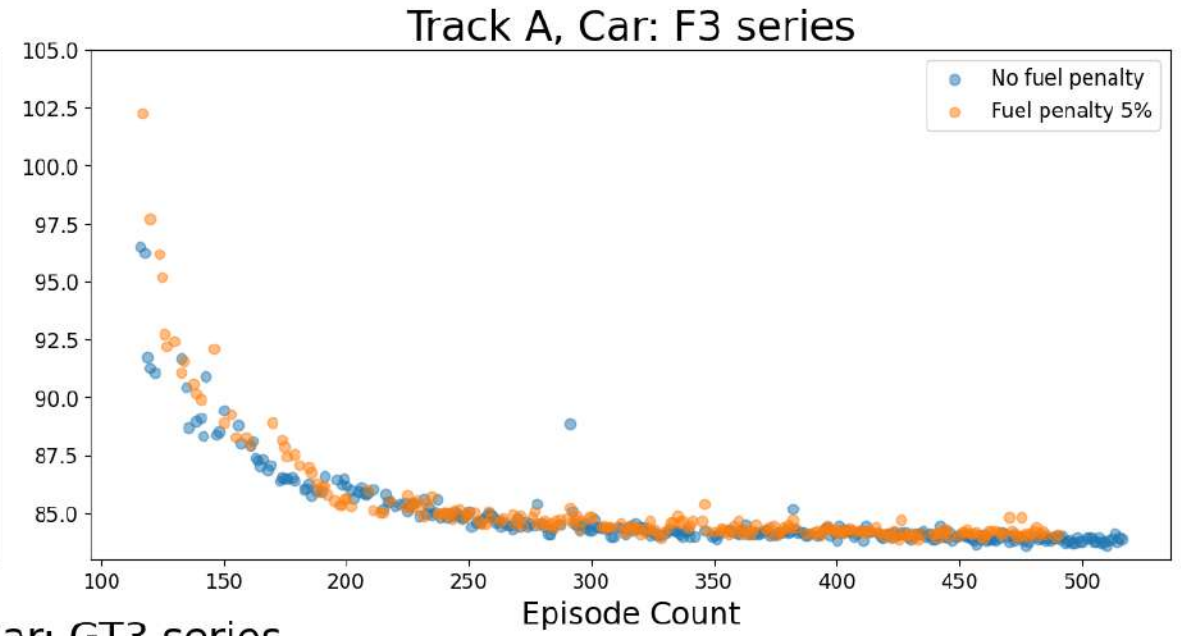
Training duration: ~500 episodes (48 hours)

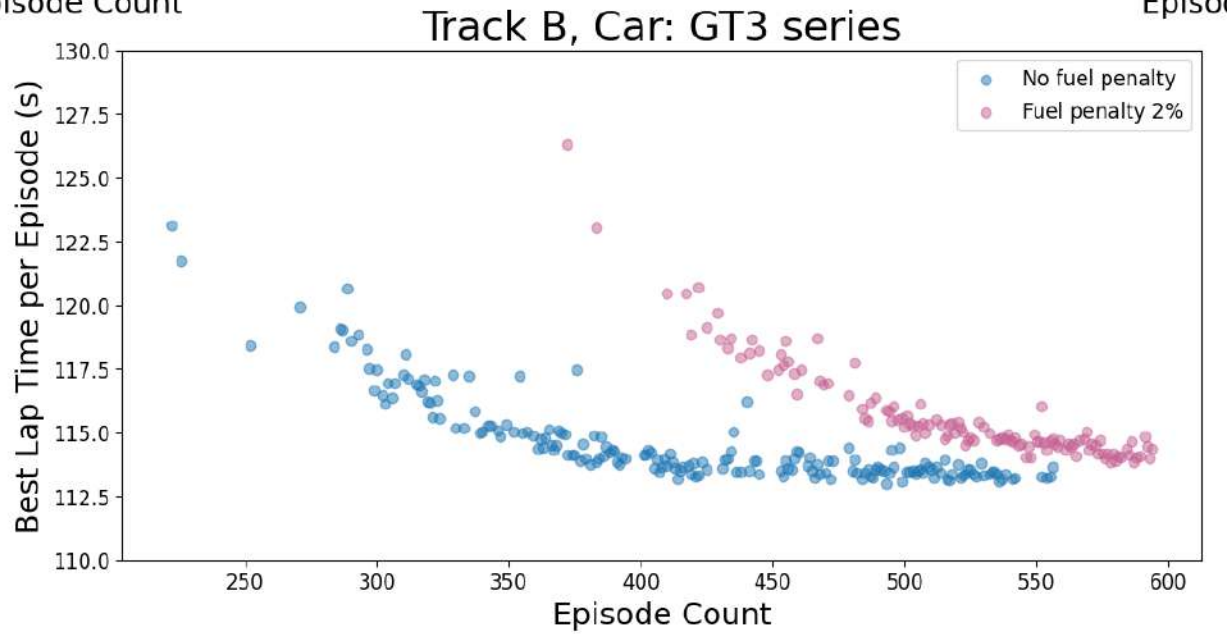
Fuel penalty levels: 0%, 2%, 5%, 10%, 20% (of mean reward per simulation step)



