Andrew Vue

SNHU CS370-H7429

Module 5 Assignment: Cartpole Problem

Initial run

Start of output

Graphical user interface, text, application, email

Description automatically generated

End of Run output

Graphical user interface, text, application, email

Description automatically generated

I made changes to the Exploration Min from 0.01 to 0.1

Text

Description automatically generated

Start of output after changes of Exploration Min.

Graphical user interface, text, application, email

Description automatically generated

End of Output:

At 799 it showed an average score of 101.05 which is one of the highest I found, but then it begins dropping again until the average came back down to about 90.29 at run 895. The highest I saw was at run 789 with a score of 322. The score started decreasing after this point.

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application

Description automatically generated

**Assignment Questions:**

Cartpole is a situation where a pole with a heavy object is placed on the top of a pole. This pole is then placed into a cart that can move freely across a plane. The goal of the situation is to keep the pole in balance for as long as possible. In the beginning of the simulation, the program knows nothing about this situation and by the end it should know how to maintain the perfect balance of the pole.

**States:** The condition of the pole on the cart can be considered a condition of the situation. Having the pole balanced is the desired end state. This state can be read as pixels on the screen, or as information about the pole or cart, including movements and angles. Each frame the state of the simulation will change because the cart is needing to move to maintain balance. “Suppose the pole is starting straight, if we go left, the pole is mostly to go right, which is a new state. Therefore, during each time-step, any action we make will always lead to a different state [(Reinforcement Learning Concept on Cart-Pole with DQN | by Vitou Phy | Towards Data Science](https://towardsdatascience.com/reinforcement-learning-concept-on-cart-pole-with-dqn-799105ca670)).”

The program we use provides the application with past experiences and values that are used for the next runs. The scores of the current run affect the model because it can change to help solver the problem or hinder it. The discount factor is used and helps with the changes being made with the current run of the problem. The discount factor can lead to improvements, or it can negatively impact the learning. Changing the values can make for shorter or longer times to sole the problem. The larger the value of the learning rate the longer it will take to solve.