Humans and machines solve problems in different ways. A human can use logic, memory, and sometimes intuition to figure out a solution. People can picture the maze, test a path, and turn back if it does not work. A machine does not “think” this way. Instead, it follows a set of rules and learns from feedback to slowly improve its choices.

If a person were in the maze, they would probably look for open spaces, walk one way, and backtrack if they hit a wall. They would repeat this until they reached the treasure. My agent works differently. It starts in the first cell, checks what moves are allowed, and then chooses one. Each move gives it either a reward or a penalty. The agent remembers these outcomes, and over time, it figures out the path that leads to the treasure more directly.

Both humans and the agent learn from trial and error, but the agent does it through numbers and rewards instead of memory and intuition. Humans can sometimes guess a shortcut, but the agent must play the maze many times before finding the best route.

The point of having this agent is to make sure the treasure can be reached in a reliable way. It is built to explore options, avoid wasting time, and use what it learns to improve. In reinforcement learning, *exploration* means trying out new moves, and *exploitation* means sticking with the move it already knows works best. For this maze, it is good to start with more exploration so the agent learns the maze, then slowly move toward more exploitation to use what it has learned.

Reinforcement learning helps because it connects actions with rewards. The pirate agent gets feedback for every step it takes. Over time, it learns which paths give bad results and which ones get closer to the treasure.

For this project, I used deep Q-learning. Instead of saving Q-values in a table, I built a neural network. The network takes in the state of the maze and gives values for each action. During training, the agent updates the network by comparing its predictions with the actual rewards. This lets the agent improve and find the treasure more efficiently.

**References**

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., et al. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529–533. https://doi.org/10.1038/nature14236

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.