

# Application of Reinforcement Learning Algorithms for Formula One F1 Vehicle for Route Efficiency

**Authors:** Mr. Muhammad Taimoor Khan (20BSCS045) \*, Mr. Muhammad Hamza (20BSCS011)

**Submitted to:** Dr Sanam Narejo (Associate Professor, CSE, MUET Jamshoro)

**Keywords:** Reinforcement Learning, NeuroEvolution of Augmenting Topologies, Vehicle Simulation, Fitness Based Evaluation

## **Abstract:**

This research focuses on applying reinforcement learning (RL) to improve the route efficiency and performance of an F1 car within a simulation. The simulation is implemented using Python, NEAT (NeuroEvolution of Augmenting Topologies), and PyGame to create an evolving environment where neural networks control the car's navigation. Reinforcement learning refers to a machine learning technique where an agent (the F1 car) learns to make optimal decisions through fitness-based reward system by interacting with its environment. The virtual car receives feedback from its surroundings through sensors (radars) that detect obstacles and measure distances. The RL algorithm adjusts the car's speed and steering based on this data to avoid collisions and optimize movement. Over time, the car improves its driving ability by refining its speed and directional control, ensuring maximum time and distance coverage while avoiding borders that signify crashes. Through a fitness-based evaluation system, the simulation tracks progress by calculating how far and long car travels without crashing. The system updates the neural networks across generations, leading to a measurable increase in lap efficiency and lower lap times. Data collection throughout the simulation, such as the best and average fitness scores, provides insights into the car's evolving performance. The results demonstrate the ability of RL to enhance autonomous driving capabilities, enabling the car to navigate challenging environments and improve decision-making over multiple generation.

# DATA SCIENCE AND ANALYTICS

## Introduction:

Reinforcement Learning (RL) is a branch of machine learning focused on making decisions to maximize cumulative rewards in each situation [1]. Unlike supervised learning, which relies on a training dataset with predefined answers, RL involves learning through experience. In RL, an agent learns to achieve a goal in an uncertain, potentially complex environment by performing actions and receiving feedback through rewards or penalties.

## Area Significance & Importance:

In recent years, the demand for statistics and analyzing data in sports has increased drastically. [2] and the automotive industry is a special industry. Many businesses in the industry, like Google, Tesla, NVIDIA [13], Uber, and Baidu, are committed to creating cutting-edge autonomous vehicles because they have the potential to significantly improve people's lives in the real world. Conversely, several games have seen the successful application of the deep reinforcement learning technique [14] [15]. To keep the passengers' safety, any accident is unacceptable. Therefore, reliability and security must satisfy the stringent standard [3]. As the significance of autonomous driving technologies has surged in recent years, particularly in high-stakes environments like Formula 1 racing. The deep reinforcement learning algorithm's success demonstrates that optimizing policy-guided agents in high-dimensional state and action space can naturally solve control problems in real-world environments [11]. These advancements promise to revolutionize the automotive industry, pushing the boundaries of vehicle performance and safety. By integrating machine learning, specifically reinforcement learning, this research aims to contribute to the growing field of intelligent autonomous systems. However, due to the complexity and unpredictability of our reality, the perception problem is extremely challenging to resolve [11].

## Problems & Challenges:

Research on autonomous driving [12] is ongoing in the fields of control systems and computer vision. Despite the potential benefits, numerous challenges persist in the application of reinforcement learning to autonomous driving. The complexity of racing environments, characterized by dynamic obstacles and unpredictable scenarios, presents significant hurdles. Moreover, ensuring real-time decision-making capabilities is critical for performance, as even minor delays can lead to crashes or suboptimal navigation. Additionally, the balance between exploration and exploitation in the learning process poses a challenge, requiring careful tuning

## DATA SCIENCE AND ANALYTICS

of algorithm parameters to achieve optimal results. Addressing these issues is essential for the successful implementation of RL in high-speed environments.

### **Proposed Methodology:**

This study employs a simulation approach to address the outlined challenges, utilizing the NEAT algorithm to train an autonomous F1 car in a controlled environment. The methodology involves real-time data collection from the car's sensors, which detect distances to obstacles and inform the neural network's decision-making. The simulation leverages PyGame to create a graphical representation of the racing environment, allowing for a comprehensive evaluation of the car's navigation capabilities.

