A HYBRID INTELLIGENCE SYSTEM WITH THE OPTIMAL LEVEL OF HUMAN PARTICIPATION FOR PERFORMANCE ENHANCEMENT IN DECISION MAKING

SO KEAN FONG

UNIVERSITI TUNKU ABDUL RAHMAN

A HYBRID INTELLIGENCE SYSTEM WITH THE OPTIMAL LEVEL OF HUMAN PARTICIPATION FOR PERFORMANCE ENHANCEMENT IN DECISION MAKING

SO KEAN FONG

A project report submitted in partial fulfilment of the requirements for the award of Bachelor of Science (Honours) Software Engineering

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

September 2022

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature :

Name : So Kean Fong

ID No. : 1806887

Date : 8 September 2022

APPROVAL FOR SUBMISSION

I certify that this project report entitled **"A HYBRID INTELLIGENCE SYSTEM WITH THE OPTIMAL LEVEL OF HUMAN PARTICIPATION FOR PERFORMANCE ENHANCEMENT IN DECISION MAKING"** was prepared by **SO KEAN FONG** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Science (Honours) Software Engineering at Universiti Tunku Abdul Rahman.

Approved by,

Signature : KLYAU

Supervisor : Prof. Dr Yau Kok Lim

Date : 8 September 2022

Signature : ______________________

Co-Supervisor : ______________________

Date : ______________________

ABSTRACT

Hybrid Intelligence System is a system allows human and machine work in together to make decision. The collaboration between human and machine can improve the ability of decision making. Human can make a different or complex decision than machine, human’s knowledge can determine whether the decision they make is suitable or not while machine have to predict the correct answer with a lot of training session with dataset. The purpose of this project is to construct a hybrid intelligence system to analyse whether the human knowledge can influence machine decision making then identify the order of human participate in decision making of machine under different circumstances and compare the performance of Human In The Loop between Reinforcement Learning and Hybrid Intelligence System. The approach of machine learning will be used is Reinforcement Learning and the solution use to solve the problem is Markov Decision Process. Therefore, the approach and solution will involved human input to change the result of the enhanced code.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

RL	Reinforcement Learning
HITL	Human In The Loop
CTIL	Computer In The Loop
ML	Machine Learning
MDP	Markov Decision Process

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

This project develops an artificial intelligence algorithm that adjusts and provides the right level of human participation in hybrid intelligence, which integrates both artificial intelligence and human intelligence. Specifically, we integrate the optimal level of human participation or inputs in an artificial intelligence approach called reinforcement learning (RL) (Najar and Chetouani, 2021), contributing to the optimal collaboration between humans and machines under a diverse range of circumstances, such as different levels of dynamicity in the operating environment.

1.1 Background

RL, which is a traditional machine learning method, enables a decision maker (i.e., agent) to learn the right course of actions for increasing its rewards (or reducing its penalties) in order to achieve its goals under a dynamic and complex environment. Through computer simulation, an agent can learn through performing actions in a trial-and-error manner. Appropriate behaviour is rewarded, while inappropriate behaviour is penalized. As a result, as time passes, an agent learns the best potential actions in various circumstances.

1.2 Significance of the Project

This is a significant project contributing to the national agendas. From the Malaysian Shared Prosperity Vision (SPV) 2030 (MEA, 2019), one of the key priority areas is to provide a high quality of life for the Malaysian citizens by 2030. The government has published a new framework called 10-10 Malaysian Science, Technology, Innovation and Economic (MySTIE), which is designed to transform the old economy of the country into a new knowledge-intensive economy. The goal of this framework is to improve Malaysia's position in the technology logistics system worldwide by generating shared economic success throughout different ecosystems in the country. Ultimately, it enhances the competitiveness and survivability of all industries in the country. The MySTIE framework includes *ten* science and technology (S&T) and *ten* socio-economic drivers, respectively. In the S&T domain, it raises the quality of life of

Malaysian citizens through the discovery of new technology for creating new impact in different sectors. In the socio-economic sector, it and identifies factors helping to raise the quality of life of Malaysian citizens. This project adjusts and provides the right level of human participation in hybrid intelligence, and so it contributes to two different S&T drivers, namely advanced intelligent systems (i.e., artificial intelligence) and neurotechnology (e.g., decision making and human-computer interface).

1.3 Background of the Problem

RL is a popular artificial intelligence approach used widely in the IT and business sectors for improving various applications, such as the marketing recommendation system that recommends the right products to customers based on their behaviours. During the Covid-19 pandemic, most of the customers purchase their daily stuff, food, and entertainment products (e.g., online games) through online platforms. Since customer behaviours and preferences change dynamically, RL has been proposed to identify and recommend suitable products to customers based on a large dataset comprised of customer reviews and ratings of products.

There are three main problems when using RL in hybrid intelligence. Firstly, there is lack of context awareness in decision making. While machines make decisions based on data, humans make decisions based on their knowledge and feelings in a diverse range of circumstances. Secondly, a longer time is needed for machines to process a large dataset. Thirdly, inaccurate samples can cause incorrect decisions.

1.4 Problem Statement

RL is expected to improve the hybrid intelligence system based on a large dataset. In RL, machines create value function (or knowledge) through trials and errors. In other words, the machine updates value function through its experience in interacting with the operating environment. The machine enhances its value function to provide updated knowledge as time goes by. This process repeats until the machine achieves the best possible value functions and convergence speed.

1.3.1 Context aware of decision making

The decision making of the machine expects to be more accurate and safer. When under different circumstances such as driving on the road, the car in front might stop which we can see the car and starting to slow down. For AI, the way they perform are similar to human, but they need more time and data to improve. However, AI systems will not be able to prioritize various inputs based on their importance, a prerequisite for making decisions and taking the right action. (Thiopoulos C, 2020). Markov Decision Process (MDP) is used in this project to figure out the low risk and maximize the probability of reward.

1.3.2 Large dataset and sample correctness

When under the circumstances such as marketing recommendation system, since there is a large size of dataset, the machine expects to recommend suitable product to the customer with review and rating of the product. However, the interest of the people will keep changing in a short period, that will create a large dataset which the sample does not fully correct. MDP is used in this project to reduce the unused and less efficiency dataset.

