**Experiment :- 9**

|  |
| --- |
| **Title:**  **Write a program to implement approximate value iteration (AVI) algorithm and API (Approximate value iteration)** |

# Objective:

# Books/ Journals/ Websites referred:

* Markov Decision Processes in Artificial Intelligence MDPs, Beyond MDPs and Applications, Edited by Olivier Sigaud, Olivier Buffet, Wiley Publications, 2010
* https://gym.openai.com/docs/
* https://www.analyticsvidhya.com/blog/2018/09/reinforcement-learning-model-based-planning-dynamic-programming/

# Resources used:

# Python

# Gym

# Theory:

Approximate value iteration (AVI) is a reinforcement learning algorithm that seeks to approximate the optimal value function in a given environment. The algorithm iteratively updates an estimate of the value function by bootstrapping from estimates of future values.

At each iteration, AVI updates the value function estimate by computing the expected value of the next state, given the current state and the available actions. This expected value is estimated using a function approximator, such as a neural network or a linear function approximator. The function approximator is trained to minimize the difference between the predicted and actual values of the next state.

The key advantage of AVI over other value iteration methods is that it can handle large state spaces and continuous state and action spaces. This is because it uses a function approximator to estimate the value function, rather than computing the exact value for each state-action pair.

The Approximate Value Iteration (AVI) algorithm is a reinforcement learning algorithm that combines ideas from dynamic programming and function approximation to compute an approximation of the optimal value function for a given Markov decision process (MDP).

The basic idea behind AVI is to iteratively improve an initial estimate of the value function using a combination of Bellman's equation and function approximation. The algorithm starts with an initial estimate of the value function, and then iteratively updates this estimate until it converges to a fixed point.

Here is the high-level pseudocode for the AVI algorithm:

1) Initialize the value function V(s) for all states s.

2) Repeat until convergence:

2.1) For each state s, compute the TD target:

TD\_target(s) = R(s) + gamma \* max\_a V(s')

2.2) Update the value function using function approximation:

V(s) = approximator.fit(features(s), TD\_target(s))

2.3) Repeat steps a-b for a fixed number of iterations or until convergence.

In the above pseudocode, R(s) is the immediate reward obtained for being in state s, gamma is the discount factor, max\_a represents the maximum value over all possible actions a in state s', approximator is a function that approximates the value function V(s), and features(s) is a feature vector that describes state s.

During each iteration of the AVI algorithm, the TD target is computed using Bellman's equation, which computes the expected value of being in a given state and taking the best action according to the current estimate of the value function. The value function is then updated using function approximation, which involves fitting a model to the observed data (i.e., the TD targets) and the feature vectors.

The AVI algorithm has been widely used in various applications such as robotics, control systems, and game playing, among others.

# Implementation (Code):

**Kindly find code and output in E9\_AVI.ipynb**

# Output Screenshots:

# Conclusion (Students should write in their own words):

# Overall, AVI is a powerful and versatile algorithm that can be used in a variety of reinforcement learning settings. However, the accuracy of the value function estimate depends on the quality of the function approximator and the choice of hyperparameters.

# Applications:

# The Approximate Value Iteration (AVI) algorithm is a popular reinforcement learning algorithm used in many real-world applications. Some of its applications include:

# 1) Game playing: AVI has been used for game playing applications, such as in the development of computer programs that can play games like backgammon, chess, and Go. In these applications, the AVI algorithm is used to learn the optimal value function and policy for the game, which allows the program to make intelligent moves

# 2) Robotics: AVI is also used in robotics applications, such as in the development of robots that can navigate and manipulate objects in their environment. In these applications, the AVI algorithm is used to learn the optimal value function and policy for the robot, which allows it to perform tasks such as picking and placing objects.

# 3) Finance: AVI is used in finance applications, such as in the development of automated trading systems. In these applications, the AVI algorithm is used to learn the optimal value function and policy for trading, which allows the system to make intelligent trading decisions.

# 4) Autonomous vehicles: AVI is also used in the development of autonomous vehicles. In these applications, the AVI algorithm is used to learn the optimal value function and policy for the vehicle, which allows it to navigate roads and avoid obstacles.

# Overall, the AVI algorithm is a powerful tool for developing intelligent systems that can learn optimal behaviors in complex environments. Its applications are widespread, and it continues to be an active area of research in the field of reinforcement learning.