# DATA SCIENCE AND ANALYTICS

## Literature Review:

Author, Year, Title	Research	Outcomes
Evans, B., Engelbrecht, H. A., & Jordaan, H. W. (2021) <b>Reward Signal Design for Autonomous Racing.</b>	This paper investigated reward signal design for reinforcement learning (RL) agents in autonomous racing, aiming to balance speed and crash avoidance. Three reward methodologies are evaluated: position-based, velocity-based (cross-track and heading), and action-based (minimum steering) rewards. A novel approach introduces minimum curvature trajectories as the reward reference, replacing the traditional track centreline. Agents are trained using the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm in a F1/10th autonomous racing simulator, tested under both obstacle-free and obstacle-filled conditions.	The Cross-track and Heading (CTH) reward with minimum curvature paths produced the fastest lap times but resulted in a higher crash rate due to more aggressive driving behaviour. Action-based rewards (minimum steering) improved safety by reducing crashes and increasing completion rates but led to slower lap times. The findings highlight that minimum curvature paths enhance speed performance while increasing crash risks, demonstrating the importance of reward signal design in balancing performance.
Evans, B. D., Jordaan, H. W., & Engelbrecht, H. A. (2023). <b>Safe Reinforcement Learning for High-Speed Autonomous Racing.</b>	This study explored the development of a safe reinforcement learning framework for high-speed autonomous racing that eliminates the need for crashes during training. Their approach integrates a Viability Theory-based supervisory system to ensure vehicles stay within friction limits and avoid unsafe actions. The study employed the Twin-Delayed Deep Deterministic Policy Gradient (TD3) algorithm to train the agents with continuous control, evaluating the method in the F1 Tenth simulator on four racetracks at speeds up to 6 m/s.	The results showed that the supervised agents achieved higher success rates and completed all laps without crashing, demonstrating that the system ensures safety throughout the training process. Additionally, the new framework improved sample efficiency by five times, requiring only 10,000 steps compared to the 50,000 steps needed for conventional methods. However, the supervised agents displayed more conservative behaviour, leading to slower lap times compared to non-supervised agents. This research addresses the sim-to-real gap, making it feasible to apply DRL methods on real-world vehicles.
Cardamone, L., Loiacono, D. and Lanzi, P.L., 2009, May. <b>On-line neuro-evolution applied to the open racing car simulator.</b>	The main research objective was to simulate and evaluate the use of the NEAT (NeuroEvolution of Augmenting Topologies) algorithm in training self-driving cars. The study aimed to create a virtual car simulation that can autonomously learn to navigate through obstacles, including other vehicles, pedestrians, and road features like potholes. The system's goal was to allow the car to evolve through repeated training cycles using neural networks that adjust based on performance. The virtual environment was created using Unity, where cars were equipped with sensors to detect nearby objects.	The cars evolved through generations, learning to avoid obstacles and improving their driving ability with each iteration. The simulation showed that cars became more efficient over time, successfully navigating more complex roadways and avoiding collisions. The project successfully demonstrated that NEAT can be applied to self-driving car simulations, with cars showing significant improvements in their ability to navigate and avoid obstacles over time. This validates NEAT as a powerful tool for developing autonomous driving solutions.

## DATA SCIENCE AND ANALYTICS

Remonda, A., Krebs, S., Veas, E., Luzhnica, G. and Kern, R., 2021. *Formula rl: Deep reinforcement learning for autonomous racing using telemetry data.*

The study explores the application of reinforcement learning (RL) models in autonomous racing, aiming to reduce lap time. The challenge is framed using a continuous action space and multidimensional vehicle telemetry input. Ten deep deterministic policy gradient (DDPG) variants were tested across two experiments to evaluate their effectiveness and generalization to new tracks: (i) how RL methods learn racing, and (ii) how learning conditions affect model generalization. The results show that RL-trained models not only outperform baseline bots in speed but also generalize well to unfamiliar tracks.

The findings demonstrated that, especially on challenging courses, RL algorithms—specifically those that used prioritized experience replay, or PER1M—were able to learn to drive racing vehicles more quickly than hand-crafted models. By optimizing racing lines, the "look ahead of the curve" (LAC) approach significantly enhanced model performance. Like how real drivers practice courses, models taught on difficult tracks performed better when trained on the same track, even if they generalized well to unfamiliar tracks. To improve performance, future research will concentrate on creating general models for unknown tracks, utilizing more sophisticated telemetry data, and comparing inputs based on images.

