- ## Reinforcement Learning

A major drawback of machine learning is that a tremendous amount of data is needed to train models. The more complex a model, the more data it may require. But this data may not be available to us. It may not exist or we simply may not have access to it. Further, the data collected might not be reliable. It may have false or missing values or it might be outdated.

Also, learning from a small subset of actions will not help expand the vast realm of solutions that may work for a particular problem. This is going to slow the growth that technology is capable of. Machines need to learn to perform actions by themselves and not just learn from humans.

All of these problems are overcome by reinforcement learning. In reinforcement learning, we introduce our model to a controlled environment which is modeled after the problem statement to be solved instead of using actual data to solve it.

Reinforcement learning is a sub-branch of Machine Learning that trains a model to return an optimum solution for a problem by taking a sequence of decisions by itself.

We model an environment after the problem statement. The model interacts with this environment and comes up with solutions all on its own, without human interference. To push it in the right direction, we simply give it a positive reward if it performs an action that brings it closer to its goal or a negative reward if it goes away from its goal.

To understand reinforcement learning better, consider a dog that we have to house train. Here, the dog is the agent and the house, the environment.

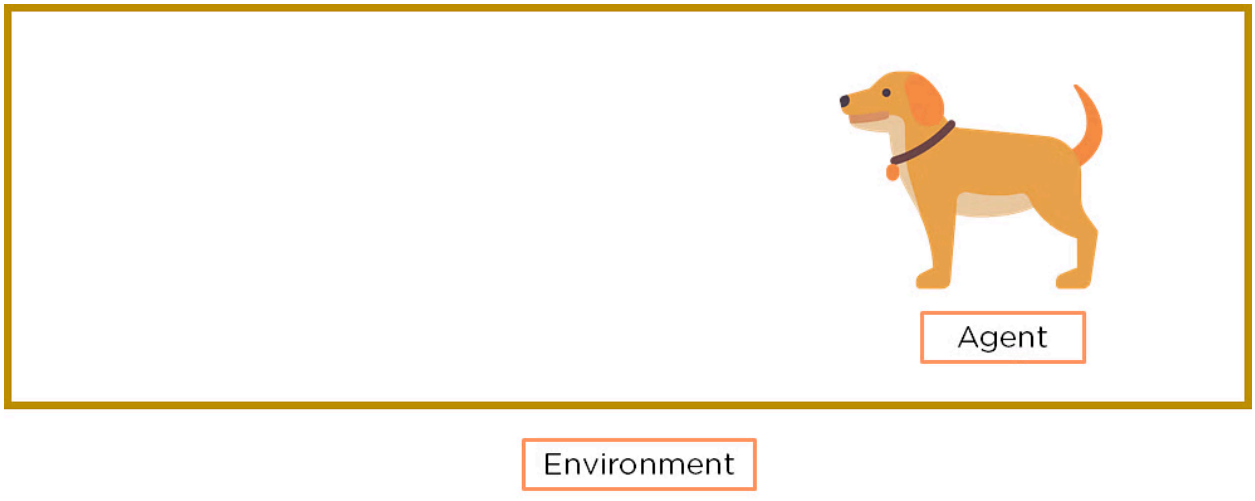

Figure 1: Agent and Environment

We can get the dog to perform various actions by offering incentives such as dog biscuits as a reward.

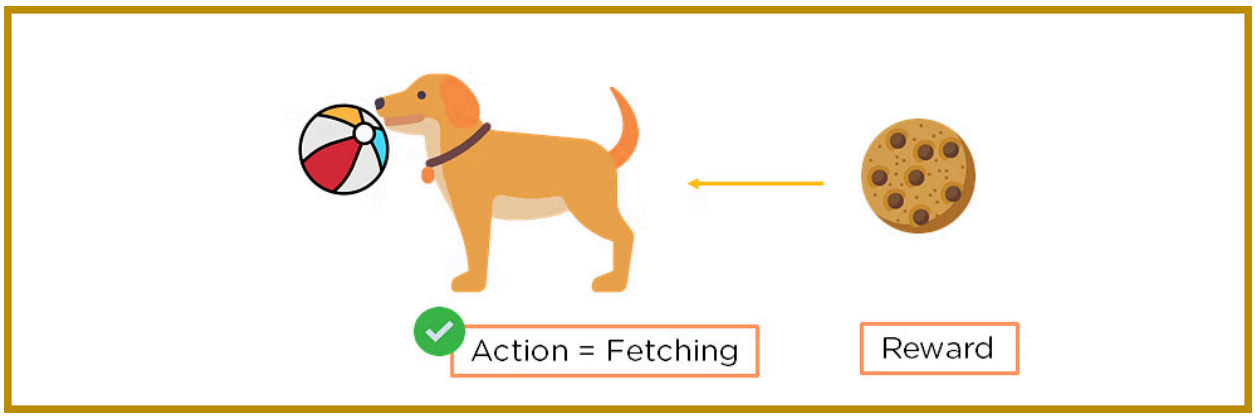

Figure 2: Performing an Action and getting Reward

The dog will follow a policy to maximize its reward and hence will follow every command and might even learn a new action, like begging, all by itself.

Figure 3: Learning new actions

The dog will also want to run around and play and explore its environment. This quality of a model is called Exploration. The tendency of the dog to maximize rewards is called Exploitation. There is always a tradeoff between exploration and exploitation, as exploration actions may lead to lesser rewards.