SAVEETHA SCHOOL OF ENGINEERING
SIMATS
MLA03 REINFORCEMENT LEARNING

Engineer to Excel

LIST OF EXPERIMENTS

S. No	Title	CO
1	Induce a Mouse-pile of cheese strategy to get maximum rewards for the mouse in 3 X 4 grid using Bellman Equation using python in a reinforcement Learning environment.	CO1 CO6
2	A Fire of value -1 and Maximum Reward of Value 1 placed on the (1,4) and (2,4) place of matrix and you are placed on the initial block of (1,1) on the matrix, through Reinforcement learning Strategy. Write a program to obtain the maximum reward using python.	CO1 CO6
3	Demonstrate the functions behind state and policies in Reinforcement Learning using Python through a 2 X 2 grid.	CO1 CO6
4	Demonstrate Bell-man equation functionality in Reinforcement Learning using Python through 3 X 3 grid.	CO1 CO6
5	Write a python program using Neural Networks for demonstrating Reinforcement Agent, Environment and Reward.	CO1 CO6
6	Demonstrate the value when exploitation mechanism is implemented into the input matrix of 6 X 4.	CO1 CO6
7	Display and visualize the difference in Learning of Exploitation and Expectation mechanisms by an agent in a Reinforcement Learning Environment using Python Programming.	CO1 CO6
8	Using Tensor flow RL library create an environment, agent and demonstrate Rewards and Punishments within the Reinforcement learning environment	CO1 CO6
9	Using Monte carlo method, induce a reinforcement Learning Environment for getting Maximum reward.	CO1 CO6
10	Consider a news recommendation System has been handled to you, an requirement of making the efficient news to be recommended is taken as the reward, through python implement how you obtain the maximum reward through TD(0) mechanism.	CO1 CO6
11	Induce the concept of dynamic programming and explain the efficiency in terms of computational complexity for reinforcement Learning environment using Python Programming	CO1 CO6
12	Consider you are playing a game of X and O, The System is getting constantly defeated by you. The System decides to enhance SARSA technique to enhance its game play strategies. Write a program for the system to plan with the help of python.	CO1 CO6
13	A Mario game is played by Agent, the agent keeps on moving over, The System decides to tough the levels, how can a system induce Q-Learning technique to enhance its game play strategies and the compiler used for the game is python.	CO2 CO7
14	You are given a 3D realistic environment in a traveler game. With a help of python interpreter and inducing temporal difference strategies, write a program to optimize the selection strategy and attain maximal travel using python program.	CO1 CO6
15	Consider three trains running on respective tracks. Each of the train is based on various algorithms of temporal difference learning say, Train A is induced with TD(0) algorithm, Train B is powered by SARSA algorithm and Train C with Q- Learning. Using a python script reveal which train outperforms the other in terms of efficiency by attaining maximal reward at less number of computational steps.	CO1 CO6
16	Consider you are playing a game of Tic-Tac-Toe, The System is getting constantly defeated by you. The System decides to enhance its reward maximization technique to enhance its game play strategies. Write a program to system for planning a Temporal difference strategy it will choose with the help of python programming.	CO1 CO6