Tomlinson, S. and Melder, N., 2014. *Representing and driving a race track for AI controlled vehicles. Game AI Pro: Collected Wisdom of Game AI Professionals.*

The study focuses on track representation in AI racing systems. While all systems include a physical track layout and a racing line, the complexity depends on the genre. In arcade-style games, the racing line prioritizes tactical routes with detours, while Formula 1 simulations use a single optimized racing line for speed. For games with diverse vehicles, real-time evaluation of parameters like optimal speed is necessary to handle varying situations. An effective AI solution combines baked-in strategic data with real-time tactical adjustments when computing resources permit.

The outcome of this paper highlights a detailed representation of racetracks and their real-time application to guide AI vehicles at the limits of speed and performance. While this level of complexity is best suited for simulation racing games, the methods can be simplified for arcade-style games. The paper also suggests improving the piecewise linear representation with splines. Further exploration should focus on learning techniques to enhance the racing line and addressing subtle factors like height variations, particularly crests, which can reduce grip by decreasing downward force.

The literature review focuses primarily focused on developing effective reward signal design and safe reinforcement learning frameworks. **Evans et al. (2021, 2023)** investigated different reward methodologies for balancing speed and safety in autonomous racing. They found that while position-based rewards can lead to faster lap times, they may also increase crash rates. The introduction of a supervisory system in their 2023 study helped to ensure safety during training and improved sample efficiency. **Cardamone et al. (2009)** explored the use of neuro-evolution for training self-driving cars in a simulated environment, demonstrating the potential of this approach for developing autonomous driving solutions. **Tomlinson, S. et al. (2014)** used the idea of using track representation rather than real world environment to test tactical detours, and efficiency of the Formula 1 Car in a gaming world.

**In conclusion,** these studies highlight the importance of carefully designing reward signals and incorporating safety mechanisms to ensure the reliability and performance of autonomous racing

## DATA SCIENCE AND ANALYTICS

vehicles. Future research should focus on further improving safety and efficiency, as well as addressing the challenges of transferring these methods to real-world environments.

### Methodology:

This study focuses on developing a simulation for an autonomous car utilizing reinforcement learning through the NEAT (NeuroEvolution of Augmenting Topologies) algorithm. The objective is to train the car to navigate a defined environment effectively while avoiding obstacles.

**Data Collection:** The simulation collects real-time data from the car's sensors, which provide information about the distances to the nearest obstacles. This data serves as input for the neural network that controls the car.

**Model Description:** The model consists of a neural network that is trained using NEAT, where the structure of the network evolves over generations. Each car in the simulation represents an individual in a population, and its performance is evaluated based on its ability to navigate the track without colliding with the boundaries. It is often enough to have a simple reward function as the algorithm itself figures things out as it goes along [4]. In addition to updating weights in the NEAT algorithm, the network architecture changes over generations and becomes more complex. [5]

**Model Definition:** The neural network takes sensor readings as inputs and outputs commands for steering and speed adjustments. The architecture of the neural network is adapted through NEAT's evolutionary strategies, promoting networks that exhibit improved performance over time.

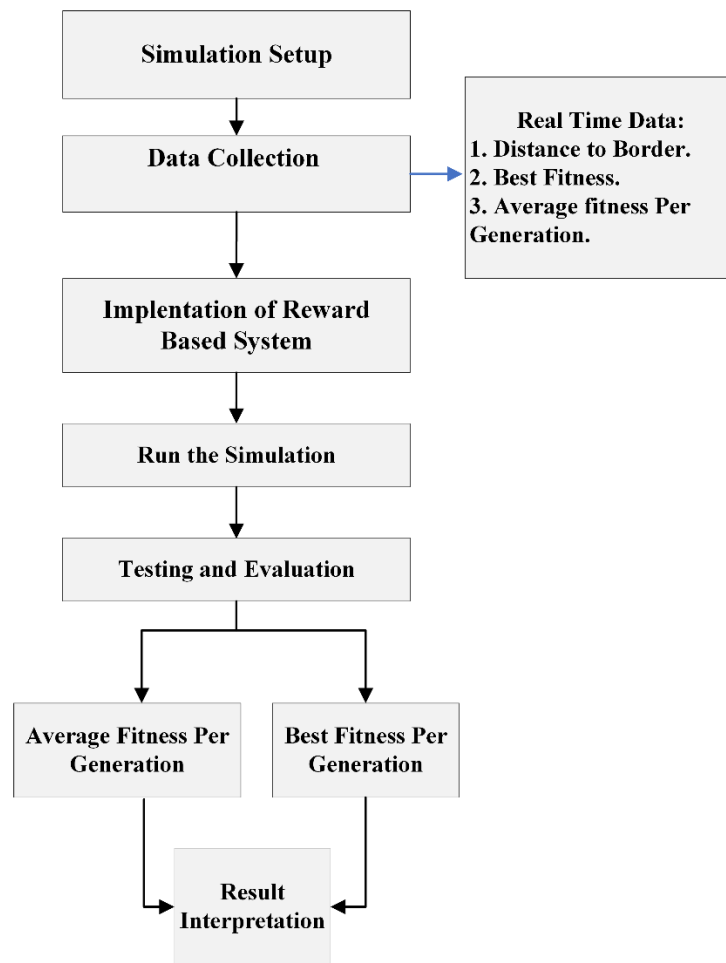
### Mathematical Equations:

