Afshin E. Ahvazi

**Professor Jason Richard**

CS 370 11205 M01

June 29, 2024

**Project 2: Design Defense**

In this project, we use the Google TensorFlow and Keras libraries to train a Q-learning model to play our custom pirate game. The objective of the game is to reach a designated target destination by navigating through an environment. This environment is represented by an 8x8 matrix, and it is filled with obstacles. To solve this problem, we must design a Q-training model that uses reinforcement learning, where the agent learns to make decisions by interacting with the environment and receiving either rewards for winning or penalties for taking a wrong turn.

Keras Q-learning utilizes a neural network to approximate the Q-function. This function estimates the expected cumulative reward for taking a specific action in a given state, facilitating the agent's decision-making process (Choudhary, 2023). During each epoch, the agent explores the environment using an epsilon-greedy equation to find an optimal choice between exploration, i.e. randomly discovering new paths, and exploitation, choosing a path based on previously learned experiences. This way, the model continuously learns to navigate through the maze.

To solve this problem, humans begin by understanding the objective of the game and how to win it. We would then study the board and locate the starting point, the destination, and identify potential winning patterns. As we move through the maze, we can look a few steps ahead for a winning path and backtrack our way back to the current state of the board. Through trial and error, we humans learn from our mistakes until the game is won.

Both humans and machines use a trial-and-error approach to explore different paths and learn from their mistakes. A common characteristic of intelligent agents is that, over time, they adapt their strategies based on the feedback they receive, thus making learning possible. Humans use their senses and cognition to input and process information, and have great ability to generalize problems and adapt based on changing circumstances (Kaushik, 2023). Machines, on the other hand, use models and predefined algorithms. In addition, machines can learn much faster because they are not limited by memory, processing power, or other factors such as distraction, human emotions, etc.

**References**

Choudhary, A. (2023, August 21). *A hands-on introduction to deep Q-learning using openai gym in python*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2019/04/introduction-deep-q-learning-python/>

Kaushik, H. (2023, September 21). *Machine learning vs human learning: The battle for future dominance*. Medium. <https://medium.com/@himanshubangalore/machine-learning-vs-human-learning-the-battle-for-future-dominance-fa7a1c99cd0c>