17	Demonstrate the need for Deep - Q- Learning as your autonomous vehicle's detecting efficiency is declining with a help of a program.	CO2 CO7
18	In a game of chess, your opponent wants to carry over mate as soon as possible but you enhance the way of deep - Q- learning method of handling the game, explain why and how will you win through a Python Program?	CO2 CO7
19	Consider you are the manager of a Finance Company; the target of the month has not been achieved and you are in trouble. You come to know that your Q Learning System not performing well as the numbers of customers have increased, the correction decision would be increasing the layers of the DQN. Write a program to enhance DQN and transform it into DDQN	CO2 CO7
20	You are given a 3D realistic environment in a traveller game. With a help of python intrepreter and inducing temporal difference strategies, explain the increase in performance difference between a DQN and DDQN through python.	CO2 CO7
21	In a maze game, the agent keeps on moving over, The System decides to tough the levels, how can a system induce a Duel DQN to enhance its game play strategies. The compiler used for the game is python.	CO2 CO7
22	You run Book my Show application, a requirement of making the efficient mapping of user interests to be recommended is the need of the hour, Your higher officials suggest to increase the depth of the Q- Network but you defend them by saying Duelling would help. Prove them by providing the code of lines using python.	CO2 CO7
23	Consider two robots are running on respective tracks. Each of the robots is built with the various Q-learning models say, Robot 1 is induced with DQN, and Robot 2 is powered by DDQN and Robot 3 with DDDQN. On using a python script reveal which Robot would outperforms the other in terms of efficiency by attaining maximal reward at less number of computational steps.	CO2 CO7
24	Consider we are playing a game of Mice and Cheese. The Mice is getting only single pile of cheese instead of the group of cheese. The System decides to enhance its reward maximization technique to enhance its game play strategies. Enhance a Q-table and train the mice to get the Big pile of cheese.	CO1 CO6
25	Consider you are playing a game of Mice and Cheese. The Mice is located next to the poison block. The System wrongly detects Minimum reward and instructs the mouse to move over to the minimal Reward block. What will save the mouse by inducing the Duelling DQN network strategy to enhance and maximize its reward maximization technique to enhance its game play strategies? Enhance a Q-table and train the mice to get the Big pile of cheese.	CO2 CO7
26	You are Mario in the given game and the main task is to guide the queen to castle, you have been instructed to use the Prioritised Experience Replay Strategy to guide the queen from obstacles and reach castle at the earliest move, Write the code for the solution.	CO1 CO6
27	A Scientist in a product development company which produces a stick which helps Visually Challenged People, the company wants to optimize the cost of the stick as the rule-based instructions make them higher time complexity, so they ask the Scientist to develop a stick based on the learning strategy for the model which is to be based on the previous experiences present in the video recordings of the dataset. Help the scientist to successfully tweak the code and induce the strategy.	CO1 CO6
28	A Scientist in a Autonomous Car development company which produces self-drive cars. But some obstacles are shown wrongly leading do driving complexity, so they ask the Scientist to develop a Robust Learning algorithm with the help of deep Q learning strategies. You advise the Scientist to decide whether its DQN or DDQN or Dueling DQN strategy to overcome this problem addressed with the help of python code.	CO2 CO7
29	Consider a line following robot are running on a line. The robot’s efficiency is very less so enhance the robot’s steps by trial and error method by inducing a Vanilla Policy Gradient mechanism.	CO1 CO6
30	You are a Stock Market advisor, now there is a need to develop a learning engine which will advice you get maximum Profit investment through Probabilistic values of the historical data processing, Use Vanilla Policy Gradient for structuring the	CO2 CO7

	highest return Policies.	
31	A robot dog you are building that is to be running on a zig zag lines. The robot's efficiency is very less as it needs to enhance both actions simultaneously and asynchronously, propose a method how you will induce an algorithm to make it work efficiently using Actor-critic method.	CO3 CO6
32	We are instructed by your mentor to build a stop clock which will be running asynchronously showing variation of different timing around the world. Now, we need to find which of two methods will be suitable for the development whether A2C or A3C through experimentation in python.	CO3 CO6
33	We are instructed by your mentor to build a stop clock which will be running asynchronously showing variation of different timing around the world. Implement A2C or A3C and give the Optimal policy designing framework.	CO3 CO6
34	We are a Stock Market advisor, now there is a need to develop a learning engine which will advice you get maximum Profit investment through Probabilistic values of the historical data processing, where Vanilla Policy Gradient is returning Lower values, so move onto the Advance Gradient policy Methods and decide which gives us the best output.	CO1 CO6
35	We run Google Maps, a discrepancy of less quality policies are returning to customers which makes them to choose low optimal strategies, CEO advises you to choose PPO for inducing optimality. Prove it through python coding.	CO4 CO6
36	We are running a Money lending company, now there is a need to develop a learning engine which will advise we get maximum Profit investment through Probabilistic values of the historical data processing of the Customers, where Vanilla Policy Gradient is returning Lower values, so move onto the Advance Gradient policy Methods and decide which gives you the best output.	CO2 CO7
37	In a maze game is, the agent keeps on moving over, but still not retained an Optimal Policy. As we are the developer of the Game keep an update to Policy options through DDPG algorithm	CO4 CO6
38	We are driving a bus in Simulation environment, a discrepancy of less quality policies are returning you a low value points in your simulation Quality which makes them to choose low optimal strategies, based on the necessity we decide to choose DDPG for inducing optimality. Prove it through python code.	CO3 CO6