

**Data Visualisation:  
AI Agent for Car Racing Game**

Version: 3.0  
Author: José Henrique Nóbrega Pereira  
Fontys ICT  
Date: May 30, 2025

|  |
| --- |
| A black background with purple text  AI-generated content may be incorrect. |

**Introduction**

After an AI is trained the result is a model which the performance can be observed and judged by humans, but this will result in subjective evaluations of how good the AI really is. A way to solve this and rate the AI objectively is to have a score that reflects the AI performance, but even then, just looking at the numbers will not end up with significant conclusions.

That is when Data Visualization comes in, this involves transforming the data into graphs and other graphic representations so that humans can easily look at it and extract meaningful information, turning abstract data into insightful graphs.

For my model, this representation will allow me to better understand the training process, how it is behaving, and how certain changes I do impact the training process. Although the setup consisted of some python scripts and excel, it greatly improved my insight and expertise on how machine learning works. This document will reflect and explain how the training of the model went and what can be concluded from it.

**First training**

The first training consisted of 1000 episodes, where the AI was free to learn and explore the environment as it wished, the graph below shows the reward level the AI achieved across the 1000 episodes.

A graph showing a blue line

AI-generated content may be incorrect.

Figure 1. Graph of the reward data across the first 1000 episodes

As you can see the graph does not look like a line, as you might have expected, but a very up and down line, with sudden increases and decreases in performance. This is because the AI when training is constantly experimenting with new ideas and techniques to try and improve in the future. Some of these work, hence the high reward values, but clearly some also do not, as seen with the low reward values. Overall, it is still possible to make out that the AI is improving overtime with higher highs and lower lows.

But it would be nice to have a steady line that is easier to follow and understand the AI performance, for that we plot on the same graph a moving average. This is an average of the last 100 episodes, which means that only the performance of the last 100 episodes is taken into account when plotting the line in a certain point. This creates a nice line that does not spike as much when there are episodes with a reward value much higher or much lower than the average.

The graph below shows this line in red, we can conclude from this that the model had a below zero reward in the first few episodes, a steady improvement across the 1000 episodes and some breakthrough discoveries at around the 175 episode and 450-episode.

A graph showing a graph of a graph

AI-generated content may be incorrect.

Figure 2. Graph of the 1000 reward values (Blue) and the 100-episode mean average (Red). When comparing both lines, the red line is much easier to look at and judge exactly how the AI improved across the training process.

**Training 2**

The average final performance was still mediocre when compared of what it is capable, with an average a bit higher than 100, but a best episode of almost 600. So, a second training was started, this time for 2000 episodes, to try and push the AI to get better and make sure if there is more performance to get out of the AI.

The graph below shows the moving averages of the first training (orange) compared to the second training (yellow). It is visible that the yellow line achieved much higher rewards due to the increase in episodes number.

A graph with orange lines

AI-generated content may be incorrect.

Figure 3. Graph comparing the 50-episode average of the first training (Orange) to the second training (Yellow). Up to the 1000 episode mark (where training 1 stopped) the steepness and the value at episode 1000 are very similar.

Comparing both graphs some conclusions can be taken:

* At the 1000-episode mark, both trainings have similar average reward value.
* Both training sessions had a negative reward at first before having a breakthrough that pushed them into the positive rewards territory.
* Even after 2000 episodes, the line is trending upwards, indicating that there is more performance to be extracted (convergence not hit).

**Final Training**

Since the previous graph clearly showed that the AI was still improving when it ended, a final training was started, this time it would only stop when convergence was hit and the performance stagnated. The hyperparameters were also updated to allow a bigger number of changes after each episode to the AI behaviour, which should lead to a more unstable training but also faster.

The AI ended up training for 7500 episodes, the graph below shows the moving average of the last training in green, compared to the other two training sessions. The line is much steeper when compared to the other two trainings, and at the 1000-episode mark, where the other trainings were at around 130 average reward, this last training was already above 600. The performance that the second training achieved in 2000 episodes was matched by this last training in only 600 episodes, showing the power of correctly tuning the hyperparameters and how they can optimize the learning process.

A graph with green line

AI-generated content may be incorrect.

Figure 4. Graph comparing the 50-episode moving averages of the first training (Orange) and second training (Yellow) to the third training (Green). Green line is much steeper due to changes in the hyperparameters before the training, allowing more learning freedom to the AI when compared to the previous two trainings

The number of episodes trained for seems to be also excessive, since after the 4500-episode mark, there is no substantial improvement and the performance stays at about the same level for the next 3000 episodes. This is because the convergence checking logic was too strict, stopping only when the average reward changed less than 1% ten times in row, checking it once every 50 episodes.

After training being complete, the AI performance can be compared to a Human player side by side. This allows us to check if the AI can really be better than a Human player and if some improvements can be taken from the model’s way of playing.

From the video, some conclusions can be taken about the main differences from the AI and the Human player shown in the following table.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Excellent  4 | Good  3 | Acceptable  2 | Unacceptable  1 | Human  Player | AI  Player |
| Straights | During the Straights the throttle is at maximum application and does not go to the grass | Keeps the car in a straight line, but does not apply full throttle | Can keep the car on track, zig zags during the straights | Can not keep the car in a straight line a goes off track | 4 | 3 |
| Fast Corners | Makes the corners with minimal to no braking at great speed | Makes the corner with some braking and at slow speed | Too much braking, almost coming to a full stop and makes the corner | Does not make the corner and goes off track | 4 | 3 |
| Slow Corners | Makes the corners with minimal to no braking at great speed | Makes the corner with some braking and at slow speed | Too much braking, almost coming to a full stop and makes the corner | Does not make the corner and goes off track | 2 | 4 |
| Start | Immediately applies full throttle | Stars slowly and gradually increases the speed | The start is slow and take too much time getting up to speed | Does not go off the line | 4 | 3 |
| Out of track | Never goes off track | Goes off track occasionally but can come back to the track | Goes off track sometimes and cannot come back | Goes off track in the first corner | 3 | 4 |
| Total |  |  |  |  | 17 | 17 |

From the table and by observing the video comparison some more statements can be made:

* The AI has a much more consistent speed along the whole lap
* The AI has much more control over the car, being able to do smooth inputs
* The Human player is on a keyboard, so the inputs are digital, maybe with a controller or other input devices that would allow the input to be analog there would be more control over the car
* The AI does not go full throttle often, has a slow start and does not maximize speed on the straights

The image below is a screenshot from the video comparing the two performances.

Screens screenshot of a video game

AI-generated content may be incorrect.

**Conclusion**

This document presents a detailed exploration of how data visualization can support the training and evaluation of a reinforcement learning-based AI agent for a car racing game. Through successive training sessions, the analysis shows how various adjustments—including training duration and hyperparameter tuning—significantly influenced the agent’s learning trajectory. The use of moving averages and comparative graphs enabled clear insights into the AI's development, uncovering both strengths and limitations of the training process. By combining quantitative data with qualitative comparisons between human and AI performance, the document not only illustrates the effectiveness of the training methods used but also highlights the value of visualization as a crucial tool for understanding and improving AI systems. Overall, it demonstrates how thoughtful analysis of training data can inform smarter AI development and foster a deeper understanding of machine learning behaviour in practice.

On a personal level, data visualization turned out be more than just plotting some graphs, it was a tool to help better understand what is going on during the training and give insight on what could be done to improve the AI. Watching the average line move up and down made it feel real and tangible. I now see Data visualization as not only a way to present data, but also a way to reflect and understand the AI model.