CS747: Programming Assignment 3 Report
Optimal Driving Control

Name: Erata Maharshi
Roll No.: 210050049

13 April 2025

1 Problem Statement

In this assignment, we will design an agent to drive a car on various tracks at high speeds. The car gets local information about the track from the environment and has two controls: accelerator and steering. Our task is to develop a policy that maps information from the environment to the car controls. The maximum duration of each episode is fixed; the episode can also end if the car is off the road. The objective of our agent will be to cover as much distance as possible within each episode. Our policy should be generalized to different tracks.

2 Environment Details

- Observation: A 13×13 binary matrix, aligned with the car, where 1's denote lane centers.
- Controls:
  - Acceleration $\in [-1, 1]$ (mapped to $[-5, 5]$ m/s^2)
  - Steering $\in [-1, 1]$ (mapped to $[-\pi/4, \pi/4]$ radians)

3 Policy Design

- Our agent uses a hybrid rule-based policy enhanced by a tunable PID controller and adaptive speed modulation. The final parameters are optimized using CMA-ES.
- The PID (Proportional–Integral–Derivative) controller was chosen for this assignment because it is a well-established, versatile, and effective control strategy commonly used in dynamic systems, especially for tasks involving motion control and path tracking—like autonomous driving.
- P reacts to the current error, I corrects past errors,D sees about future errors. PID provides continuous feedback control.

3.1 Steering Control (PID)

The steering is governed by a PID controller that minimizes the deviation from the lane center. The components are:

$$\begin{aligned}
\text{error} &= \frac{\text{target_x} - 6}{6} \\
\text{integral} &\leftarrow 0.95 \cdot \text{integral} + \text{error} \\
\text{derivative} &= \text{error} - \text{prev_error} \\
\text{steering} &= K_p \cdot \text{error} + K_i \cdot \text{integral} + K_d \cdot \text{derivative}
\end{aligned}$$

To prevent oscillation, we clip the steering output to $[-1, 1]$ and apply windup protection on the integral term.

3.2 Adaptive Speed Control

We modulate the target speed based on Curvature (larger lane offset $\rightarrow$ slower)

$$\text{curvature\_penalty} = \text{clip}\left(\frac{6 - \text{road\_width}}{6}, 0, 1\right)$$
$$\text{lane\_offset} = \left|\frac{\text{target\_x} - 6}{6}\right|$$
$$\text{speed\_target} \leftarrow \text{target\_speed} \cdot (1 - \text{curvature\_penalty} \cdot s_f) \cdot (1 - 0.6 \cdot \text{lane\_offset})$$

The final acceleration is based on normalized speed error:

$$\text{acceleration} = \text{clip}\left(0.6 \cdot \text{error} + 0.4 \cdot \tanh(2 \cdot \text{error}), -1, 1\right)$$

4 Fitness Function

- The fitness function is designed to reward the agent based on three main factors: distance traveled, staying on the road, and completion of the tracks. In addition, penalties are applied for going off-road early or stopping before making significant progress.
- The score is computed by accumulating the ratio of distance to steps, along with a speed term $\frac{\text{speed}}{20}$, provided that the agent stays on the road. If the agent goes off-road, a large penalty is applied. Furthermore, tracks where more than 250 units are covered receive a 1.5x bonus to the score.

5 Parameter Optimization (CMA-ES)

We used CMA-ES to optimize 6 parameters:

$$\theta = [K_p, K_i, K_d, \text{target\_speed}, \text{lookahead}, \text{speed\_factor}]$$

CMA-ES Configuration

- **Generations:** 100
- **Population size:** 6
- **Policy parameters:** 6
- **Initial mean:** Gaussian samples
- **Fitness function:** Total score over 6 tracks