Learning Curves







Training Results Comparison

Track A, Car: GT3 series

Fuel Penalty	Random Seed	Best Lap Time (s) ↓	Fuel/Lap for Best Lap (L) ↓
0%	0	94.97	1.88
0%	1	92.85	1.82
5%	0	96.59	1.88
5%	1	**92.57**	1.59
10%	1	94.35	1.44
10%	2	93.80	1.43
20%	1	115.93	**1.01**

Track A, Car: F3 series

Fuel Penalty	Best Lap Time (s) ↓	Fuel/Lap for Best Lap (L) ↓
0%	**83.62**	0.97
5%	83.86	**0.90**

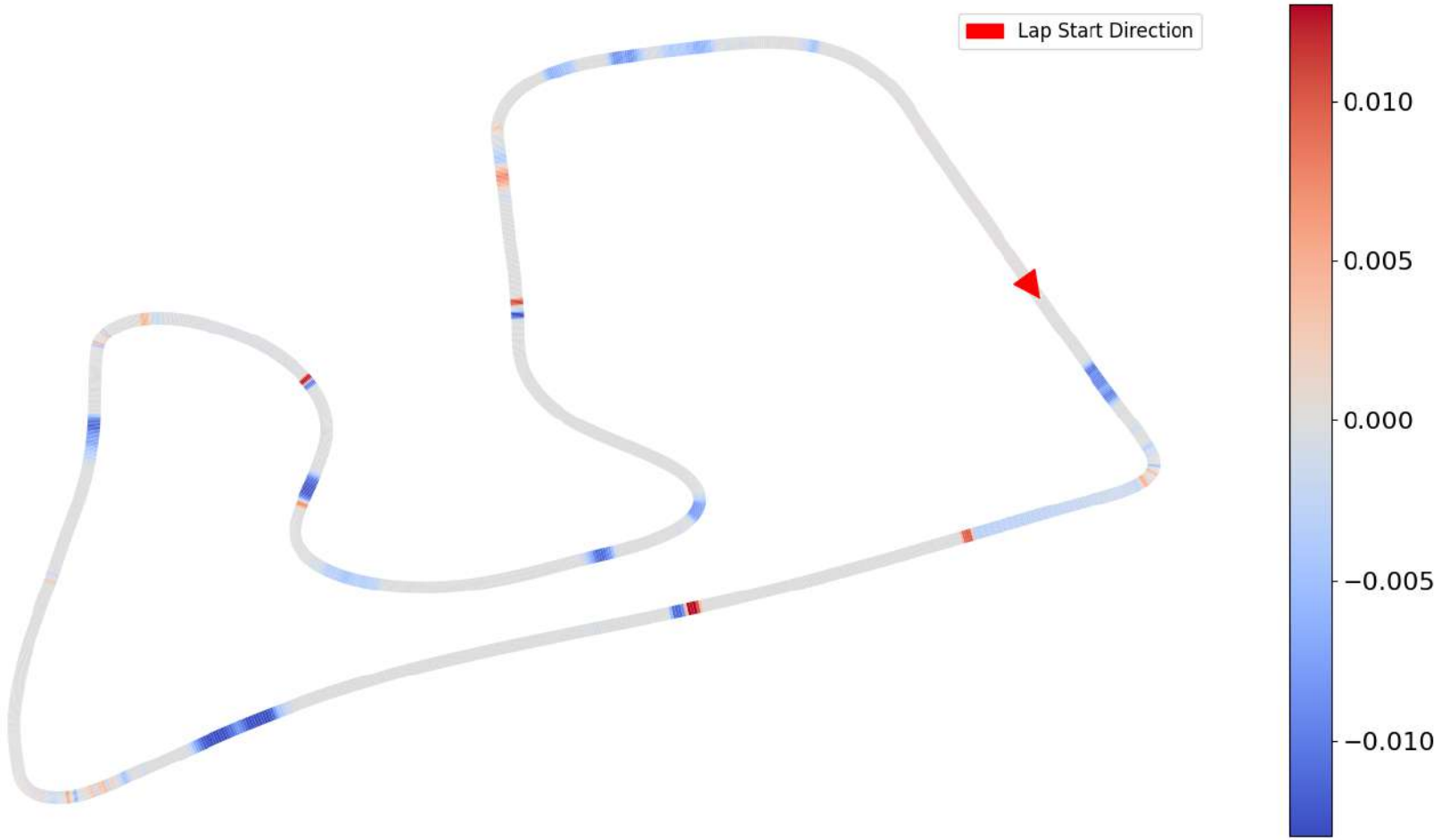
Track B, Car: GT3 series

Fuel Penalty	Best Lap Time (s)	Fuel/Lap for Best Lap (L)
0%	**112.99**	2.45
2%	113.85	**2.39**

Fuel Consumption (Penalty – No Penalty)

10.7% decrease in fuel usage

- sections: 5% fuel penalty model saves fuel
- sections: 5% fuel penalty model consumes more



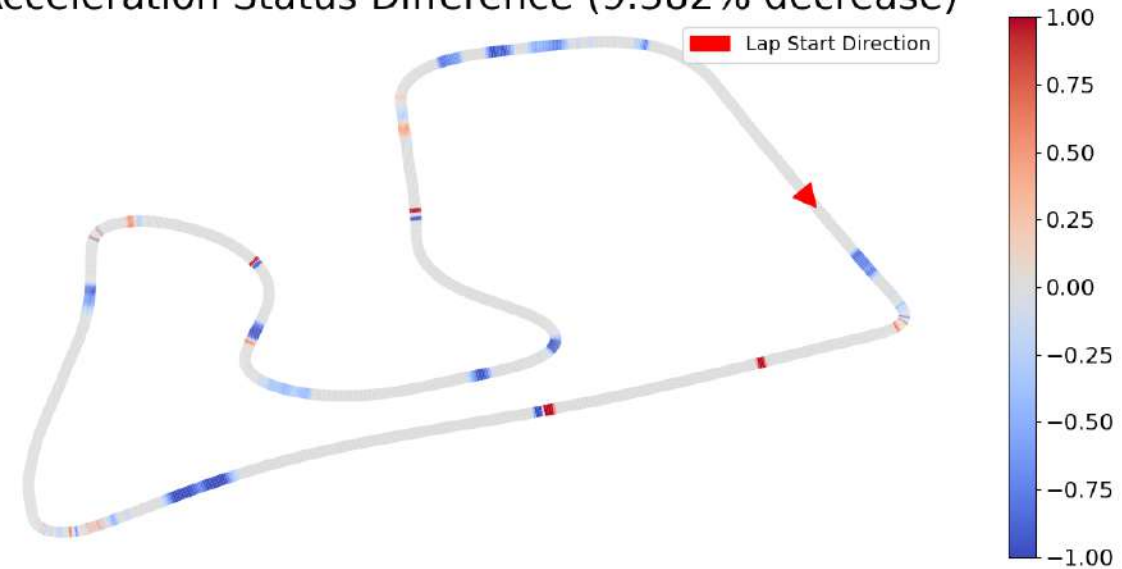
Quantifying Feature Impact on Fuel Efficiency

