**Batch: ML1 Roll No.: 1911004**

**Experiment / Assignment / Tutorial No. 8**

|  |
| --- |
| **Title: Case study of reinforcement learning application** |

**Aim:** Report writing on reinforcement learning case study.

**Expected Outcome of Experiment:**

CO5: Describe the use of neural networks and reinforcement learning.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Books/ Journals/ Websites referred:**

[**https://arxiv.org/pdf/1811.00260.pdf**](https://arxiv.org/pdf/1811.00260.pdf)

[**https://analyticsindiamag.com/facebook-reinforcement-learning-platform-horizon-is-shaping-the-future-of-intelligent-interaction/#:~:text=In%20Opinions-,Facebook's%20Reinforcement%20Learning%20Platform%20'Horizon'%20Is%20Shaping,The%20Future%20Of%20Intelligent%20Interaction&text=Reinforcement%20Learning%20is%20a%20reward,in%20state%20of%20the%20environment**](https://analyticsindiamag.com/facebook-reinforcement-learning-platform-horizon-is-shaping-the-future-of-intelligent-interaction/#:~:text=In%20Opinions-,Facebook's%20Reinforcement%20Learning%20Platform%20'Horizon'%20Is%20Shaping,The%20Future%20Of%20Intelligent%20Interaction&text=Reinforcement%20Learning%20is%20a%20reward,in%20state%20of%20the%20environment)**.**

**What is reinforcement learning ?**

Reinforcement learning is an area of Machine Learning. It is about taking suitable action to maximize reward in a particular situation. It is employed by various software and machines to find the best possible behaviour or path it should take in a specific situation. Reinforcement learning differs from supervised learning in a way that in supervised learning the training data has the answer key with it so the model is trained with the correct answer itself whereas in reinforcement learning, there is no answer but the reinforcement agent decides what to do to perform the given task. In the absence of a training dataset, it is bound to learn from its experience.

**Name of application & details:**

In the engineering frontier, Facebook has developed an open-source reinforcement learning platform — **Horizon.** The platform uses reinforcement learning to optimize large-scale production systems. Facebook has used Horizon internally:

1. to personalize suggestions
2. deliver more meaningful notifications to users
3. optimize video streaming quality.

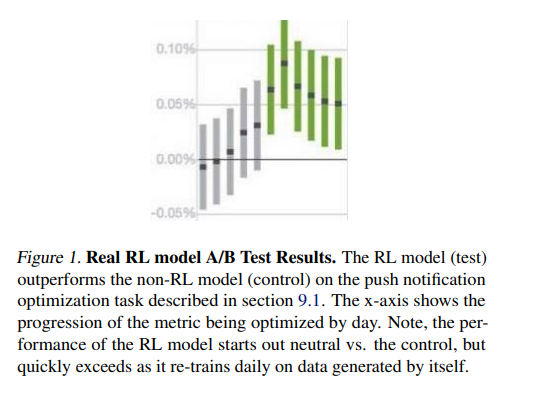
**Horizon** also contains workflows for:

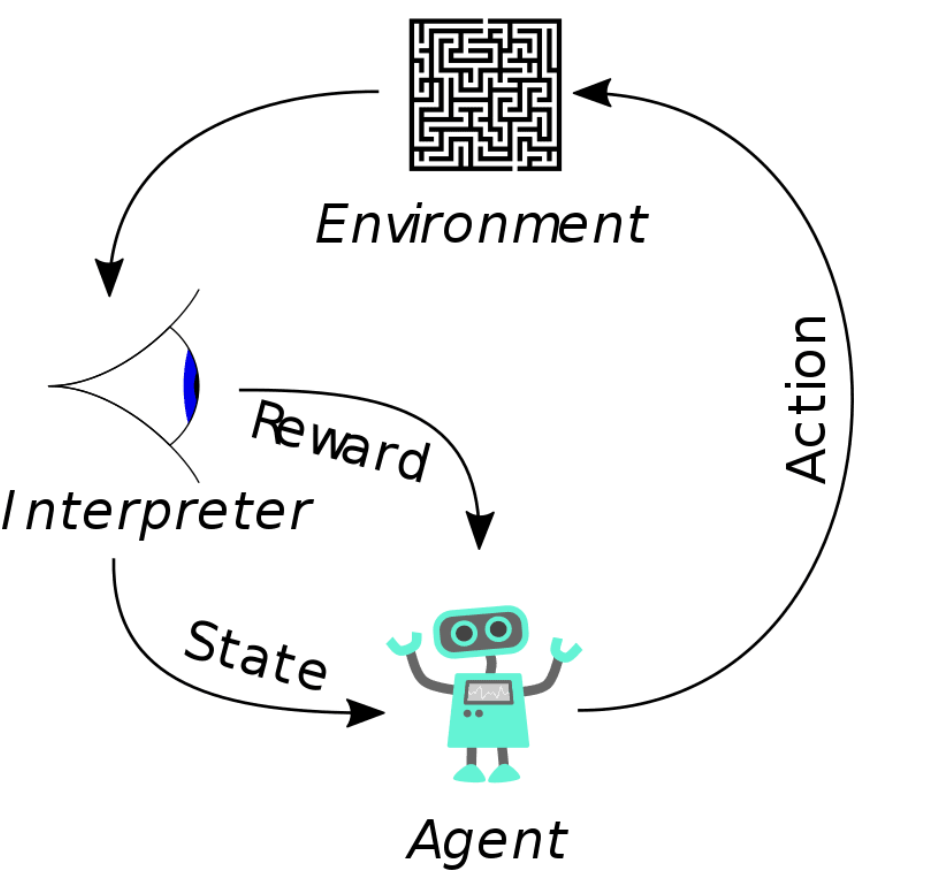
1. simulated environments
2. a distributed platform for data pre-processing
3. training and exporting models in production.
4. A classic example of reinforcement learning in video display is serving a user a low or high bit rate video based on the state of the video buffers and estimates from other machine learning systems.

Horizon trains a **Discrete Action Deep Q-Networks (DQN)** to learn a policy to determine whether to send a notification or not based on the state represented by many features. The Markov Decision Process (MDP) here is based on a sequence of notification candidates for a particular person. The actions here refers to sending and dropping of notifications and the state describes the set of features about the person and the notification candidate and rewards are given for each interaction and activity on Facebook with a penalty to control the volume of notifications sent. This compared to the Supervised model optimizes for the long-term value and is capable of capturing incremental effects on sending the notification by comparing the Q-values of the send and don’t send action.

**Horizon** is capable of handling production-like concerns such as:

1. deploying at scale
2. feature normalization
3. distributed learning
4. serving and handling datasets with high-dimensional data and thousands of feature types.

****

****

**Type of Reinforcement learning in application:**

Facebook's Reinforcement Learning Platform 'Horizon' Is Shaping The Future Of Intelligent Interaction. Reinforcement Learning is a **reward-based learning algorithm** in which an agent or a system interacts with its environment, looks for a reward and adapts to the change in state of the environment**.**

**Advantages of Horizon :**

Horizon was designed considering the following principles:

1. Ability to handle large datasets efficiently
2. Ability to pre-process data automatically and efficiently
3. Competitive algorithmic performance
4. Algorithm performance estimates before launch
5. Flexible model serving in production
6. Platform reliability

**A new problem that can be solved using reinforcement learning and how:**

**The strength of RL algorithms is being applied in solving various business scenarios in the real world where task automation is required.**

**Manufacturing**: Manual tasks of manufacturing which usually require tremendous labour hours and human efforts are performed with automated robots with high accuracy and speed. A Japanese company named Fanuc manufactures robots who can self-learn for a broader range of industries. The robots made by this company can pick the right objects out of a box with few annotations and sensor technology, thus lessening the training efforts drastically.

**Resource Optimization:** Creating solutions for resource management tasks such as allocating computers to several awaiting jobs can be challenging, requiring human intervention. RL algorithms can be effectively used to learn about the vacancy and allocate resources to the waiting jobs, resulting in less delay.

**Conclusion:** Thus, we have successfully learnt & understood about what reinforcement learning is, and one of the applications of Reinforcement learning in Facebook’s Horizon.