$$\delta = \frac{c_1 E}{N} + \frac{c_2 D}{N} + c_3 \cdot \overline{W}$$

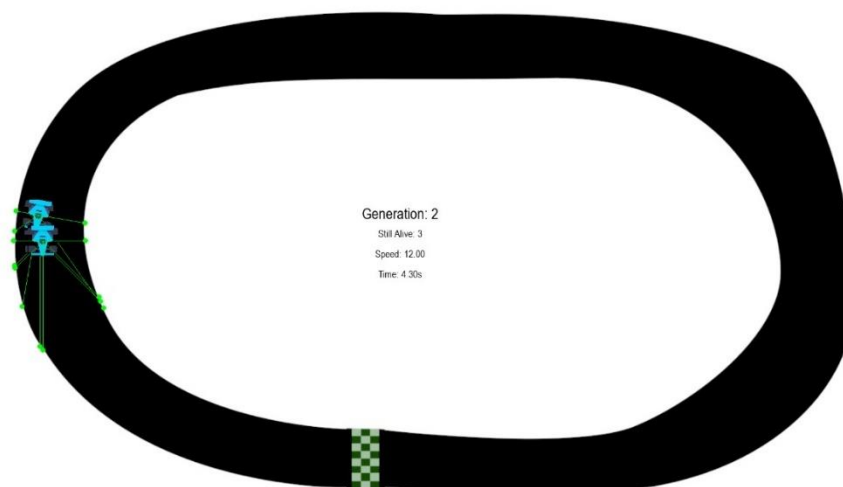
$$f'_i = \frac{f_i}{\sum_{j=1}^n sh(\delta(i,j))}$$

# DATA SCIENCE AND ANALYTICS

## Flowchart:

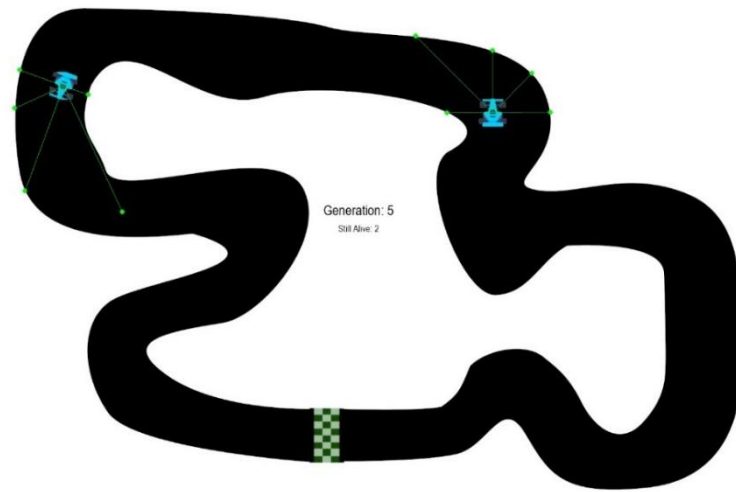


**Experimental Setups:** The simulation is set up using PyGame to create a graphical environment, with a population size of 100 cars per generation. Each car's neural network is evaluated over a maximum of 20 generations, allowing for gradual improvement in performance. The configuration parameters for NEAT include mutation rates, selection methods, and fitness evaluation metrics to ensure effective training of the neural networks.

