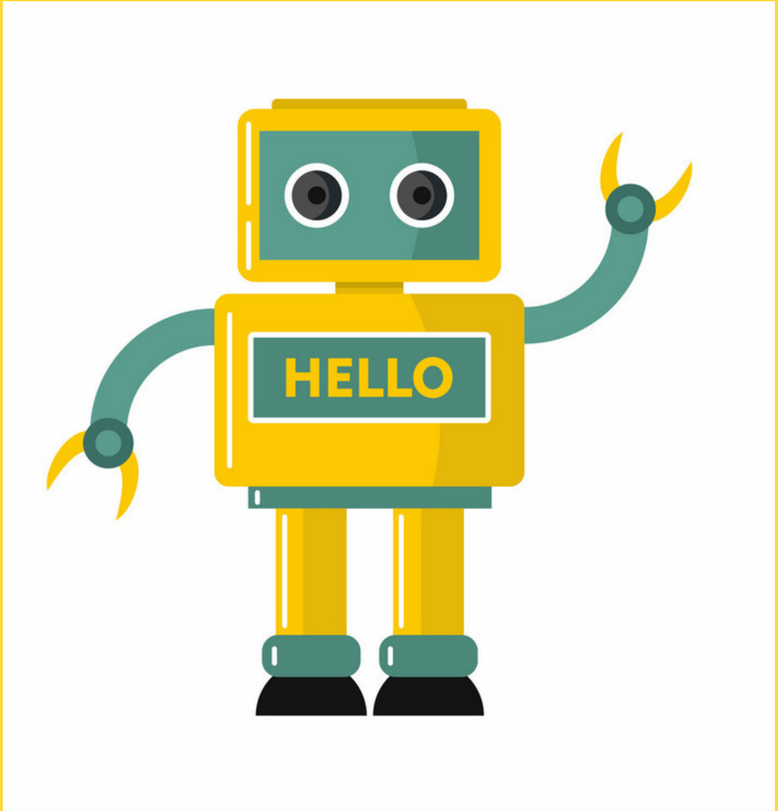
Reinforcement Learning

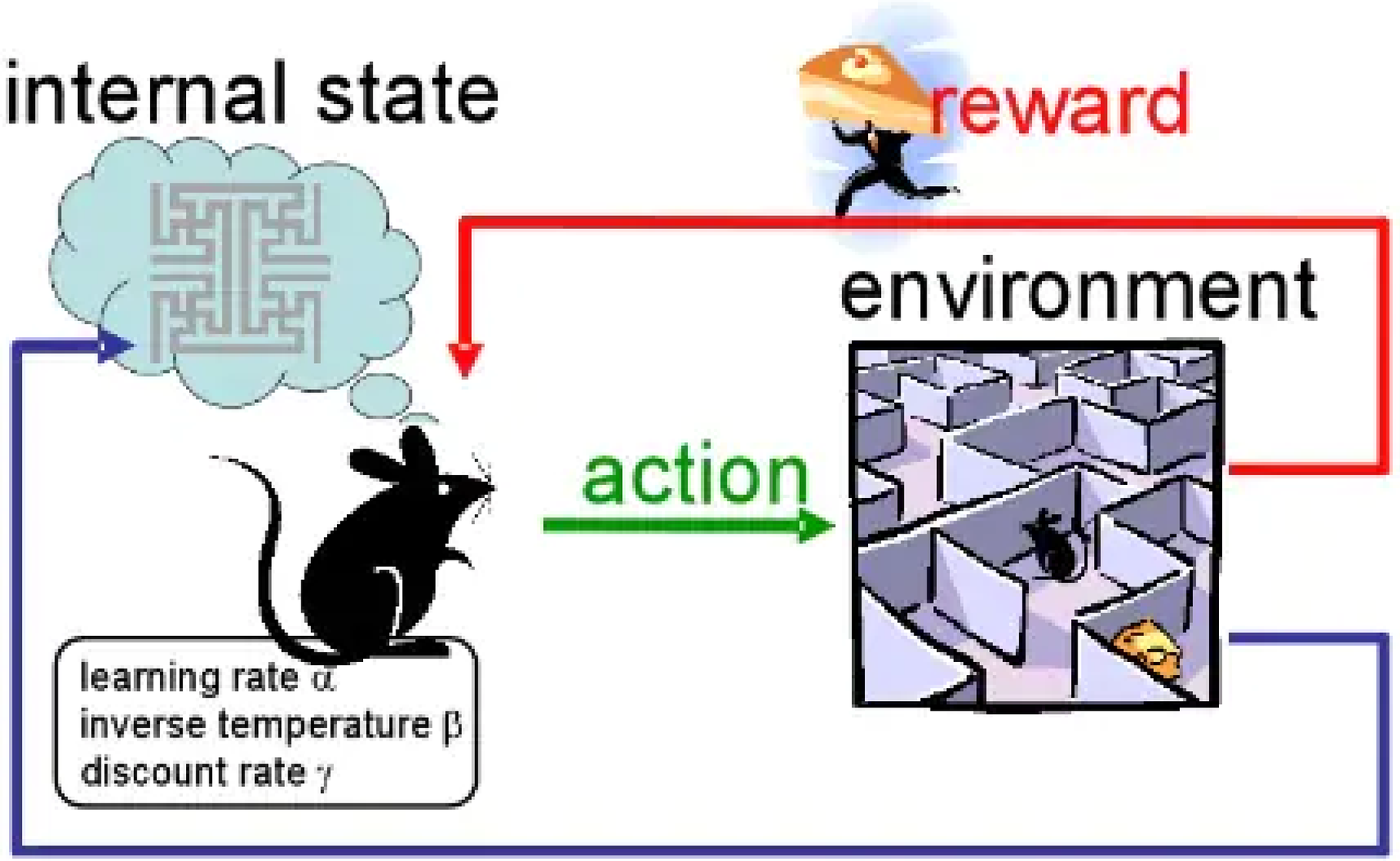
GWOC'21 Contribution



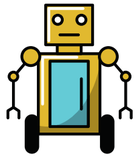
Content:

- RL defintion
- pre- requisutes
- RL vocabularly
- RL in action
- Types
- Learning Models
- Limitations
- Comparison
- Real life application
- conclusion
- Algorithms

What is reinforcement learning?

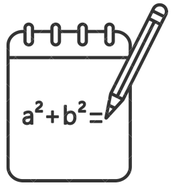


Pre-requistes



HIGHLIGHT 1

Proficiency in Python



HIGHLIGHT 2

Calculus, Linear Algebra

HIGHLIGHT 3

Basic Probability and Statistics

HIGHLIGHT 4

basic knowledge on machine learning

Types

2 types of RL are:

- Positive reinforcement is a reward for doing something well. It strengthens the occurrence of the action that will generate more reward

- Negative reinforcement occurs when an aversive stimulus (a ‘bad consequence’) is removed after a good behavior is exhibited. Set the minimum action required to achieve the reward or assigned value function.

Algorithms

3 algorithms in RL

- Value based: tries to set the maximum value function for the agent. The agent gains the reward in return.
- Policy based: This RL method devise a policy to the agent by which maximum reward can be obtained. E.g., SARSA (State-action-reward-state-action)
- Model based: A simulated, or artificial environment is provided to the agent to perform the action.

Limitations

challenges of RL

- agent needs to be trained in the simulated environment and devise a strategy to achieve the provided target. It is however in need of large amount of time and expected trail and error come in its way. In real life situation, this become a matter of concern especially in autonomous vehicle where real environment changes swiftly and health risk safety concerns are paramount. Also, the algorithms once fed to agent cannot be altered except of changing parameters like reward and punishment hence it creates an issue when modification to previously defined algorithm is needed.

