

**Subject: Current/ Emerging Trends**

**Topic: Project Two**

**Hamza Malik**

**Table of Content**

[Introduction 2](#_Toc138377699)

[Analysis 2](#_Toc138377700)

[Differences between Human and Machine Approaches to Problem Solving: 2](#_Toc138377701)

[How to Solve Human Problems in the Maze: 3](#_Toc138377702)

[Pathfinding Problem Solution by the Intelligent Agent: 3](#_Toc138377703)

[Comparing and contrasting human and machine approaches 4](#_Toc138377704)

[Purpose of the Intelligent Agent: 5](#_Toc138377705)

[References: 7](#_Toc138377706)

# Introduction

The goal of the Treasure Hunt game, an engaging maze navigation task, is to lead a pirate to a hidden treasure within a maze. In this project, we created an intelligent agent that can solve the Treasure Hunt game using the Q-learning method.

The game's setting is a grid-based, two-dimensional maze with specific places for the treasure, walls, and obstacles. The intelligent agent can manoeuvre the pirate in four directions: up, down, left, and right. The agent's objective is to find his way out of the maze, dodge traps and walls, and eventually find the treasure.

# Analysis

## Differences between Human and Machine Approaches to Problem Solving:

* **Cognitive Skills:** Complex cognitive skills including logic, intuition, creativity, and prior knowledge are essential to human problem solving. On the other side, machines lack creativity and intuition and instead process problems using algorithms and processing capacity.
* Humans are capable of learning from their mistakes, adapting their approaches, and using the information they have gleaned from previous issues to solve new ones. Algorithms and training data allow machines to learn as well, but they still need explicit programming or training to change their behavior.
* **Processing Speed and Accuracy:** Machines are excellent at processing swiftly, which enables them to carry out complicated calculations and quickly analyses huge amounts of data. Though they may not be as quick as machines, humans may still make good decisions by using their intuition and capacity for pattern identification.
* Human problem-solving skills are influenced by emotional states, individual experiences, and social relationships. Machines only focus on logical thinking and impersonal decision-making since they lack emotions.

## How to Solve Human Problems in the Maze:

* A person would carefully examine the maze, noting its layout, any walls or other impediments, as well as where the prize was.
* Mental Mapping: A person would mentally draw a map of the maze, imagining different paths and spotting obstacles in the way.
* The human would create a route based on the mental map, taking alternate routes and any barriers into consideration.
* Execution: A person would physically make their way around the maze, choosing their path based on visual cues and their memories.
* Adaptation: If the initially selected path encounters a barrier or dead end, the human would change their plan, go back, and attempt several routes until they located the reward.

## Pathfinding Problem Solution by the Intelligent Agent:

* **Initialization:** The intelligent agent configures the neural network model, initializes the Q-table, and initializes the labyrinth environment.
* **Exploration and exploitation:** The agent explore the maze by performing arbitrary activities in order to collect information and discover the rewards connected to various states and behaviors. It gradually moves from exploration to exploitation while making decisions based on the acquired Q-values.
* **Learning and updating:** The agent modifies the neural network weights to roughly represent the Q-values and modifies the Q-table based on the rewards attained via actions.
* **Making a decision:** The agent chooses the appropriate course of action to pursue in order to maximize the anticipated rewards using the current state and the Q-values.
* **Evaluation and Execution:** The agent carries out the selected action, is rewarded by the environment, and assesses the altered state. The method is repeated until the prize is found or a termination condition is reached.

## Comparing and contrasting human and machine approaches

**Similarities:**

* Both people and robots want to solve the issue and accomplish the intended result.
* Whether through training data or firsthand observations, both rely on learning through experience.
* Both strategies include making choices based on the facts at hand and assessing the results.

**Differences:**

* Humans' decision-making is influenced by their intuition, emotions, and social circumstances. Machines only use computational logic and algorithms.
* Machines are faster and more accurate than humans at processing massive amounts of data quickly and performing complex calculations.
* While computers need explicit programming or instruction to adjust their behavior, humans can change their techniques based on existing knowledge and experience.
* Machines interpret information using mathematical models and algorithms, whereas humans can develop mental representations and use cognitive methods.