1.5 Objectives

The objectives of this project are:

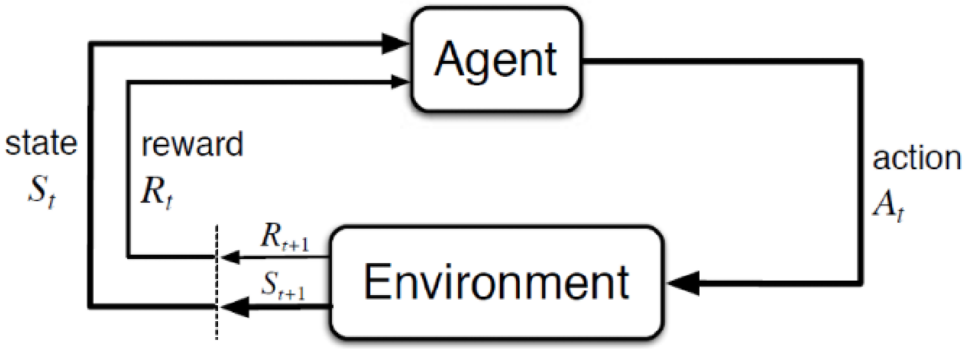
I. To construct a hybrid intelligence system to analyse whether the human knowledge can influence machine decision making.

II. To create a hybrid intelligence system that can identify the order of human participation in decision making of machine under different circumstances.

III. To compare the performance of Human In The Loop in Decision Making between traditional RL approach and hybrid intelligence system.

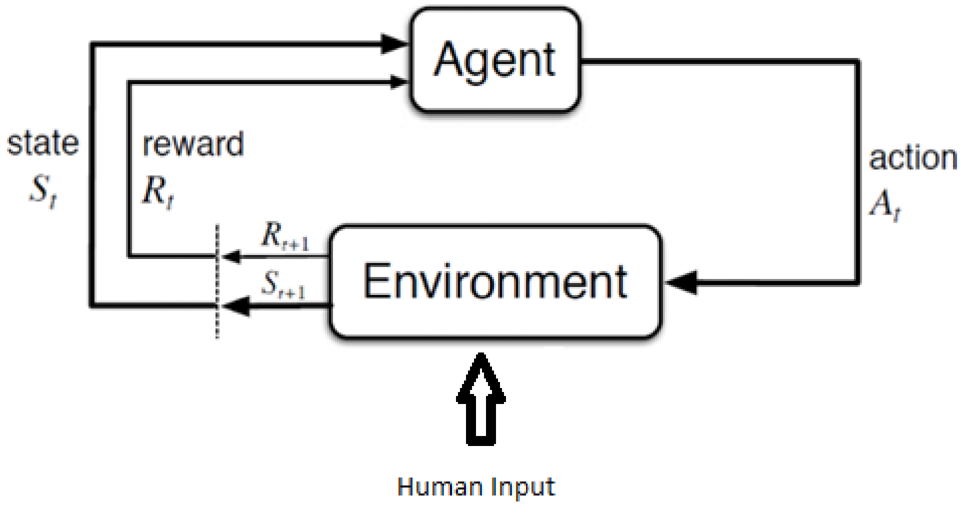
1.6 Proposed Solution

Figure 1: Markov Decision Process (MDP)



In this project, Markov Decision Process will be the the solution for the problem. The basic elements of RL are environment, state, reward, policy and value. MDP is using a framework based on mathematic to explain the environment for RL (Sreenath14, 2020). The agent interact with the environment each discrete time step, for example t = 0, 1, 2, 3,... The environment state S_t will provide information to the agent when it reach the S_t again. The agent carries out A_t according to the current status of the environment on time t. In the next moment, the agent also receives a reward signal with R_{t+1}. As a result, a series of numbers like as State_0, Action_0, Reward_1, S_1, A_1, R_2... emerge. Variables selected at random R_t and S_t have well-defined discrete probability distributions. By virtue of the Markov Property, the PD are simply reliant on the previous state and action. Let S stand for states, A stand for actions, and R stand rewards, respectively. The probability values of S_t, A_t, and R_t with the preceding state s adopting the values s', r, and an is then given by, $p(s', r|s, a) = P\ \{S_t = s', R_t = r|S_{t-1} = s, A_{t-1} = a\}$. P will be the function to control the dynamics of the process. [5]

Figure 2: Markov Decision Process (MDP) with human input



From Markov Decision Process, we implements human input into the solution which user inputs value into the system to change the environment value and run the simulation. Therefore the environment such as reward will change the value of it and the agent will make a new action depends on the new reward value.

1.7 Project Approach

In this project, it will divide into three different phases. The first phase is Data preprocessing, we will have an analysis of the dataset will be used for performance enhancement. The dataset finds on Kaggle and Google. The data stimulates on the tool that will use is Jupyter Notebook. Second phase is training the model. After we select a model from OpenAi Gym, we use our data to improve the decision making of the agent. In the training phases, we add the human knowledge to suggest the agent. Finally, when the training is complete, we test the agent with the dataset again then create another dataset to compare the difference of human involvement agent's performance and machine performance. The detail of the approach will be discussed in methodology part.

1.8 Scope of the Project

This project implements a hybrid intelligence system for making better decision makings. The system allows user to participate in the learning process of the Artificial Intelligence. The user can input the value to change the result and let

AI agent to learn from the new value. When the agent completes the learning process, it pauses a few second for user to determine want to continue the learning process or not. When user participate in the process, the agent will determine the participate level of the user's knowledge level. If the knowledge level of the user does not reach the right level, the agent lowers down the participate level of the users.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this chapter, we discuss about the studies about 'A Hybrid Intelligence System with the Optimal Level of Human Participartion for Performance Enhancement in Decision Makings'. There are three literature review study in Sufficient Dataset for hybrid intelligence system in decision makings, Context Awareness for AI in Decision Making and Traditional AI approach and hybrid intelligence system.

2.2 Sufficient Dataset for hybrid intelligence system in decision makings

The By making decision, we need many data to confirm the decision we have make is precise. AI will also make dicision depends on algorithm set by the programmer. For hybrid AI system, human will participate in the decision making with AI. Human can change the algorithm of the AI to make the decision they want. Paullada A, Raji ID, Bender EM, Denton E andHanna A (2021) have mentioned in their research paper that dataset always seen as a negative impact to advancement in algorithmic and scientific development, with just a few benchmark datasets available. Some dataset will be cause to make wrong decision, human have to examine the dataset is suitable to use or not. The correctness of dataset can influence the performance of hybrid intelligence in decision making.

In another research, Mahmud H, AKMN Islam, SI Ahmed and K Smolander (2021) have mentioned that the algorithm will make sure the people satisfied with the choice when the algorithm provides a better choice than human decision. Most of the time, the AI algorithm used will make sure the people satisfied with the choice when the AI algorithm provides a better choice than human decision choice. The people would like to reject the AI algorithm and choose to use human decision making algorithm. When people choose to use algorithm to decide, they will not feel satisfied because the decision might not their desired choice.