- Pearson Partial Correlation Coefficient
- Fuel usage increases with **acceleration**, **RPM** and **steer angle amplitude**, decreases with **braking**

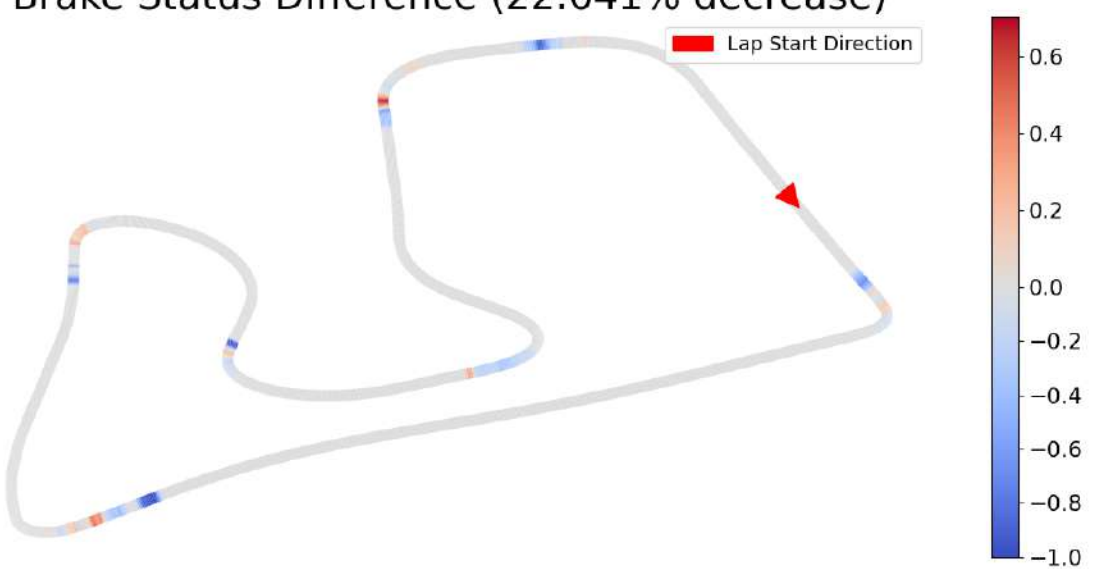
Feature	**Partial correlation (r)**	**P-value**	**Confidence Interval (95%)**
Acceleration	0.94	0.00	[0.93, 0.94]
Engine RPM	0.59	5.95×10^{-218}	[0.56, 0.62]
Braking	-0.29	5.84×10^{-47}	[-0.33, -0.26]
Steer Angle Amplitude	0.20	4.50×10^{-23}	[0.16, 0.24]
Speed	0.08	1.16×10^{-4}	[0.04, 0.12]
Roll	0.07	2.31×10^{-3}	[0.03, 0.11]
Yaw	-0.01	6.12×10^{-1}	[-0.05, 0.03]

Adapted Driving Behaviors (Penalty – No Penalty)

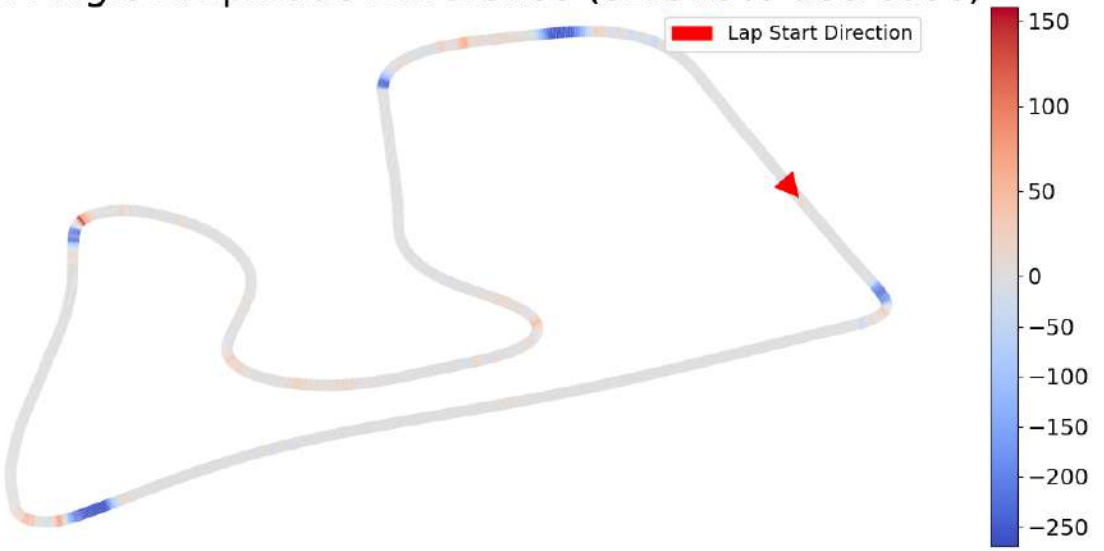
Acceleration Status Difference (9.382% decrease)



