Page No.
Date

Experiment NO:-5

Aim:- Reinforcement learning

(a) Calculate Reward
(b) Discounted Reward
(c) Calculating optimal
(d) Implementing Q learning
(e) Setting up an Optimal Action.

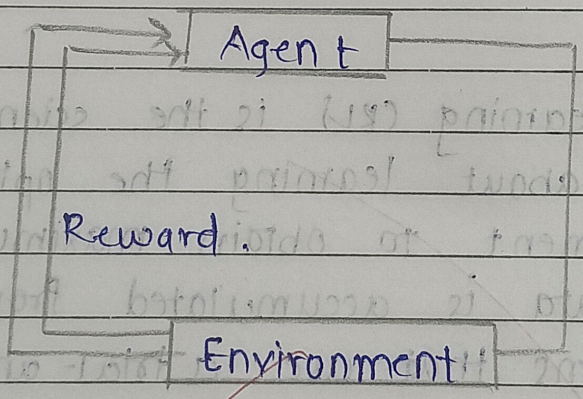
Objective:- To study and implement
(a) Calculating Reward
(b) Discounted Reward
(c) Calculating optimal
(d) Implementing Q learning
(e) Setting up an Optimal Action.

Theory:-
Reinforcement learning (RL) is the science of decision making. It is about learning the optimal behaviour in an environment to obtain maximum reward. In RL, the data is accumulated from machine learning systems that use a trial-and-error method. Data is not part of the input that we would find in supervised or unsupervised machine learning.

Reinforcement learning uses algorithms that learn from outcomes and decide which action to take next. After each action, the algorithm receives feedback that helps it determine whether the choice it made was correct, neutral or incorrect. It is a good technique to use for automated systems that have to make a lot of small decisions without human guidance.

Reinforcement learning is an autonomous, self-learning system that essentially learns by trial and error. It performs actions with the aim of maximizing rewards, or in other words, it is learning by doing in order to achieve the best outcomes.

(a) & Diagram:-



(a) Calculating Reward:-

In Reinforcement learning, the reward is a numerical signal that indicates how well an agent is performing in its environment. The calculation of rewards can vary depending on the specific problem domain.

b) Discounted Reward:-
Discounted Reward, also known as discounted return, is a concept in reinforcement learning where future rewards are given less importance compared to immediate rewards. This is done by introducing a discount factor, typically denoted by γ (gamma), which is a value between 0 and 1. The discount factor determines the importance of future rewards relative to immediate rewards.

c) Calculating Optimal:
Calculating the optimal policy or optimal value function in reinforcement learning refers to finding the best strategy or set of actions that maximizes the cummulative reward over time. This is typically done through value iteration or policy iteration algorithms. Value iteration iteratively updates the value function until convergence, while policy iteration

d) Implementing Q-learning:-
Q-learning is a popular model-free reinforcement learning algorithm that aims to find an optimal policy for a given environment. The Q-learning algorithm learns an action-value function, denoted as Q(s,a), which represents the expected cumulative reward when taking action 'a' in state 's' and following the optimal policy thereafter.