Reinforcement Learning

- We don't tell the learner which actions to take, but let the agent discover the actions yielding the best reward by experimenting in a simulation

- The training system gives the agent input from the environment and rewards the agent when the outcome of the simulation (or game) being played is positive.

- You assign a reward for a given behavior and overtime, the agents learn to reproduce that behavior in order to receive more rewards.

- Some kind of evaluation signal that comes from the environment specifying how well you are performing in this task.

- We need evaluation measure- somebody clapping, falling down(it will be translated to numeric scale)

- The goal of agent is to learn a policy which is a mapping of states to actions, to maximize measure of long term performance.