Learning Models

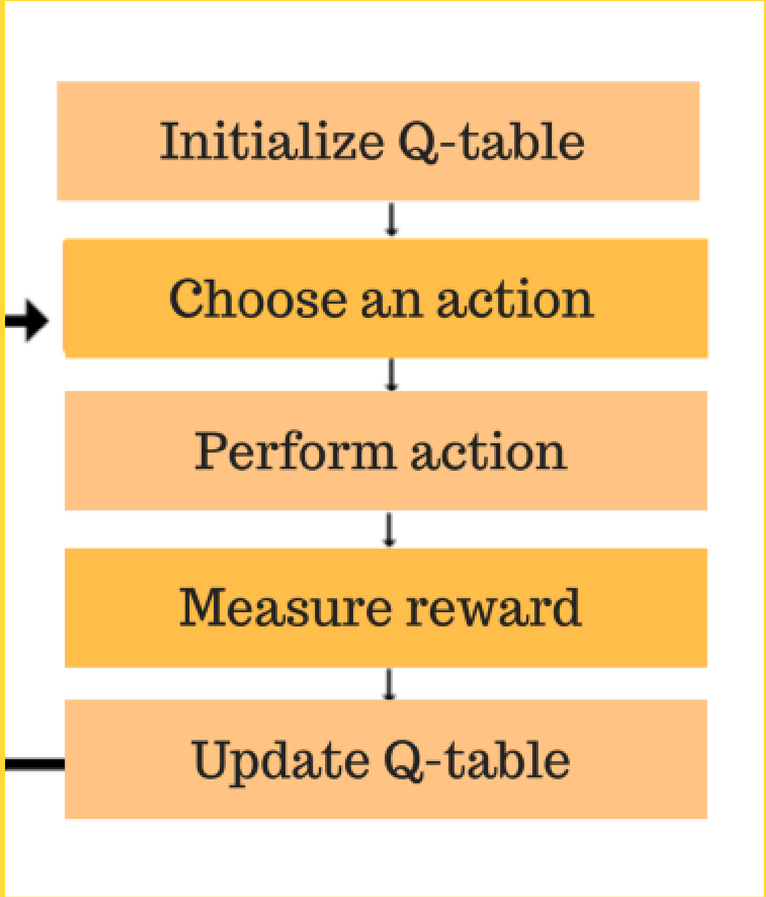
MDPs

Markov Decision Process

Markov's decision process or MDPs is based upon the long-term impact of the choice. Markov's property state that the reward of action is dependent upon present condition only. MDP consist of 5 components

1. Agent
2. State
3. Environment
4. Action
5. Reward

Q-learning



Comparison

Supervised ML

In supervised learning, model is trained by providing the labelled dataset

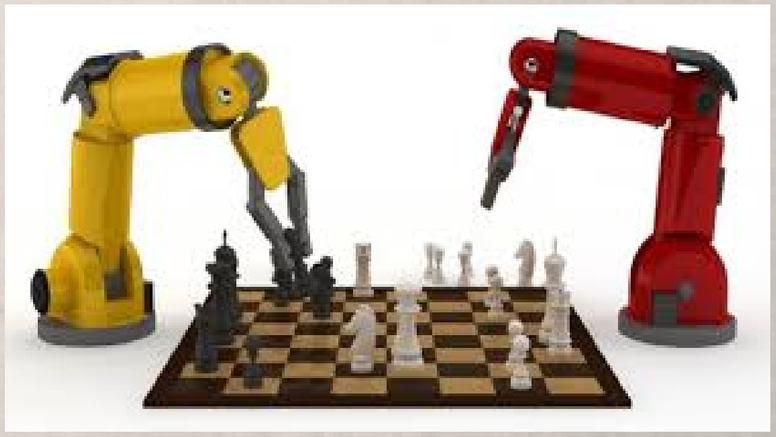
Unsupervised ML

In unsupervised learning, large unlabelled data is fed into system with the set of complex algorithms that enable that model to itself distinguish between different elements of dataset provided.

Reinforcement Learning

RL takes an entirely different approach, it provided the agent with defined targets and rewards and punishment and agent must develop the strategy to gain rewards and minimize penalty while achieving the target during its interaction with the environment. Due to resonance with supervised learning on the provision of certain defined parameters to the machine, it is also referred to the subset of supervised learning.

Applications



Conclusion

- RL is the subset of machine learning that require an agent to be fed by complex algorithms, perform stipulated target, gain the rewards, and avoid penalties.
- RL is often beneficial when large no of dataset is unavailable
- downside is that it increases the time lapse of the execution of the action.
- The most widely known application of RL is gaming like AlphaGo, Chess etc.