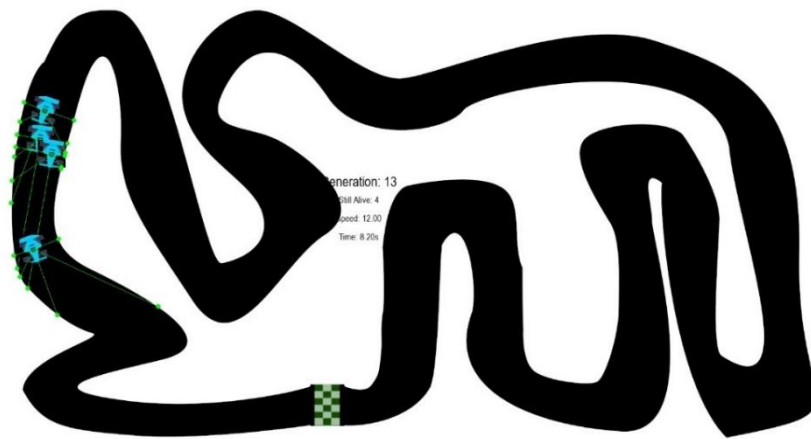


Map 1: Racetrack

## DATA SCIENCE AND ANALYTICS



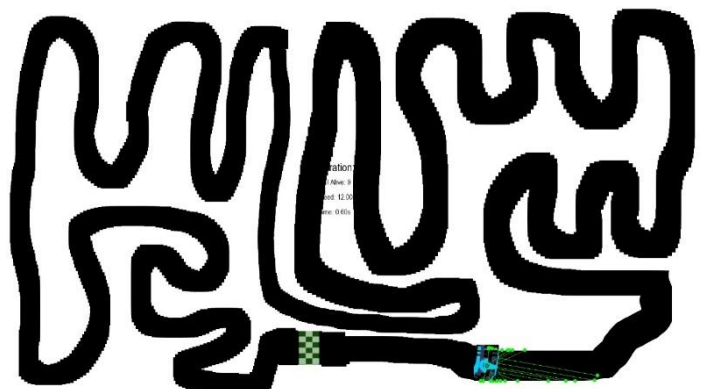
Map 2: Racetrack



Map 3: Racetrack



Map 4: Racetrack



Map 5: Racetrack



## Results:

The experiment used a genetic algorithm to train a neural network to control a car in a simulated environment. The fitness score of the car was used to measure its performance. **As shown in Figure 1 (a) and Figure 1 (b)**, both the best fitness (solid line) and the average fitness (dashed line) increased over the course of the simulation, which demonstrates that the neural network was successfully learning to control the car."

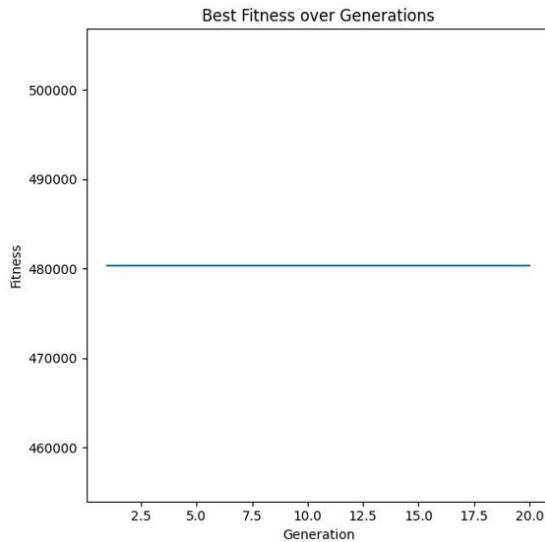


Figure 1 (a)

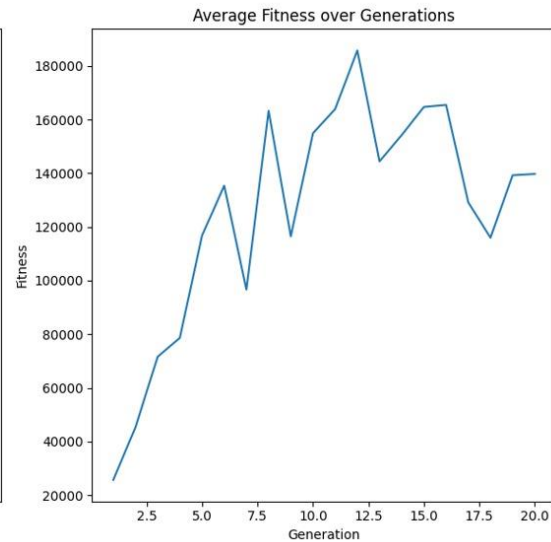


Figure 1 (b)

**In map 2**, the neural network's performance exhibited a distinct pattern of plateaus followed by sudden increases in fitness, **as shown in Figure 2 (a)**. This suggests that the network was encountering challenges in adapting to the new environment, but was able to overcome these challenges through occasional breakthroughs. The average fitness also showed a similar pattern **as shown in Figure 2 (b)**, although the overall improvement was less pronounced, indicating that the population as a whole was struggling to adapt as quickly as the best performers.

