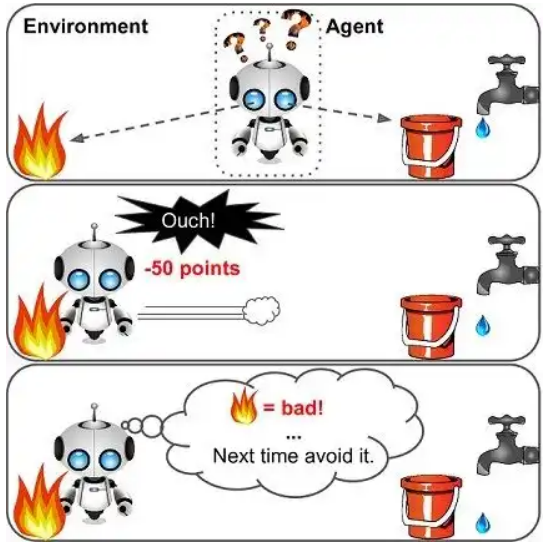
TECNOLÓGICO NACIONAL DE MÉXICO

1.3 Interacción Agente - Entorno

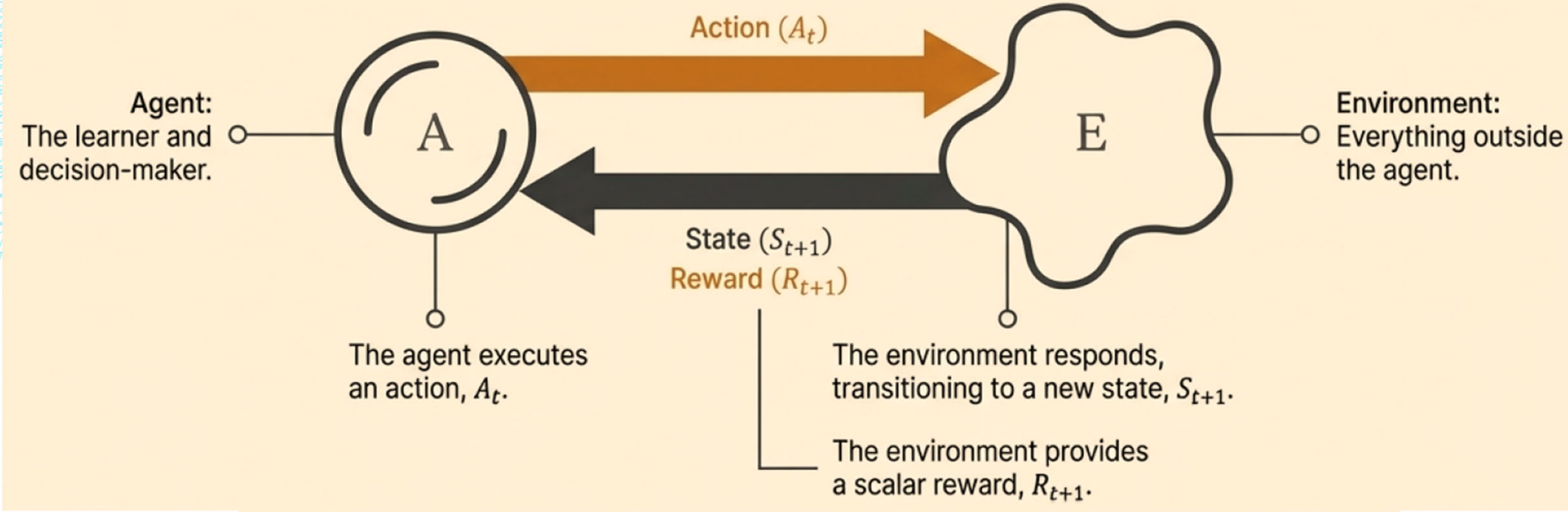
EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Reinforcement Learning is learning what to do — how to map situations to actions — so as to maximize a numerical reward signal.