Table 1: Evaluation Results for Various Parameter Sets Across 6 Tracks

Kp	Ki	Kd	T.S	L.A.	S_f	Track 0	Track 1	Track 2	Track 3	Track 4	Track 5
2.4751	0.0312	5.1855	15.92	6.26	0.2004	1064.71	1046.78	1039.26	930.24	1024.74	864.45
2.4751	0.0212	3.1855	15.92	6.26	0.2004	1106.78	1029.06	1008.40	927.21	1020.85	884.43
2.4751	0.0312	3.1855	15.92	6.26	0.2004	1048.76	1032.11	1022.42	915.58	1017.95	880.04
2.4751	0.0375	0.0319	15.92	6.26	0.2004	1055.53	1018.06	994.06	919.72	986.56	183.35
1.6173	0.0312	0.0088	14.37	3.83	0.0429	950.34	931.47	919.89	901.98	876.58	184.14
2.4751	0.0312	2.7949	14.37	6.26	0.2004	938.29	930.24	924.73	922.13	813.91	806.08
2.4751	0.0312	2.7949	14.37	3.83	0.2923	897.48	892.22	882.94	884.19	749.16	778.68
1.6173	0.0312	2.7949	14.37	3.83	0.4292	891.65	884.54	865.76	888.26	778.26	773.99
2.4751	0.0375	0.0519	15.92	6.26	0.2004	1064.01	1025.40	990.16	912.53	186.66	196.70
1.3309	0.0264	0.3317	14.01	5.44	0.1468	910.90	894.99	873.98	836.43	873.98	190.90

Final Learned Parameters

K_p = 2.4751, K_i = 0.0312, K_d = 5.1855, target_speed = 15.9205, lookahead = 6.2583, s_f = 0.2004