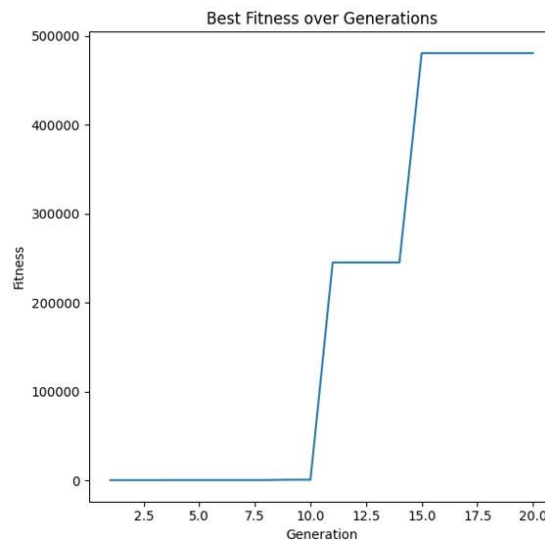


Figure 2 (a)

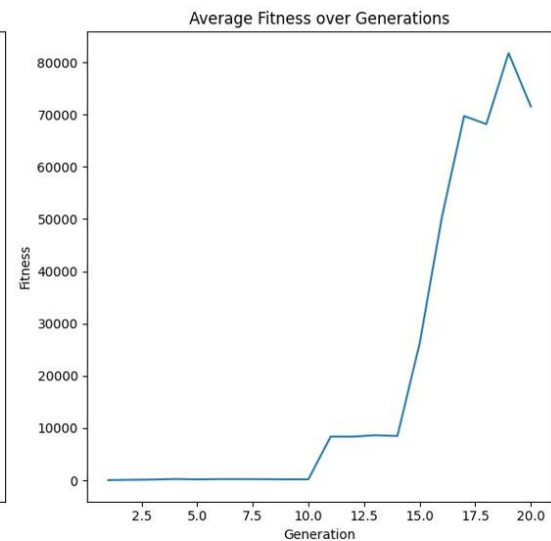


Figure 2 (b)

## DATA SCIENCE AND ANALYTICS

**In map 3**, the neural network's performance exhibited a similar pattern of plateaus and sudden increases **as observed in Figure 3 (a)**. This suggests that the challenges encountered by the network were similar in both environments. However, the overall level of fitness achieved was lower on map 3, indicating that the environment was more difficult to navigate. The average fitness also showed a similar pattern **as seen in Figure 3 (b)**, but with a less steep increase, suggesting that the population as a whole struggled to adapt as quickly.

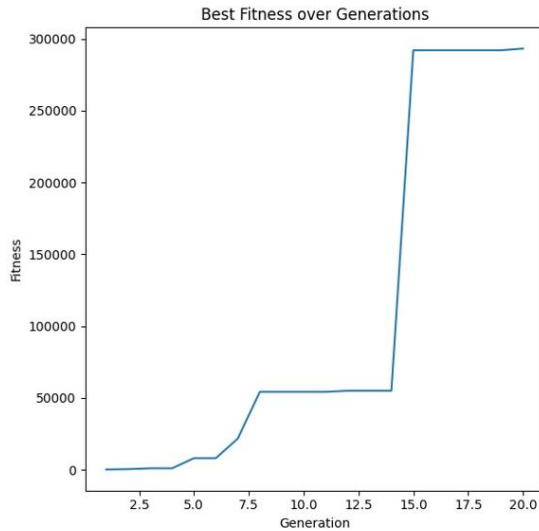


Figure 3 (a)

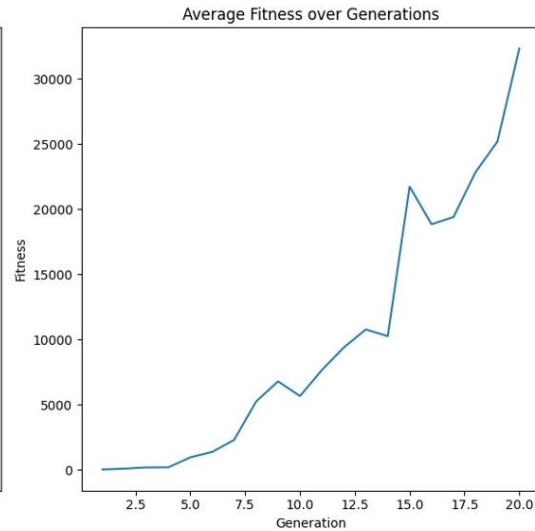


Figure 3 (b)

