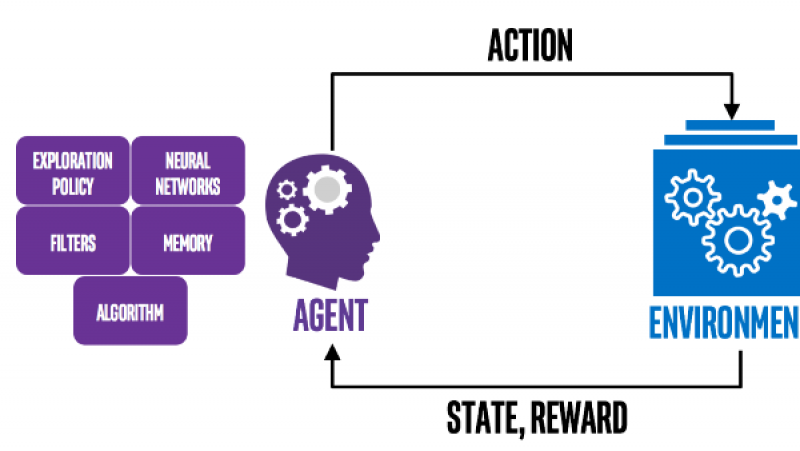
**What is Reinforcement Learning?**

* **923ACTIVITIES**

Reinforcement Learning is an area of Machine Learning that allows machines and software agents to automatically determine the ideal behavior within a specific context, in order to maximize its performance. [Software agents](https://www.google.com/url?q=https://en.wikipedia.org/wiki/Software_agent&sa=D&ust=1563173776718000) take [actions](https://www.google.com/url?q=https://en.wikipedia.org/wiki/Action_selection&sa=D&ust=1563173776718000) in environments to maximize some notion of cumulative reward. Simple reward feedback known as the reinforcement signal is required for the agent to learn its behavior.

Reinforcement learning is one of three basic machine learning paradigms, alongside [supervised learning](https://en.wikipedia.org/wiki/Supervised_learning&sa=D&ust=1563173776719000) and [unsupervised learning](https://en.wikipedia.org/wiki/Unsupervised_learning&sa=D&ust=1563173776719000).

Reinforcement learning contrasts with other machine learning approaches in that the algorithm is not explicitly told how to perform a task, but works through the problem on its own.

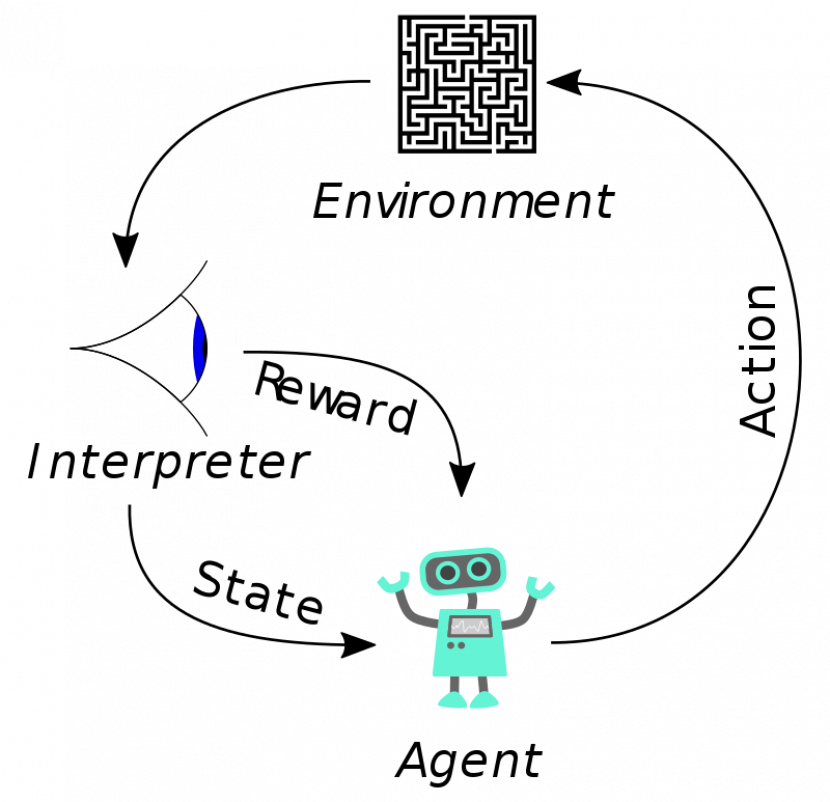
There are many different algorithms for Reinforcement Learning. It is defined by a specific type of problem and all its solutions are classed as Reinforcement Learning algorithms.

In the problem, an agent is supposed decide the best action to select based on his current state. When this step is repeated or [dynamic programming](https://en.wikipedia.org/wiki/Dynamic_programming&sa=D&ust=1563173776720000) techniques are utilized, the problem is known as a Markov Decision Process (MDP).

The main difference between the classical dynamic programming methods and reinforcement learning algorithms is that reinforcement learning algorithms do not assume knowledge of an exact mathematical model of the MDP and they target large MDPs where exact methods become infeasible.

Reinforcement learning problems involve an agent interacting with an environment. The agent must learn about the environment and must also discover how to act optimally in that environment. There are both statistical component (learning about the environment) and computational component (deciding how to act).

The typical frame of a Reinforcement Learning whereby an agent takes actions in an environment, which is interpreted into a reward and a representation of the state, which are fed back into the agent, is shown below:



EXAMPLE 1: SELF-DRIVING CAR

As an agent, it could be a self-driving car that interacts with its environment, receives a reward state depending on how it performs, such as driving to destination safely .Conversely, the agent receives a penalty for performing incorrectly, such as going off the road or violating traffic rules.



EXAMPLE 2: ONLINE CHESS

As an agent, it could be a program playing chess which interacts with its environment, receives a reward state depending on how it performs, such as winning a game. Conversely, the agent receives a penalty for performing incorrectly, such as being checkmated.

The agent in both the examples over time makes decisions to maximize its reward and minimize its penalty using dynamic programming. The advantage of this approach is that it allows an AI program to learn without a programmer spelling out how an agent should perform the task.

**The reinforcement learning model consists of:**

1. A set of environment and agent states S.
2. A set of actions A of the agent.
3. Policies of transitioning from states to actions.
4. Rules that determine the scalar immediate reward of a transition.
5. Rules that describe what the agent observes.

A task is defined by a set of states, s∈S, a set of actions, a∈A, a state-action transition function,

T: S×A→S, and a reward function, R: S×A→R. At each time step, the learner (also called the agent) selects an action, and then as a result, given a reward and its new state.

The goal of reinforcement learning is to learn a policy, a mapping from states to actions, Π: S →A that maximizes the sum of its reward over time.

**Elements of Reinforcement Learning:**

Except for the agent and the environment, we have four sub-elements of reinforcement learning system:

1. Policy: It defines the learning agent’s way of behaving at a given time.
2. Reward function: It defines the goal in reinforcement learning problem.
3. Value function: It specifies what is good in the long run.
4. Model of the environment (optional): Models are used for planning, by which we mean any way of deciding on a course of action by considering possible future situations before they are actually experienced.

Rewards are in a sense primary, whereas values, as predictions of rewards, are secondary. Without rewards, there could be no values, and the only purpose of estimating values is to achieve more reward.

**Challenges in reinforcement learning research:**

 (a) Scaling up to large problem.

 (b) Handling partially-observable Markov decision problems where the agent cannot sense the entire state of the environment.