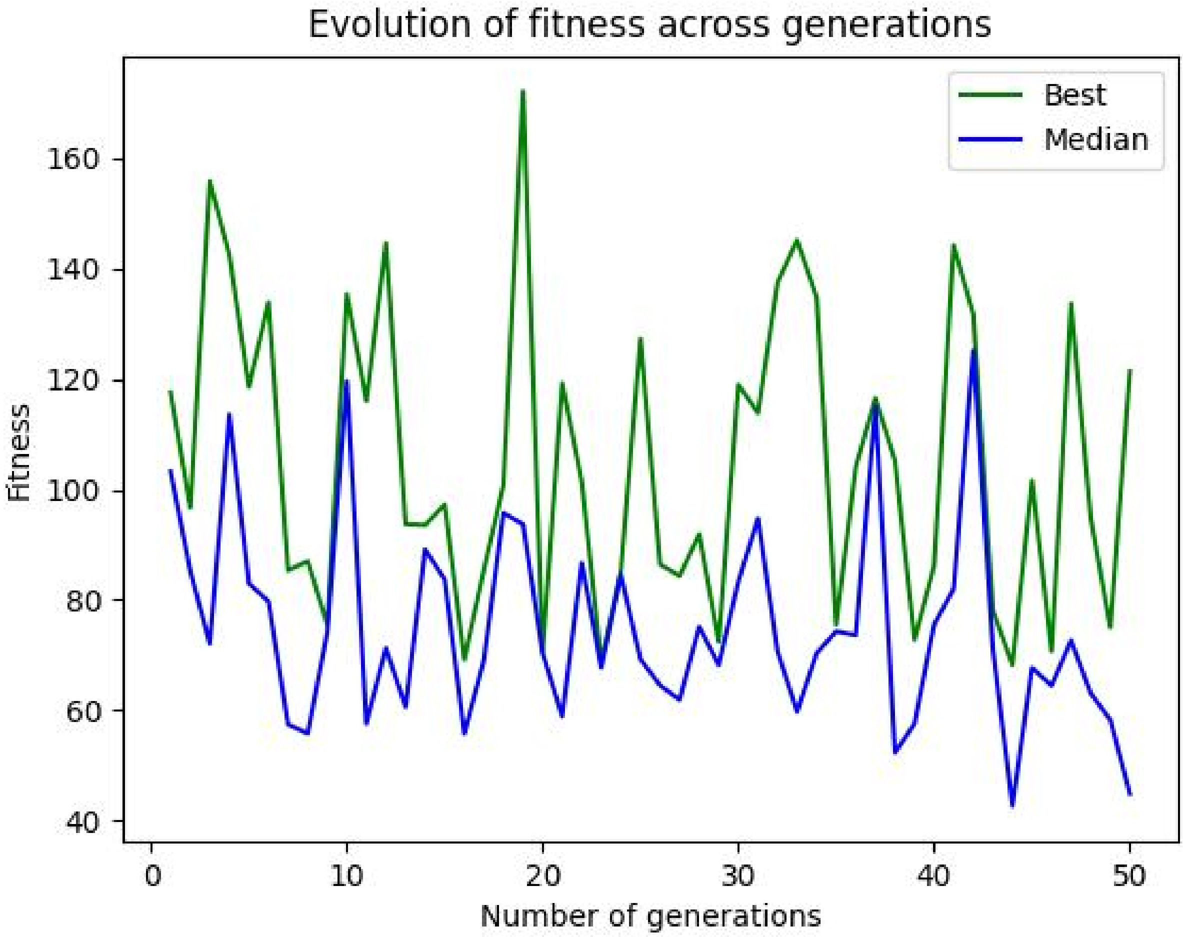

Figure 1: Learning curve with numgen=50 and genpop = 2

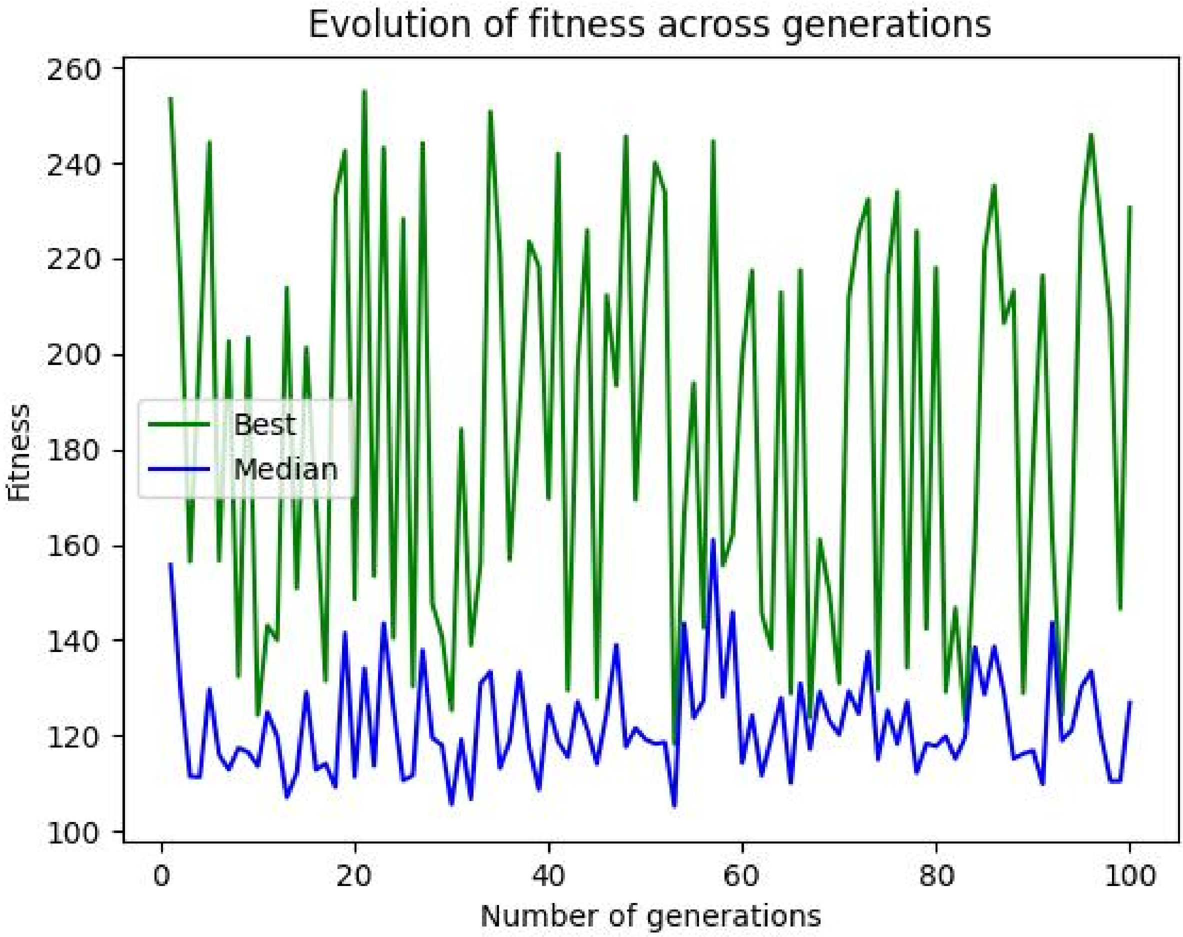

Figure 2: Learning curve with numgen=100 and genpop = 6

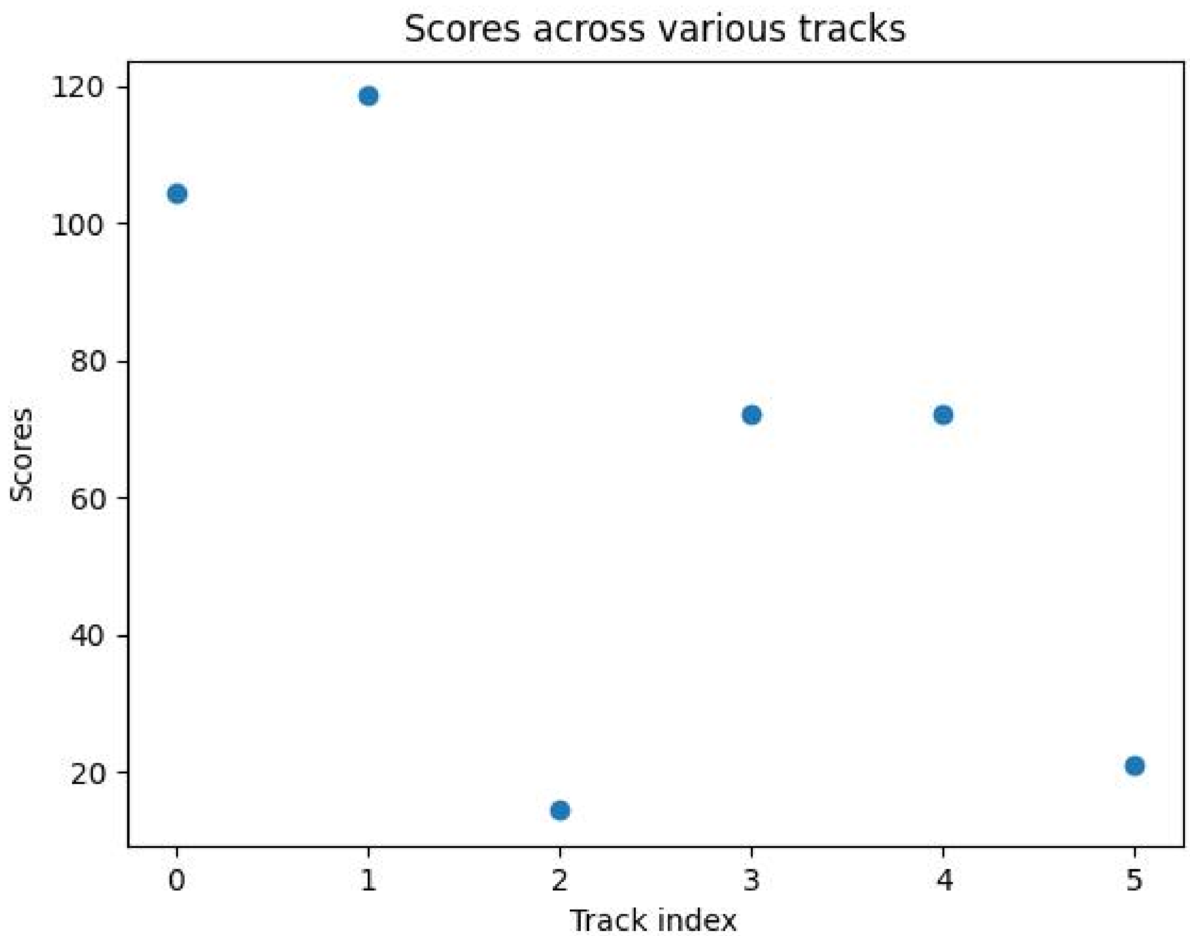

Figure 3: Evaluation scores for numgen=50 and genpop = 2

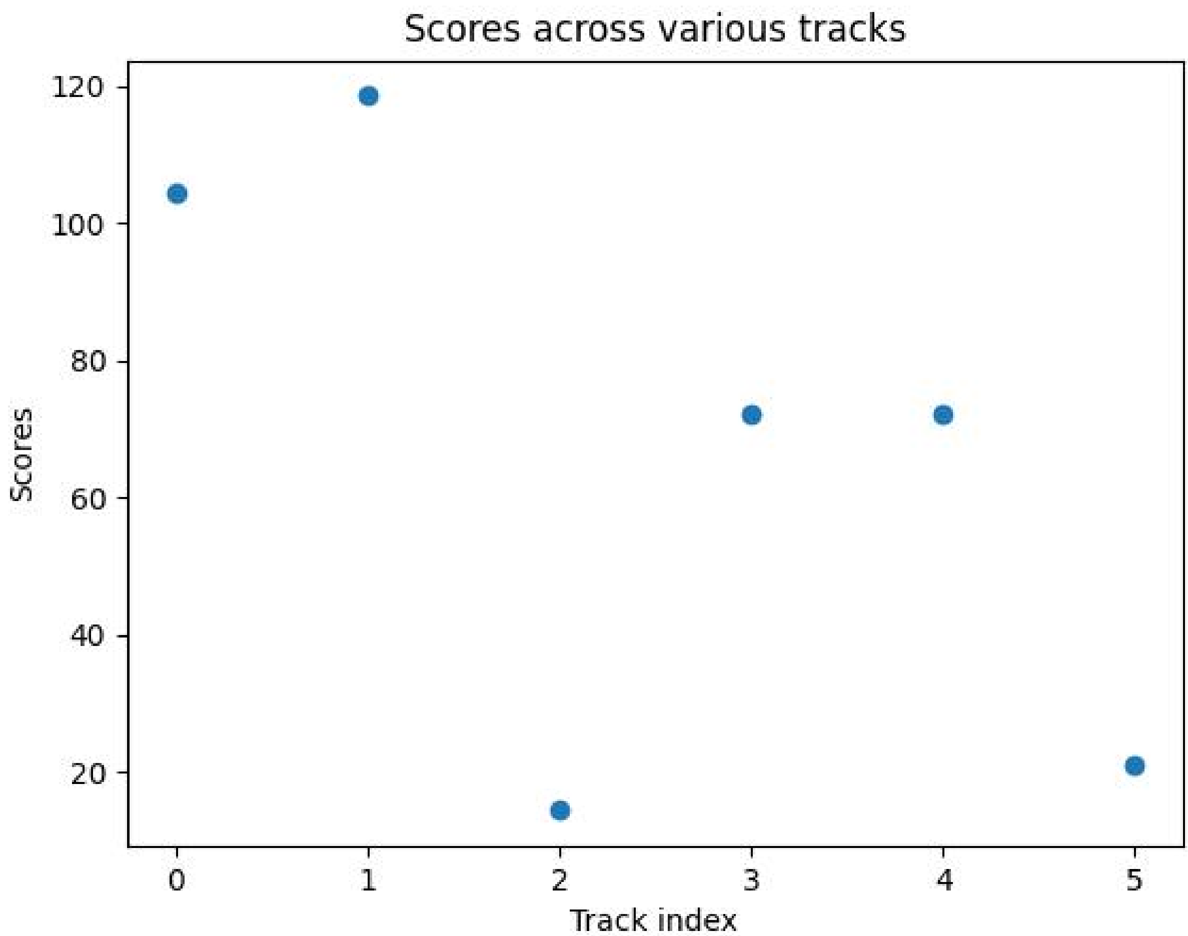

Figure 4: Evaluation scores for numgen=100 and genpop = 6

6 Experiments and Results

- Evaluated on all 6 public tracks
- Policy shows strong generalization to unseen tracks
- Clear improvement with learned parameters vs. manual tuning

7 Challenges and Observations

- Abrupt curves led to frequent exits.
- Adaptive speed control based on curvature greatly improved stability.

8 Conclusion

For the given parameter values. The agent preforms well in evaluation. The table with many values are obtained by stopping the training in middle and then slighly tweaking the values according to PID concepts to attain max scores.

9 References

- `https://en.wikipedia.org/wiki/CMA-ES`
- `https://github.com/Farama-Foundation/HighwayEnv`
- `https://www.cse.iitb.ac.in/~shivaram/teaching/cs747-s2025/pa-3/pa3.html`
- `https://www.cse.iitb.ac.in/~shivaram/teaching/cs747-s2025/lectures/cs747a2022l19.pdf`
- `https://en.wikipedia.org/wiki/Proportional%E2%80%93integral%E2%80%93derivative_controller`
- `https://medium.com/@madhusudhan.d/tuning-pid-controller-for-self-driving-cars-3813f7f18eb0`

Text suggestions provided by Writefull (overleaf) were used in this report, which was compiled using Overleaf.
For writing report, referred overleaf.com.