

Assignment

Reinforcement Learning

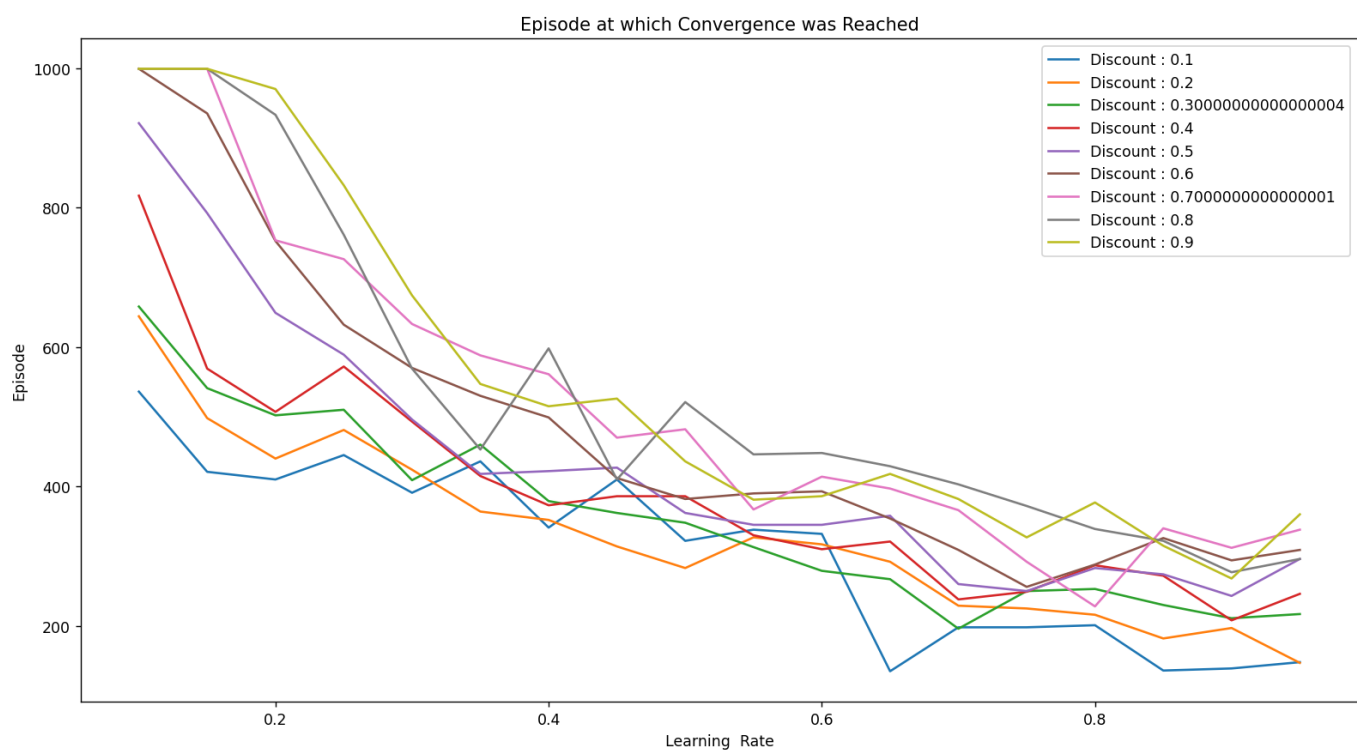
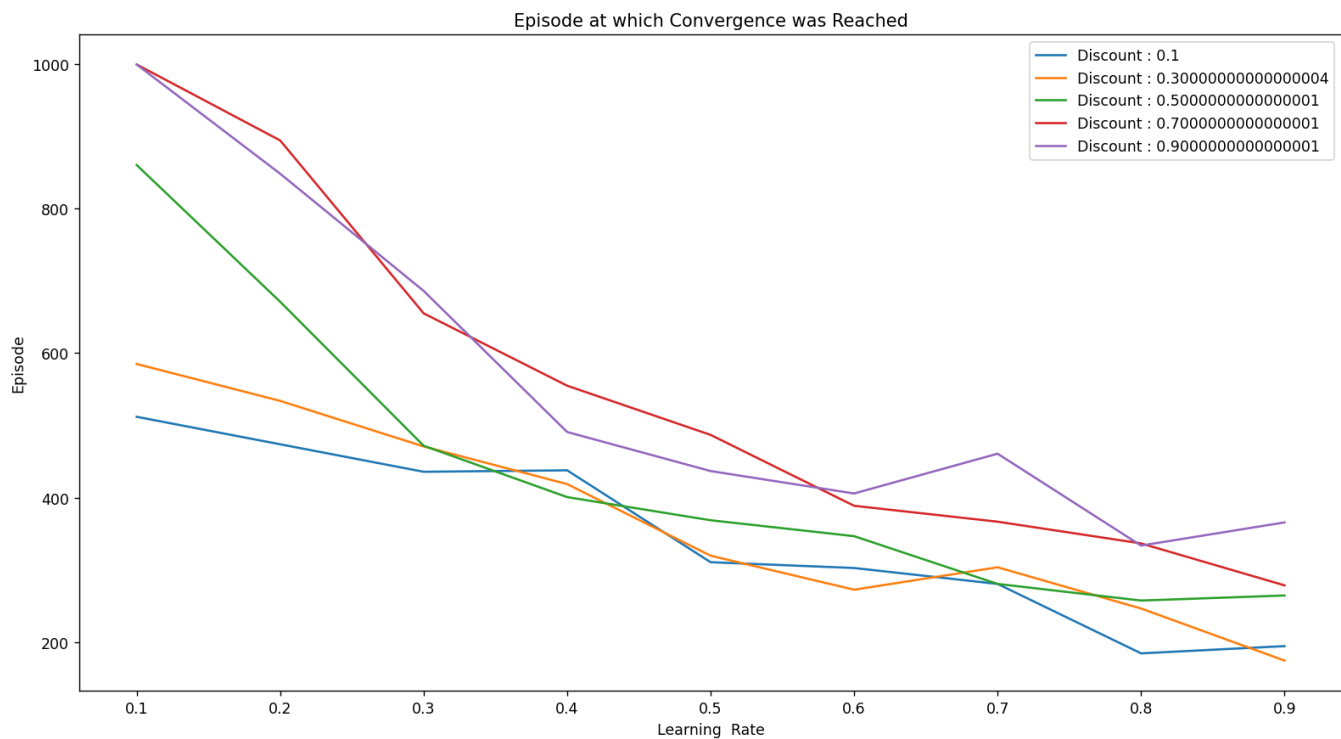
GROUP 36

Rafshan M.R.M 210501B

Ranasinghe A.A.D.H 210515V

Ranasinghe K.S. 210518H

Ranasinghe R.P.C.G.J 210520G



By analyzing the above graphs we can see the following;

- Decreasing the Discount Rate reduces the number of episodes needed for convergence.
- Increasing the Learning Rate reduces the number of episodes needed for convergence.

Reasoning

- Low Discount Rate prioritizes immediate rewards, making the agent myopic and focusing on short-term gains making the learning process quicker. But there is also a possibility of this leading to suboptimal policies as the agent ignores long-term consequences, especially in tasks where delayed rewards play a crucial role.
- High Learning Rate converges faster as it rapidly updates Q-values. But this implementation can fail to converge to the best policy and instead overshoot the ideal values. Because of its sensitivity to noisy incentives, it may diverge or oscillate around the best solution.