In a conclusion, dataset is important to traditional AI and hybrid intelligence system. Without data, we cannot make a precise decision due to the unknown result of each choice. Human has to provide a accurate dataset and less mistake rate algorithm for machine which can reduce the error rate to make a decision.

2.3 Context Awareness for AI in Decision Making

The AI have different context awareness will have different decision making. Mostly context awareness for AI will be used in business which the business can collect a lot of data about user preference for buying product online to train the AI to improve their decision makings to recommend the suitable product for the customer. Kulkarni S, S.F Rodd (2020) have mentioned in their report that hybrid recommendation system using different method can earn different benefits which still need to continue investigate but it also limited by it's contend based and CF technique. The recommendation system mostly fails due to the speed of the user preference changes over the time. For Mozgovoy M (2018), the aim of his research is to demonstrate how good a hybrid of machine learning and human labour may help build a robust gaming Artificial Intelligence system that can act like a person. The limitation of the AI is it need more and more experience and data from the game to improve their react. The strength is the correct combination of the analysis, machine learning approach, learning data and decision point will produce a better AI. The machine learning and context awareness method combine together and used in the environment which data input and output are clearly marked and also a suitable decision making.

Another research done by MDCR Hernandez (2021) is to provide a survey of intelligent mobile Context-Awareness Recommendation System which focus on mobile computing and AI approach. In this research, the author proposes a framework to analyze the context awareness recommendation systems. It also uses some related literature review and AI approach to complete the research. The AI approach such as traditional supervised learning approach and deep learning approach. The limitation of context awareness in decision makings are due to the user will keep moving around when using mobile devices, the context awareness recommendation system will also need more data to figure out the situational of user then only can provide a better and satisfied choice for users. The context awareness recommendation system is increase for different application scenario and also a wide range of items which contain daily life stuff, books, food, games, and other items. Nowadays, using AI in context awareness recommendation system will be much easier to predict the user preferences with the large amount of data collected on the internet.

In a nutshell, the context awareness is important for machine to decide the correct action to give human. If there is one wrong decision has made, it will cause a huge problem. So that human also needs to be aware of select the correct dataset to build up the context awareness of the machine. The best option is still create a hybrid intelligence system with the combination of human and machine. It can provide better result with human provide the knowledge and experience and machine take the dataset relevant to specific context.

2.4 Traditional AI approach and hybrid intelligence system

Traditional AI needs a large amount of dataset to have training to get a precise decision making algorithm. For hybrid intelligence system, the machine learns the knowledge from human, in the training process, human also can learn something new from the machine. The limitation of hybrid intelligence system is it needs to have communicate with human then only can make them more efficiency. In Wiethof, C and Bittner (EAC 2021)'s report, they aims to support Hybrid Intelligence by providing a knowledge of collaborative learning and a shared ground methodologies between machines and humans. They find out that human can interact with the machine learning algorithm by optimizing the algorithm with knowledge of human and machine to have better performance. But the limitations are if they need to implement the Human In The Loop (HITL) Hybrid Intelligence System, there will need the context-aware and scenario like human feedback, knowledge and experience to improve the algorithm of the machine learning. The performance of HITL hybrid intelligence system are improve by the instruction given.

In another research done by Wells L and Bednarz T (2021), it aims to investigate the present state of Explained Artificial Intelligence (XAI) in the field of Reinforcement Learning (RL). They found that the Reinforcement

Learning agent will keep learn the behavior which will keep them unique and stronger. From the research, there are few limitation in the field of Explained Artificial Intelligence in RL which is less of new algorithm, less tested by user, too complex to explain, basic visualizations and less of opensource code. The Reinforcement Learning agent has two different agent which uses dataset and not use any training data. The RL agent in hybrid intelligence system also using human replay data and feedback to improve the performance. According to Dellermann D, Calma A, Lipusch N and Weber T (2019), they mentioned that the aims in the research is to show how hybrid intelligence system work and developed, and also conceptualize hybrid intelligence systems and present a first design knowledge taxonomy for construction. The hybrid intelligence system aims to combine human knowledge and AI together which can perform well in complex decision makings and it can reduce the error in making decision.

In conclusion, the hybrid intelligence needs human participate with AI to provide a precise decision. The communication between human and AI are very important because AI needs the feedback and instruction of human to perform the next important action to avoid error.

2.5 Summary

In conclusion, hybrid intelligence system which combine with the knowledge of human and machine will be making more precise decision. AI is providing human a good help on decision making. The limitation for them is needed a large amount of dataset to train themselves. When the AI combines with human are call as hybrid intelligence system, the machine can learn the knowledge from human while the human also learn from the machine too. With sufficient data and precise dataset, the machine can no longer giving a wrong decision which they learn in the data. If the dataset have error, it will cause faulty decision. In some specific context, the machine should have the high accurate decision to avoid risk but it needs the dataset relevant to the specific context. With human participation, machine can combine the dataset with the knowledge of human to make sure more accurate and safety decision. Traditional AI still provide a good help to human in many field, but dataset need to be re-new to make sure the current dataset can cope with the context. Hybrid intelligence system would provided more useful and precise decision because the machine learn the knowledge of human and the dataset have been collected to train themselves. The traditional AI and hybrid intelligence still have their own strengths and limitations which will affect human to choose which one to implement in the system.

CHAPTER 3

METHODOLOGY AND WORK PLAN

3.1 Introduction

The methodology and work plan of the project will be discussed in this chapter. The methodology that use in this project is Quantitative Research. The phases of this methodology are conceptual stage, the design and planning stage, empirical stage and analytic phase. In other hand, Gantt chart will also be present in this section.

3.2 Quantitative Research

The quantitative research phases will be discussed in this section. The phases are conceptual stage, the design and planning stage, empirical stage and analytic phase.

3.2.1 Conceptual phase

Firstly, identify the concept of hybrid intelleigence system. During this phases, we also discussed about the research purpose and understand the requirement of from the question. At the same time, point out the problem faced with the research title and determine the purpose of this research and identify the project scope. The similar literature about the topic will be review. The literature review part will be discussed in Chapter 2.