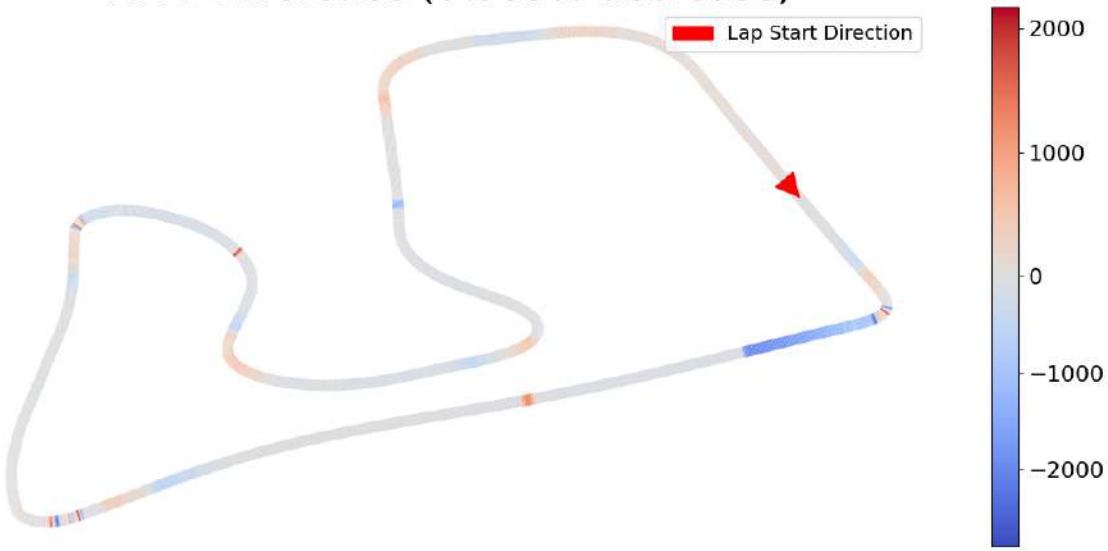
Brake Status Difference (22.641% decrease)



Steer Angle Amplitude Difference (32.915% decrease)



RPM Difference (0.869% decrease)



Key Findings

- Matching and even outperforming baseline
- Maintaining lap time while reducing fuel usage (-10.7% mean)
- Smoother acceleration and steering profiles
- Lowered engine RPM by staying in higher gears
- No rapid breaks due to fuel saving (!)

Limitations

- Tested on two cars and tracks only (narrow sample size)
- No manual gear shifting support
- Limited to 1X simulation speed

Future work

- Expand to different cars and tracks (especially simpler ones)
- Explore manual gear shifting
- Apply different reward function designs
- Apply rewards to other RL-based algorithms or MPC

References

1. https://medium.com/intro-to-artificial-intelligence/soft-actor-critic-reinforcement-learning-algorithm-1934a2c3087f
2. https://assetto-corsa-gym.github.io/
3. https://en.wikipedia.org/wiki/Reinforcement_learning
4. https://www.ibm.com/think/topics/reinforcement-learning

Q&A