# Purpose of the Intelligent Agent:

* Finding the best route from a starting point to a goal requires navigating through a maze or other environment, which is the job of the intelligent agent used in pathfinding. The agent gains knowledge from its interactions with the environment, modifies its knowledge or policy in response to incentives or punishments, and eventually strengthens its ability to make decisions to achieve the objective effectively.
* Exploitation differs from exploration in that it relates to the agent's approach of selecting activities that, in light of its present knowledge, are considered to produce the largest predicted benefits. Utilizing newly acquired knowledge entails choosing actions that, in light of the agent's present understanding, are deemed to be the most optimal.

On the other hand, exploration entails doing acts that might not have high immediate returns but offer the chance to discover new knowledge or about unknown areas of the environment. In order to increase its knowledge and ultimately make better decisions, exploration seeks to find potentially better activities that were not immediately obvious.

**Ideal Proportion of Exploitation and Exploration:** Depending on the learning stage and the degree of environmental ambiguity, the ideal proportion of exploitation and exploration in the pathfinding problem varies. To discover the rewards connected to various states and activities, the agent must first thoroughly study its surroundings.

The agent can move towards exploitation as it learns more and becomes more aware of its surroundings, concentrating on applying what it has learned to make decisions that maximize rewards. To guarantee that the agent keeps exploring uncharted territory and stays out of local optima, it is crucial to keep up some level of exploration even during the exploitation phase.

Depending on the particular issue and the degree of uncertainty present, the precise split between exploitation and exploration may change. Finding the ideal balance that enables the agent to sufficiently investigate the environment while making use of the knowledge acquired to make effective judgements is frequently accomplished through testing and fine-tuning.

Using a reward signal offered by the environment, reinforcement learning can assist the agent in figuring out the best route to the target (the treasure). Based on its activities and the consequent states, the agent is rewarded or punished. The agent learns to link particular behaviors with bigger rewards through repeated interactions with the environment, and it adapts its decision-making process accordingly.

The agent can learn an ideal policy or value function that directs its activities in the direction of the objective through reinforcement learning. In order to favor acts that result in more rewards, the agent investigates various paths, receives feedback in the form of prizes, and updates its knowledge.

# Use of algorithms to solve complex problems

By offering organized and effective methods for addressing complicated issues, algorithms play a significant part in finding solutions. When used properly, algorithms can provide step-by-step instructions for solving difficult issues and can divide large problems into smaller, more manageable subproblems. To assess how well algorithms are used to solve complicated issues, consider the following main points:

* **Effectiveness:** The goal of algorithms is to offer effective answers with the least amount of time and resources needed. An algorithm's efficiency is frequently expressed in terms of time complexity and space complexity, which determine how well the algorithm performs as input size increases.
* **Accuracy:** Algorithms should produce accurate results that satisfy the problem's requirements. To deliver accurate and dependable results, they must take into account all pertinent variables and limitations.
* c. Scalability: Algorithms must be scalable in order to handle more complex issue cases without suffering appreciable performance deterioration. They ought to be built to manage expanding input volumes and adjust to shifting problem complexities.
* d. Optimality: In some circumstances, algorithms look for the best answers by maximising or minimising particular goals. The best feasible outcome for a given criterion is guaranteed by optimal algorithms, assuring that no other outcome can produce superior outcomes.
* a. Robustness: Algorithms should be tolerant of changes in input data and capable of gracefully navigating unforeseen circumstances. They ought to be built to deal with edge cases and exceptions without producing mistakes or unexpected results.

# Deep Q-Learning implementation employing neural networks:

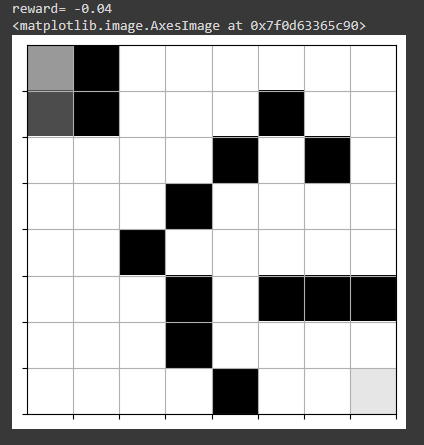
In order to approximate the Q-function, which calculates the predicted rewards for each potential action in a given state, deep Q-learning is used in this game. An overview of the use of deep Q-learning is given below:

* **State Representation:** The game environment is represented by states, which typically entail encoding the current maze layout, the location of the pirate, and any other pertinent data.
* **Neural Network Architecture:** To approach the Q-function, neural network models like the deep Q-network (DQN) are constructed. The state representation is the network's input, and its output is a group of Q-values corresponding to each conceivable course of action.
* **Experience Replay:** An experience replay buffer is used to increase learning's stability and effectiveness. By interacting with the environment, the agent gathers experiences and stores them in a memory buffer. A random sample of experiences is taken during training to update the neural network's weights.
* **Trade-Off Between Exploration and Extraction** An epsilon-greedy approach is frequently employed to strike a balance between exploration and exploitation. In order to explore the environment and gain new insights, the agent periodically chooses random actions in addition to actions based on the Q-values predicted by the neural network.
* **Q-Learning Updates:** The Q-learning algorithm is used to train the neural network. The Bellman equation is used to calculate the loss function, which contrasts the goal Q-values obtained from the reward and the next state with the expected Q-values. To reduce the loss, the weights of the network are updated via gradient descent.
* **Iterative Training:** Up until the agent's performance converges or reaches the target level, the process of gathering experiences, updating the neural network, and fine-tuning the Q-function is performed iteratively.

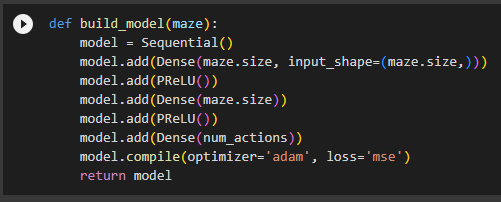
The agent can discover the best strategy for navigating the maze and getting to the destination (the treasure) by utilizing neural networks to approximate the Q-function. With the help of reinforcement learning, the neural network captures the underlying dependencies and patterns in the state-action space, empowering the agent to take well-informed decisions and continuously improve.

# Project file

First in this project I explore the all code files and related commented then, I will try to do some experiments on the code what each function does and what it returns. First I draw the maze matrix using the draw method.

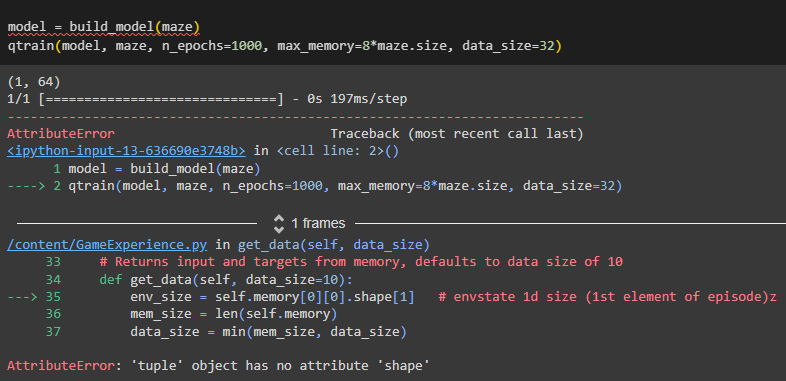


Here is the matrix, then I the basic neural network for training the model it uses the most commonly used activation function which is the RELU, I add one or more layer to the network to experiments things out.



Here I also add the adam optimizer which is most commonly used optimizer and used the loss function which mean squared error ‘mse’.

Here I add the code for TODO project but I try many times to cover up this but still I am getting this error even though I tried different techniques to print out the each step which is the best way to debug your code but I didn’t make I am adding the error which I got.



# References:

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533. <https://www.nature.com/articles/nature14236>

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT Press. <https://mitpress.mit.edu/9780262039246/reinforcement-learning/>

<https://www.scaler.com/topics/differences-between-artificial-intelligence-and-human-intelligence/>

<https://www.frontiersin.org/articles/10.3389/frai.2021.622364/full>

<https://slejournal.springeropen.com/articles/10.1186/s40561-019-0088-z>

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484-489. <https://www.nature.com/articles/nature16961>

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., ... & Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. In International conference on machine learning (pp. 1928-1937). <https://proceedings.mlr.press/v48/mniha16.html>

<https://www.scaler.com/topics/differences-between-artificial-intelligence-and-human-intelligence/>

<https://dataconomy.com/2022/04/20/is-artificial-intelligence-better-than-human-intelligence/>