3.2.2 Design and Planning phase

In this phases, we are going to choose one of the AI approach to complete this research. The approach will be used is reinforcement learning approach with human in the loop which means human will participates in the training phases of the Artificial Intelligence. For example, The data will be used is get from Google and Kaggle. Before using the data from the internet, figure out which one will be suitable for used to train the AI agent. Then understand whats the result look like. Therefore, the data will be used in this project is Q-Learning Theory with an Example for Beginners. It used the reinforcement learning agent in a environment which is a maze to run the simulation.

3.2.3 Emprical phase

In this phases, analyse all useable data collect from Kaggle and Google and prepared for next steps. We use the Q-Learning Theory to train the agent and check the first test run result of the agent. With Q-Learning Theory, the initialize Q value to zero then the agent will choose the action according to the value of Q and perform the action whether it want to move left, right, up or down. After the action performs, the reward will be calculate and update the Q-value in the Q-table for the next training iteration. We change the environment of baseline code with enlarge the environment and use the algorithm which is human inputs to train the agent again.

3.2.4 Analytic phase

In analytic phases, we will continue using the same Q-Learning Theory, enlarged environment and new algorithm to perform training of the agent in the OpenAI - GYM. The algorithm will be performed is human in the loop which also known as human input. After that, we run the simulation with the algorithm and collect the results.

3.3 Project Plan

3.3.1 Work Breakdown Structure (WBS)

Table 1: Work Breakdown Structure

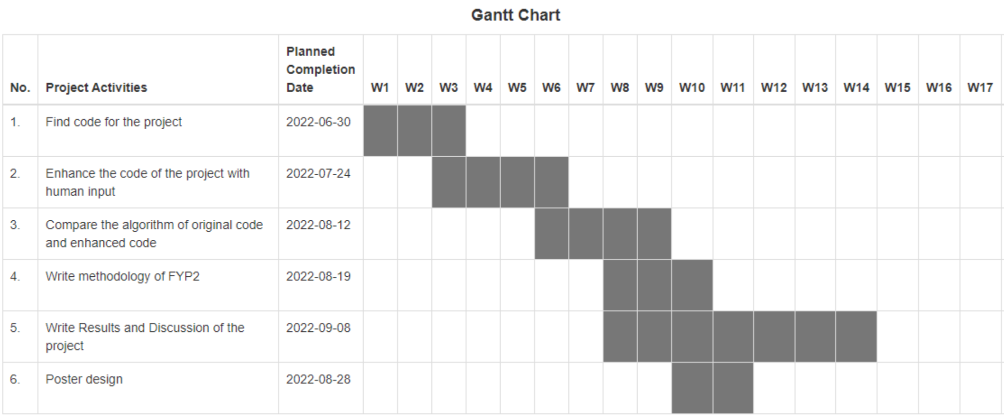
	Name	Description
1.1	Project research	Find suitable topic for project research
1.1.1	Planing and research about the project title	Understand the project title
1.1.2	Create Gantt chart	To ensure every progress finishs on time
1.1.3	Gathering the information of the project	Collects all usedable information from Science Direct, IEEE Explore, doi.org and other more

1.1.4	Identify the problem statement and background of the problem	Find out the problem that will be face
1.1.5	Identify the project objective and project scope	Create objective to reach aim
1.1.6	Proposed solution of the problem and determine the project approach	Create a solution to solve the problem of the project
1.2	Literature Review	Find similar topic of the project and theory of the algorithm
1.2.1	Sufficient Dataset for hybrid intelligence system in decision makings	If the dataset is a lot, the agent will have more data to predict decision. But its cost more time to have training.
1.2.2	Context Awareness for AI in Decision Making	In different situation, AI will make a different choice than human as they will have their on ethics and principles
1.2.3	Traditional AI approach and hybrid intelligence system	Hybrid Intelligence that involve human knowledge will provide a different answer compare to traditional AI in the same question
2.0	Preparation	Prepare the tool and algorithm
2.0.1	Determine research development tools	Find out the tools will be used such as Kaggle, Jupyter Notebook and Open AI-GYM
2.0.2	Determine algorithm	Find out the algorithm can be used to solve the problem of the project
2.1	Analyse the result of the training	To analyse the result from 30 iteration of training
2.1.1	Create environment for AI agent to train	Create environment which is 4 column x 4 row of grid world

2.1.2	Record the result of the agent to reach goal	Collect the result of agent with the algorithm of Markov Decision Process
2.1.3	Enlarge the environment	Enlarge the environment to 5 column x 5 row of the grid world
2.1.4	Create algorithm of the training	The algorithm will be used contain the human input
2.1.5	Record the final result of the agent to reach goal	Collect result of the agent that have used algorithm with human input
3.0	Results and Discussion	Make discussion of the result
3.1	Run simulation and collect result	Run 30 times of simulation to get the result from MDP algorithm and MDP with human input
3.2	Compare the result of baseline code and enhanced code	Compare the agent that used MDP algorithm and agent that used MDP with human input

3.3.2 Gantt Chart

Figure 3: Gantt Chart for FYP2



3.4 Development Tools

Table 2: Development Tools

Development Tools	Description
Kaggle	Kaggle is a website that Data Science engineer and Machine Learning engineer can share their idea, the example of the code, and new coding information. In the website, we find out many dataset that we can use to complete the project
Jupyter Notebook	Jupyter Notebook, an open-source web application which people are allowed to create, edit and share the code of their project. The programming language is using Python.
OpenAI -GYM	OpenAI – GYM provides the environment for AI to train themselves. There are many environment and different agent can be used.

3.5 Summary

In conclusion, this chapter discuss about the methodology, project work plan and development tools will be used in this project. The project finished by used the steps providing in the methodology. The project work plan is to help us to complete the project on time.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

The results collected from the simulation are discussed in this chapter. First, the results and graphs of the baseline code and enhanced code will be showed. Therefore, we will compare the results of baseline code and enhanced code. We also will discuss about the algorithm and flowchart for both code.

4.2 Baseline Code

The baseline code that are used in this research is a Q-Learning Theory with an example for Beginners. In this baseline code, we will use GridWorld as the environment and Q-Learning Theory to train our agent.

4.3 Environment and Enlargement Environment

The environment is a maze which has 4 column x 4 row of grid world. The agent have to finish the maze by reachs the finish point of the maze or hit the block of the maze which is presented with red colour. Therefore, we enlarges the size of the environment to 5 column x 5 row of grid world. The reason that we decide to enlarge the maze because baseline code environment is too small for us to train our agent with human input. The block of the environment is (1,0), (1,2), (1,4), (2,4), (3,0), (3,2), (4,0), (4,2), (4,3), (4,4).

