Module 4 : Learning and Improving the Tutor using Reinforcement Learning

1. Objective

The goal of this module is to build an intelligent tutor that learns how to respond to a student based on their emotional state and performance. Instead of using fixed rules, the system uses reinforcement learning so it can improve over time and choose the best action such as motivating, giving hints, or adjusting difficulty to help the student learn better and remain emotionally stable. The main focus is to improve the student's emotional well-being. If the student's emotion improves or stays positive, the model receives a higher reward. If the student's emotion becomes worse, the model receives a negative reward.

2. Input Variables

Emotions = happy, neutral, anger, boredom, deeply anxious, overthinking, losing motivation, slightly stressed.

Performance = good, average, bad, unanswered, excellent.

Actions = motivate, give_hint, clarify_again, reduce_difficulty, increase_difficulty, idle.

3. Non-Trivial Test Case

For testing my module, I designed a non-trivial experiment where the Q-learning tutor was trained for a total of 6000 episodes. In each episode, the tutor interacted with the simulated student for 8 steps. I used a learning rate (α) of 0.25, a discount factor (γ) of 0.95, and an exploration value (ε) of 0.45, which slowly decreases over time so the model explores less and starts using what it has learned. After the training was completed, I tested the performance of the learned model Q-Tutor by running it for 1000 episodes. I also tested the fixed-rule policy for the same number of episodes so I could compare how well the learned model performs against a manually designed rule-based system.

4. Actual Output Produced by the Code

The output of code contains the mean reward and standard deviation of both Q-tutor and fixed rule based model, and it tells how many times both Q-tutor and fixed rule based model used each action. The model is trying different actions in every possible student emotion and performance state during training. Each time it takes an action, it receives a reward based on whether the student's emotional state and performance improve or get worse.