**In map 4**, the neural network's performance exhibited a similar pattern of plateaus and sudden increases which **is shown in Figure 4 (a)** and as observed with the previous maps. This suggests that the challenges encountered by the network were consistent across different environments. However, the overall level of fitness achieved on map 4 was significantly lower than on the other maps, indicating that it was the most difficult to navigate. The average fitness also showed a similar pattern, but with a less steep increase, suggesting that the population struggled to adapt as quickly, this **can be observed in Figure 4 (b)**.

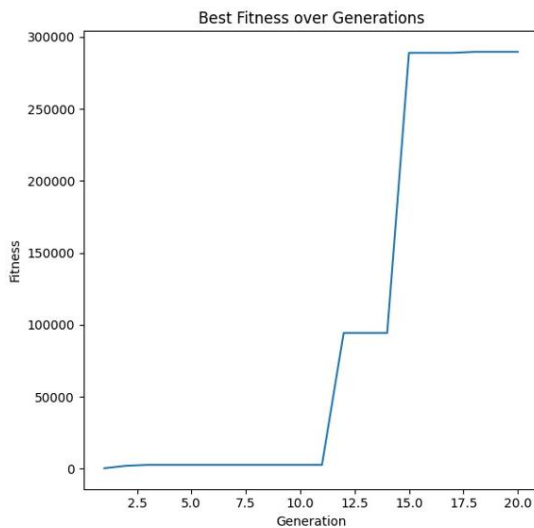


Figure 4 (a)

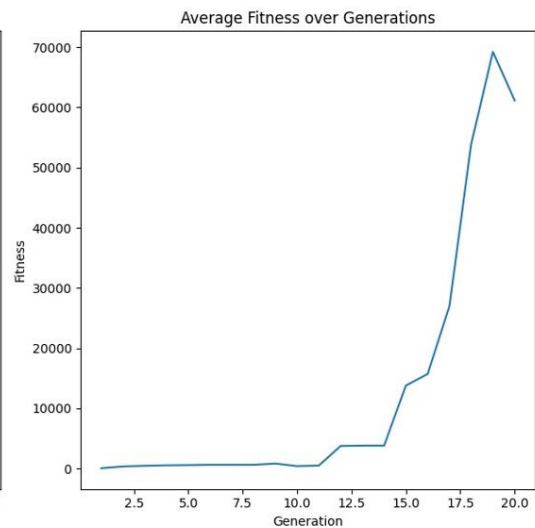
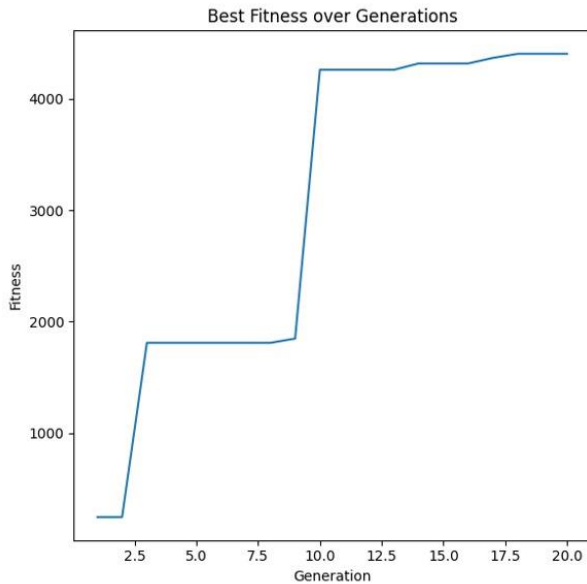
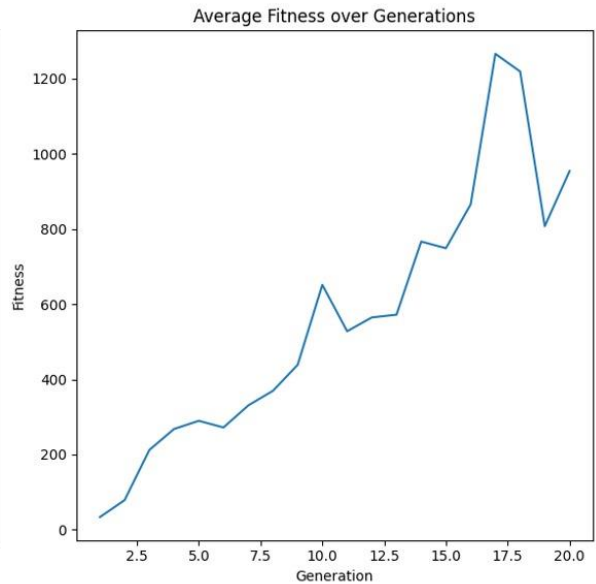