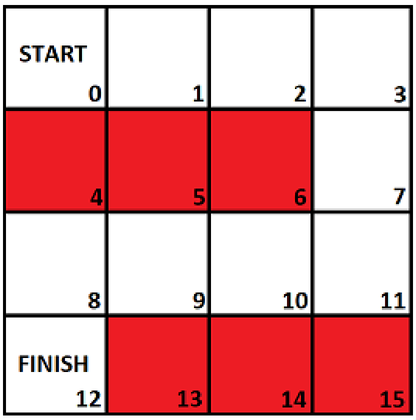


Figure 4: 4 column x 4 row of grid world

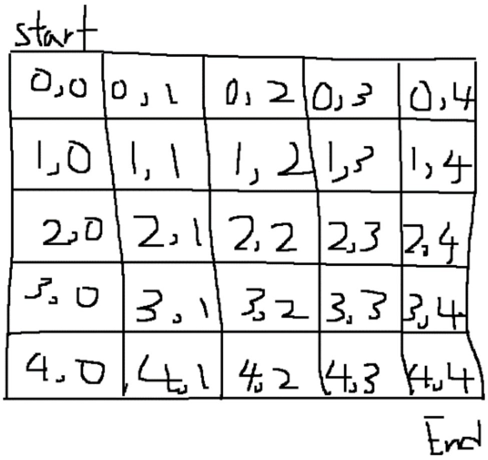


Figure 5: 5 column x 5 row of grid world

4.3.1 Simulation Parameter

In this section, the figure below shows the parameter of the simulation such as environment info and agent info which contains number of states, number of action, epsilon, learning rate, discount and random seed. In figure 6, it is the default value of number of state which is 5 states and number of action which is 2 actions for baseline code and enhance code. For number of state as there is five state in the environment and number of action has two actions in the state which the agent can move up, down, left and right. In figure 7(a) and 7(b), it is the environment info, the info discusses at section 4.3 and agent info for baseline code shows number of actions is 4, number of states is 16, epsilon is 0.01, learning rate is 0.1, discount is 1 and random seed is 0. In figure 8(a) and 8(b), the environment info is enlarged to 25 state space and discusses at section 4.3 and the agent info has number of actions is 4, the number of states change to 25, epsilon is 0.01 then increase the learning rate of the agent from 0.1 to 1.0, discount is 1 and random seed is 0.

```
num_States = 5 # we have 5 states in our state space
num_Actions = 2 # we have 2 actions in our action space
```

Figure 6 : Number of states and number of actions for two algorithm

```
envInfo = [[(0,0), (0,1), (0,2), (0,3), (1,3), (2,3), (2,2), (2,1), (2,0), (3,0)], [(1,0), (1,1), (1,2), (3,1), (3,2), (3,3)]]
# envInfo[0] => safe path list
# envInfo[1] => cliffs list
```

Figure 7(a): Environment info for baseline code

```
agentInfo = [4, 16, 0.01, 0.1, 1, 0] ###numActions, numStates, epsilon, learningRate, discount, random seed
```

Figure 7(b): Agent info for baseline code

```
envInfo = [[(0,0), (0,1), (0,2), (0,3), (1,1), (1,3), (2,0), (2,1), (2,2), (2,3), (3,1), (3,3), (4,1)],
           [(1,0), (1,2), (1,4), (2,4), (3,0), (3,2), (4,0), (4,2), (4,3), (4,4)]]
# envInfo[0] => safe path list
# envInfo[1] => cliffs list
```

Figure 8(a): Environment info for algorithm with human input

```
agentInfo = [4, 25, 0.01, 1, 1, 0] ###numActions, numStates, epsilon, learningRate, discount, random seed
```

Figure 8(b): Agent info for algorithm with human input

4.4 Enhanced Code

The enhanced code is enhanced by using baseline code with the implementation of human in the loop to achieve hybrid intelligence. In the code, we used user input as human participate in the simulation as human give feedback to the system. The code will stop each time the training result show and it will ask the user to continue training or change the reward value.

```
for i in range(30):
    print("Iteration Number: " + str(i))
    currentReturn = 0
    reward = 0
    stt = env.envStart()
    act = agent.agentStart(stt) # choose action
    rew, stt, trm = env.envStep(act)
    currentReturn += rew
    while(True):
        act = agent.agentStep(rew, stt) # update q table and choose action
        rew, stt, trm = env.envStep(act)
        currentReturn += rew
        if trm: # agent is terminated
            agent.agentEnd(rew) # make a final update to q table
            print("Reward: " + str(rew))
            print("Agent is terminated!")
            ####ask user input for accept the result or not
            answer = input("Do you accept the result? (Yes or No)\n")
            #if yes, the iteration will continue;iF no,the user can input the value to change
            if any(answer.lower() == f for f in ["no", "No", "NO", "n","N"]):
                #Get new reward
                print("Please enter new value of reward: " )#+ str(currentReturn))
                reward = int(input("New value: "))
                #print the new reward value
                currentReturn = rew + reward
                print("New Reward: " + str(currentReturn))
                returns.append(currentReturn) # save the return of the episode
                break
            else:
                print("Reward: " + str(rew))
                print("Agent is terminated!")
            returns.append(currentReturn) # save the return of the episode
            break
```

Figure 9: Enhance code with human input

```
Iteration Number: 0
Reward: 100
Agent is terminated!

Do you accept the result? (Yes or No)
```

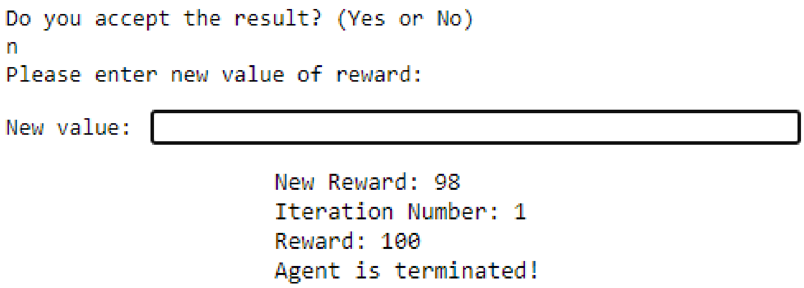
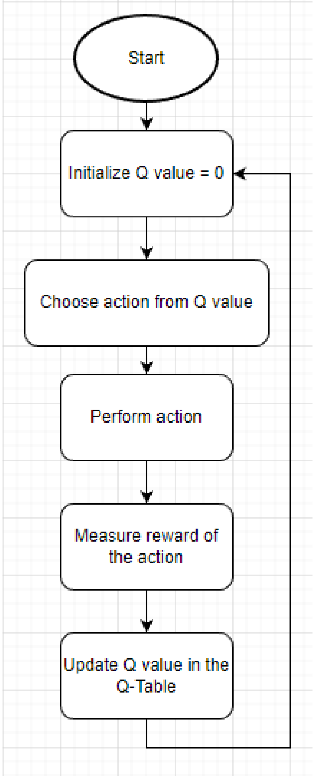


