

**Data Preparation: Training Data to Moving Average  
AI Agent for Car Racing Game**

Version: 2.0  
Author: José Henrique Nóbrega Pereira  
Fontys ICT  
Date: May 5, 2025

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

**Introduction**

For the project of this semester, an AI model was trained using Proximal Policy Optimization (PPO), a popular reinforcement learning (RL) algorithm. After the training is done, a pickle file with all the episode rewards is saved. Looking at this data is possible but would probably mean nothing and will not help providing conclusions about the training process. So, this data needs to be extracted, formatted and analysed to obtain meaningful insights from it.

This document describes the process from going to the pickle file obtained from the training to meaningful data that can be plotted and give insight on the training itself. This involved playing with multiple file formats, making some python scripts and excel calculations.

**Training 1 & 2: Python Plotting and Data Extraction**

For the first two trainings, one consisting of 1000 episodes and the other of 2000 episodes, seeing the performance overtime of the model is something important to be able to make some conclusions about the process. For this a python script (appendix 1) is written that plots all the reward on a graph as seen in the picture below. It is visible that the model has progressed overtime, but it is very “noisy” since the AI is always trying new things and these can be very good or very bad, leading to this huge range in reward.

A graph showing a blue line

AI-generated content may be incorrect.

It is hard to make out how much the AI is really improving looking at that graph, so the next step is calculating a moving average, this will allow to better see the improvement by providing a line that does not spike this much. The moving average chosen in the picture below is a 100-episode moving average, which means that each episode, the average of the last 100 episodes is calculated.

From the picture below it is much easier to see the steady improvement that took place during the training, and where the AI made breakthrough discovery on how to get more reward when there are step increases in the average reward. (e.g. around the 450-episode mark)

A graph showing a graph of a graph

AI-generated content may be incorrect.

Now what would be useful is plotting both trainings in the same graph, for that an excel sheet is used, this means that the data from the pickle file needs to be extracted. To get the data into excel, another python script (appendix 2) is made that prints the reward results in the console, these numbers are then put into Claude that reduces all the numbers to 3 decimal, numbers them and puts them into 2 columns, afterwards these can simply be plugged into an excel sheet.

With the necessary data plugged in, the formulate AVERAGE(B2:B51) can be used to calculate the moving average, this time a 50-episode moving average, to have a graph that is more responsive and better visualise the changes in performance. The picture below shows the comparison between the two training sessions.

A graph with orange lines

AI-generated content may be incorrect.

**Instance 3: Automating with Python with Excel**

The last training session was run for 7500 episodes, until convergence was hit, this resulted in a much bigger pickle file, that could not be processed by Claude due to the character limit in a chat. This meant that a new solution to extract the data into excel was needed.

A new python script (appendix 3) is then made, making use of the openpyxl library which allowed to extract the reward values into a column and added another column to number them. From the generated excel sheet it is copied to the excel sheet with the comparison of the two first trainings, the moving average is calculated for the new data and the graph below is the result.

A graph with green line

AI-generated content may be incorrect.

**Conclusion**

Through these three training sessions, I had my first experience with data preparation and how it can affect what information we can extract from the AI performance. Extracting the reward history, formatting it for analysis plotting them showed how important it is to have a good data gathering method and cleaning the data for the purpose we want.

This did not only improve my technical understanding of data but also outlined that an AI model is as important as the information we can extract, understand and analyse from it.

Some key takeaways:

* Automate as much as possible: Manual steps are useful for tasks that do not need to be repeated a lot. Beyond that, automation reduces time consumption and improves reproducibility.
* Future Idea Improvement: A script that directly converts the pickle file to an excel, including automatic moving averages and charts.

**Appendix**

A screen shot of a computer program

AI-generated content may be incorrect.

Appendix 1. Plot graph from pickle file code

A screen shot of a computer program

AI-generated content may be incorrect.

Appendix 2. Print pickle file content in terminal code

A screenshot of a computer program

AI-generated content may be incorrect.

Appendix 3. Convert reward list to excel sheet code