Figure 4 (b)

**In map 5**, the neural network's performance exhibited a similar pattern of plateaus and sudden increases as observed with the previous maps. This suggests that the challenges encountered by the network were consistent across different environments. However, the overall level of fitness achieved on map 5 was slightly higher than on map 4 which can be shown in **Figure 5 (a)**, but still lower than on maps 1, 2, and 3. The average fitness also showed a similar pattern, but with a less steep increase, suggesting that the population struggled to adapt as quickly as shown in **Figure 5 (b)**.



**Figure 5 (a)**



**Figure 5 (b)**

### Conclusion:

This research has effectively showcased the application of reinforcement learning, particularly through the NEAT (NeuroEvolution of Augmenting Topologies) algorithm, to enhance the navigation capabilities of an autonomous F1 car within a simulated environment. The results highlight how integrating reinforcement learning not only improves decision-making but also fosters adaptive learning in the face of complex and dynamic racing scenarios. Over generations, the simulation demonstrated a consistent upward trend in fitness scores, indicating that the neural networks were able to refine their strategies for obstacle avoidance and route optimization.

### References:

- [1] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. A. Riedmiller, "Playing atari with deep reinforcement learning," CoRR, 2013.
- [2] Heimbs, S., Strobl, F., Middendorf, P., Gardner, S., Eddington, B. and Key, J., 2009, May. Crash simulation of an F1 racing car front impact structure. In *7th European LS-DYNA conference* (pp. 1-8).
- [3] Zhang, Q., Du, T. and Tian, C., 2019. Self-driving scale car trained by deep reinforcement learning. *arXiv preprint arXiv:1909.03467*.
- [4] Andersson, M., 2022. "How does the performance of NEAT compare to Reinforcement Learning?"
- [5] Khamesian, S. and Malek, H., 2024. Hybrid self-attention NEAT: a novel evolutionary self-attention approach to improve the NEAT algorithm in high dimensional inputs. *Evolving Systems*, 15(2), pp.489-503.
- [6] Evans, B., Engelbrecht, H.A. and Jordaan, H.W., 2021, December. Reward signal design for autonomous racing. In *2021 20th International Conference on Advanced Robotics (ICAR)* (pp. 455-460). IEEE.
- [7] Evans, B.D., Jordaan, H.W. and Engelbrecht, H.A., 2023. Safe reinforcement learning for high-speed autonomous racing. *Cognitive Robotics*, 3, pp.107-126.
- [8] Cardamone, L., Loiacono, D. and Lanzi, P.L., 2009, May. On-line neuroevolution applied to the open racing car simulator. In *2009 IEEE Congress on Evolutionary Computation* (pp. 2622-2629). IEEE.
- [9] Remonda, A., Krebs, S., Veas, E., Luzhnica, G. and Kern, R., 2021. Formula rl: Deep reinforcement learning for autonomous racing using telemetry data. *arXiv preprint arXiv:2104.11106*.
- [10] Tomlinson, S. and Melder, N., 2014. Representing and driving a race track for AI controlled vehicles. *Game AI Pro: Collected Wisdom of Game AI Professionals*. Boca Raton: CRC Press/Taylor and Francis Group, pp.481-490.
- [11] Sallab, A.E., Abdou, M., Perot, E. and Yogamani, S., 2017. Deep reinforcement learning framework for autonomous driving. *arXiv preprint arXiv:1704.02532*.
- [12] Q. Memon, M. Ahmed, and S. Ali. Self-Driving and Driver Relaxing Vehicle. 2016.

## DATA SCIENCE AND ANALYTICS

- [13] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, et al. End to end learning for self-driving cars. arXiv preprint, 2016.
- [14] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller. Playing atari with deep reinforcement learning. NIPS, 2013.
- [15] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al. Human-level control through deep reinforcement learning. Nature, 2015.