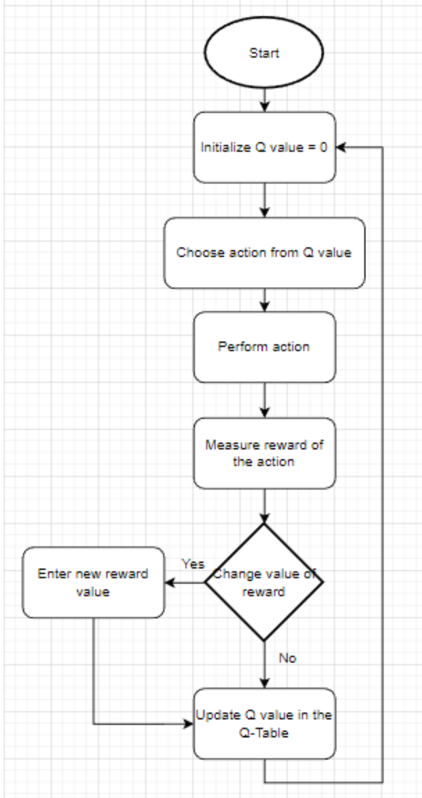
Figure 10: Human input for enhanced code

4.5 Flowchart

Below is the flowchart of Baseline code() and Enhanced code(). The difference between two flowchart is human can participate in the training state of the agent by decide to continue training or input the value to change the old value and update the Q-table.

Table 3: Comparison of two algorithm

<table>
<tr>
<td>


</td>
<td>


</td>
</tr>
<tr>
<td>Figure 11(a): Without Human Input</td>
<td>Figure 11(b): With Human Input</td>
</tr>
</table>

In the flowchart figure 8(a), the process of the Without Human Input will start from initialize Q value to zero then the agent will perform the action depend on the Q-value on the Q-learning table. After that, the reward will be changes from the updated Q-value in the Q-table. Then the next simulation will be start with the updated Q value. In the flowchart figure 8(b), the process will be the same from the start, but it will be difference after measure the reward of the action. The system will ask user decides to continue the training progress without change any reward value or the user have to enter the new reward value for the simulation, then the Q-value will be update in the Q-table.

4.6 Results of Baseline Code

The results of baseline code are shown in below. First and foremost, baseline code does not have the user input process in the training state. Therefore, the Q-Table of the agent and the line graph represents the sum of the reward in an episode.

```
Iteration Number: 0
Agent is terminated!
Iteration Number: 1
Agent is terminated!
Iteration Number: 2
Agent is terminated!
Iteration Number: 3
Agent is terminated!
Iteration Number: 4
Agent is terminated!
Iteration Number: 5
Agent is terminated!
Iteration Number: 6
Agent is terminated!
Iteration Number: 7
Agent is terminated!
Iteration Number: 8
Agent is terminated!
Iteration Number: 9
Agent is terminated!
Iteration Number: 10
Agent is terminated!
Iteration Number: 11
Agent is terminated!
Iteration Number: 12
Agent is terminated!
Iteration Number: 13
Agent is terminated!
Iteration Number: 14
Agent is terminated!
Iteration Number: 15
Agent is terminated!
Agent is terminated!
Iteration Number: 16
Agent is terminated!
Iteration Number: 17
Agent is terminated!
Iteration Number: 18
Agent is terminated!
Iteration Number: 19
Agent is terminated!
Iteration Number: 20
Agent is terminated!
Iteration Number: 21
Agent is terminated!
Iteration Number: 22
Agent is terminated!
Iteration Number: 23
Agent is terminated!
Iteration Number: 24
Agent is terminated!
Iteration Number: 25
Agent is terminated!
Iteration Number: 26
Agent is terminated!
Iteration Number: 27
Agent is terminated!
Iteration Number: 28
Agent is terminated!
Iteration Number: 29
Agent is terminated!
```

Figure 12: Training state of the baseline code

In Figure 12, it shows the iteration of 30 times for the agent to train with Markov Decision Process. When agent crashed to the block or reach to the end, it will be terminated.

```
In [9]: print(agent.qTable)

[[-19.         -10.         -19.01        -4.39972588]
 [-10.019       -3.71482728 -10.0271      -3.10843021]
 [-10.03539     -2.92784479 -10.06468304  -0.32687236]
 [-10.07868346  -2.30651771   5.51448069 -10.0504659 ]
 [  0.           0.           0.           0.        ]
 [  0.           0.           0.           0.        ]
 [  0.           0.           0.           0.        ]
 [ -1.60355784 -10.11260259  16.28859471 -19.52479871]
 [  0.           0.          95.76088417  -0.1       ]
 [-10.16108106  80.66865012 -10.18908632  -0.2388552 ]
 [-10.09978704  57.06177068 -19.59956581  -0.50967267]
 [ -1.04615008  34.66768307 -10.13266874 -10.14692177]
 [  0.           0.           0.           0.        ]
 [  0.           0.           0.           0.        ]
 [  0.           0.           0.           0.        ]
 [  0.           0.           0.           0.        ]]
```

Figure 13: Q-Table of the baseline code

In figure 13, it shows the Q-learning table of the baseline code which the agent will follow the value in qTable to perform the next action.

```
In [10]: plt.plot(returns)
         plt.ylabel('Sum of Rewards in an Episode ')
         plt.xlabel('Iterations')
         plt.show()
```

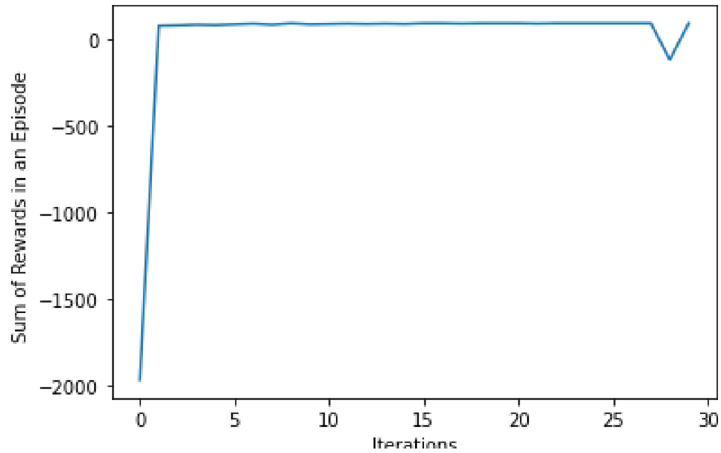


Figure 14: Without Human Input

In figure 14, it shows the sum of reward in for each episode. From iteration 0, the sum is from near -2000 graduly increase to 0 at iteration 3 or 4. Therefore, it remains at 0 until iteration 26 or 27 because the agent learns the path and reach to the end. So the reward will be 0 as the action of the agent make each time will be -1, so after the agent reached the end of the maze, it will get reward of 100.

4.7 Results of Enhanced code

The results of enhanced code are shown in below. Before each new iteration starts, user will be asked to continue training the agent or want to input the new reward value of the training. Therefore, updates the Q-Table of the agent and line graph represents the sum of the reward in an episode.

```
In [18]: plt.plot((returns))
         plt.ylabel('Sum of Rewards in an Episode ')
         plt.xlabel('Iterations')
         plt.show()
```

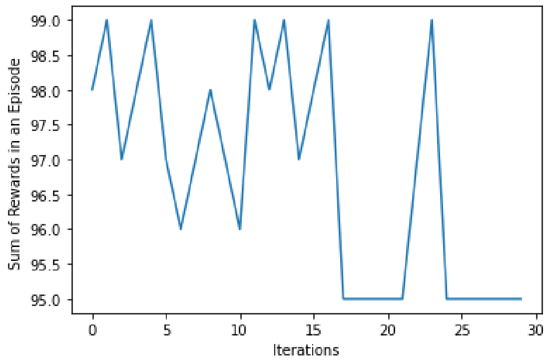


Figure 15: With Human Input

In figure 15, it shows the sum of reward for each episode of the agent uses the algorithm with human input. The graph shows fluctuate from iteration 0 until iteration 16 or 17 is because the user decides to input the value to change the reward value. From the iteration 0, the result of the reward is 98, we can figure that the agent does not perform more than few moves then it crashs to the block and terminated. So the reward will be less. From iteration 16 or 17 until 21 and 24 to 30, the results are remain straight line is because the agent has learned the path to reach the end of the maze.

4.8 Comparison of two results

In this section, we will compare the results of baseline code and enhanced code. From the graph, we can see that graph from baseline code increases sharply from iteration 0 and remains the same until iteration 23 and fluctuate until iteration 30. The reward value from iteration 3 remains at 95 until iteration 23 because the agent has learned how to avoid the block of the environment from the reward. From the graph, we can see that graph from enhanced code is fluctuate from the start until interation 16 or 17 until 21 or 22 because the user inputs the different value from the iteration 0 until iteration 16 or 17, when the reward remains the same then the graph will show a straight line.

Table 4: Comparison of two results

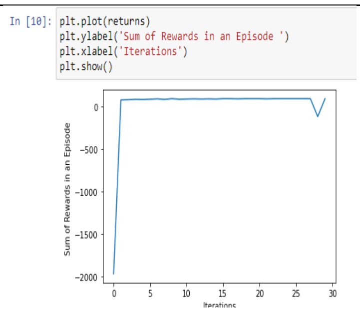
<table>
<tr>
<td>


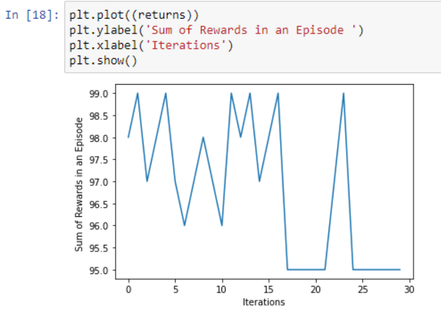
Figure 16(a): Without human input
</td>
<td>


Figure 16(b): With human input
</td>
</tr>
</table>

For figure 16(a), from iteration 0 and remains the same until iteration 23 is a straight line. There does not have user to participate into the training, so the result of the graph will be 0 as the agent performs the same action with the same route to reach the end. For figure 16(b), from iteration 16 or 17 until 21 and 24 to 30, the results are remain straight line is because the user has input the value to change the reward value of the qTable which the agent is finding the new route to reach the end. When the line straights, that’s mean the agent learns the path.

4.9 Summary

In this chapter, the results and graphs are shown for two different algorithm which is Q-Learning Theory and Q-Learning Theory with human in the loop and also compare the results of both algorithm. Therefore, we also compare the results and graphs of the algorithm with sum of the reward in each episode. There are a input function for user to decide want to change the value of the previous reward with new reward or continue training with the previous reward. The result of the RL without human input is -1964, 78, 80, 84, 82, 86, 90, 84, 92, 86, 88, 90, 88, 90, 88, 92, 92, 90, 92, 92, 92, 90, 92, 92, 92, 92, 92, 92, -118, 92. For RL with human input, the result is 98, 99, 97, 98, 99, 97, 96, 97, 98, 97, 96, 99, 98, 99, 97, 98, 99, 100, 100, 100, 100, 97, 99, 100, 100, 100, 100, 100, 100, 100. After compared the result of both algorithm, it has fulfill the objective of the project which is compared both performance of decision making. For the result presented, we can see that Human in the loop algorithm will make the results become better than baseline code as the agent keeps learned to find a optimal path to reach the endpoint with human input. Three objectives are fulfilled, we have constructed a hybrid intelligence which used reinforcement learning with human input to compare the performance of decision making of reinforcement learning without human input under different circumstances.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In conclusion, this project is to find out the can the hybrid intelligence with human in the loop can enhance the performance of AI to make decision. With the simulation, we have presented Q-Learning Theory and Markov Decision Process (MDP) to train the agent. Then we implement the human in the loop algorithm which is provide a human input function in the Markov Decision Process that the user can decide to input the value to change the result or continue the previous result to move forward.

In this project, we have construct the hybrid intelligence with human knowledge to run the simulation with human input to decide the reward value of the environment to get different result from the baseline code. Therefore, we also use human input to let the user decides the reward value of the simulation which the user wants to change the value of the reward, he or she has to ensure the value cannot be lower than -5 which means user have to input the value from 0, -1, -2, -3, -4 and -5 and the user cannot depend on their feeling to change the value. If the value is more than -5, the agent will take long time to run the simulation. At the result and discussion section, we have compared the result of baseline code which is used reinforment learning method without human input and result of hybrid intelligence is used reinforment learning with human input.

With reinforcement learning, the agent will just follow the Q value of the program and run the simulation. With human in the loop algorithm, human can help agent to decide the training should be continue or change the Q value to achieve a different training results. When human participates in the training state, human cannot decide the decision with their emotion. They have to decide the decision according to the data if the data value is too low or too much or according to the situation of the environment.

5.2 Recommendations for future work

The result has meet the basic requirement of enhance the decision making of AI, but there are still have a lot of improvements that can be improve further, such as the algorithm of human input, environment of the simulation according to the project. From these improvements, we can achieve a better and accurate results.

Human in the loop is very useful in future. It can be use in marketing business system as user preference keep changing over time and the AI might not receive the lastest in a short time. Therefore, human in the loop can reduce the number of failure made by AI agent during the training state to save more time and money.

REFERENCES

MEA (Ministry of Economic Affairs), 2019. *Shared Prosperity Vision 2030.* Putrajaya, Malaysia, 5 October 2019. Najar A and Chetouani M 2021. *Reinforcement Learning With Human Advice: A Survey*, [online] Available at: https://doi.org/10.3389/frobt.2021.584075

Yau KLA, Komisarczuk P, Teal PD. 2012. *Reinforcement Learning for context awareness and intelligence in wireless networks: Review, new features and open issues.* Journal of Network and Computer Applications [e-journal] 35(1):253-267. Available at: https://doi.org/10.1016/j.jnca.2011.08.007

Thiopoulos C, 2020. *Can AI Systems Match Human-Level Situational Awareness?* All, EIT 2020: The Intelligent Revolution. Available at: < https://www.mouser.com/blog/can-ai-systems-match-human-level-situational-awareness>

Sreenath14, 2020. *Getting to Grips with Reinforcement Learning via Markov Decision Process*. [online] Available at https://www.analyticsvidhya.com/blog/2020/11/reinforcement-learning-markov-decision-process/

Paullada A, Raji ID, Bender EM, Denton E, Hanna A, 2021. Data and its (dis)contents: A survey of dataset development and use in machine learning research. Available at https://www.sciencedirect.com/science/article/pii/S2666389921001847#bib17 [Accessed 22 March 2022]

Mahmud H, AKMN Islam, SI Ahmed, K Smolander, 2021. What influences algorithm decision-making? A systematic literature review on algorithm aversion. Available at: https://www.sciencedirect.com/science/article/pii/S0040162521008210 [Accessed 24 March 2022]

Kulkarni S, S.F Rodd, 2020. Context Aware Recommendation Systems: A review of the state of the art techniques. Available at: https://www.sciencedirect.com/science/article/pii/S1574013719301406 [Accessed 23 March 2022]

Mozgovoy M, 2018. Context-Awareness and Anticipation in a Tennis Video Game AI System. Available at: https://ieeexplore.ieee.org/document/8616123/references#references [Accessed 24 March 2022]

MDCR Hernandez, 2021. AI-based Mobile Context-Aware Recommendation Systems from an Information Management Perspective: Progress and Directions.

Available at: https://www.sciencedirect.com/science/article/pii/S0950705121000034 [Accessed 24 March 2022]

Wiethof, C and Bittner, EAC., Hybrid Intelligence – Combining the Human in the Loop with the Computer in the Loop: A Systematic Literature Review. ICIS 2021 Proceedings. 11. Available at: https://www.researchgate.net/publication/356209722_Hybrid_Intelligence_-_Combining_the_Human_in_the_Loop_with_the_Computer_in_the_Loop_A_Systematic_Literature_Review [Accessed 23 March 2022]

Wells L, Bednarz T, 2021. Explained AI and Reinforcement Learning – A Systematic Review of Current Approaches and Trends. Frontiers in Artificial Intelligence 4:550030. Available at: https://www.frontiersin.org/articles/10.3389/frai.2021.550030/full?&utm_source=Email_to_authors_&utm_medium=Email&utm_content=T1_11.5e1_author&utm_campaign=Email_publication&field=&journalName=Frontiers_in_Artificial_Intelligence&id=550030 [Accessed 23 March 2022]

Dellermann D, Calma A, Lipusch N, Weber T, 2019. The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems. Available at: https://www.researchgate.net/publication/332085078_The_Future_of_Human_AI_Collaboration_A_Taxonomy_of_Design_Knowledge_for_Hybrid_Intelligence_Systems [Accessed 23 March 2022]

APPENDICES

Appendix A: Gantt Chart for FYP2

Gantt Chart

No.	Project Activities	Planned Completion Date	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17
1.	Find code for the project	2022-06-30																	
2.	Enhance the code of the project with human input	2022-07-24																	
3.	Compare the algorithm of original code and enhanced code	2022-08-12																	
4.	Write methodology of FYP2	2022-08-19																	
5.	Write Results and Discussion of the project	2022-09-08																	
6.	Poster design	2022-08-28																	

Appendix B: Log Book for FYP 2

Week	Activities to achieve milestones	Submission Date/Status	Ackowledge By	Comments
Week 2	Meeting with Supervisor - discuss about the progress for next few week - find some code and compare the performance for different algorithm	2022-06-27 02:54:13	Reviewed by supervisor	Advised the student to gather baseline results.
Week 4	Meeting with supervisor - have gather the baseline results - show graph of the result - working to explain the graph	2022-07-14 11:18:49	Reviewed by supervisor	Advised to start developing proposed solutions for comparison with baseline.
Week 6	Meeting with supervisor - Question of the proposed solution - Looking for Human input solution	2022-07-20 14:03:52	Satisfactory from Supervisor	Advice is given to incorporate human inputs.
Week 8	Meeting with supervisor -Enlarge the environment -Have human input to control the action of the AI agent	2022-08-10 14:02:20	Reviewed by supervisor	Student is advised to explore incorporating human inputs into the programme through some keys on the keyboard.
Week 10	Meeting with supervisor -Discuss about the report -Highlight the charges from the enhanced code (human input) -provide flowchart of original code and enhanced code (if possible)	2022-08-24 14:31:13	Satisfactory from Supervisor	The student is advised to run simulation and gather results.
Week 12	Meeting with supervisor -Discuss about the report -change some part of the report (objective, literature review and results) -add table and explanation for result	2022-08-30 19:40:44	Satisfactory from Supervisor	Student is progressing well. Report needs to be updated to reflect latest work in the project.