Sutton Richard and Barto Andrew. Reinforcement Learning An Introduction



1.3 Interacción Agente - Entorno

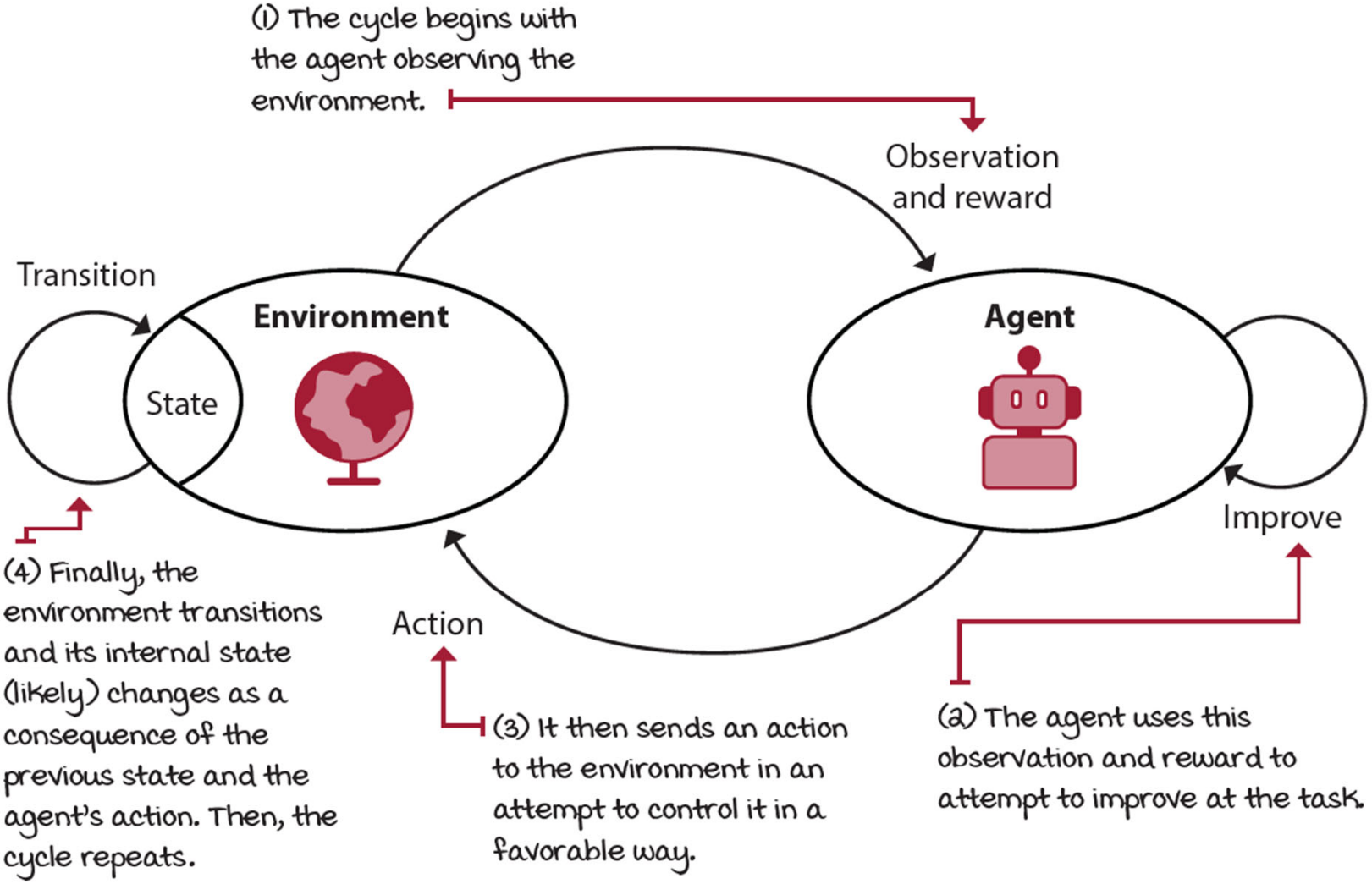


At each time step t:

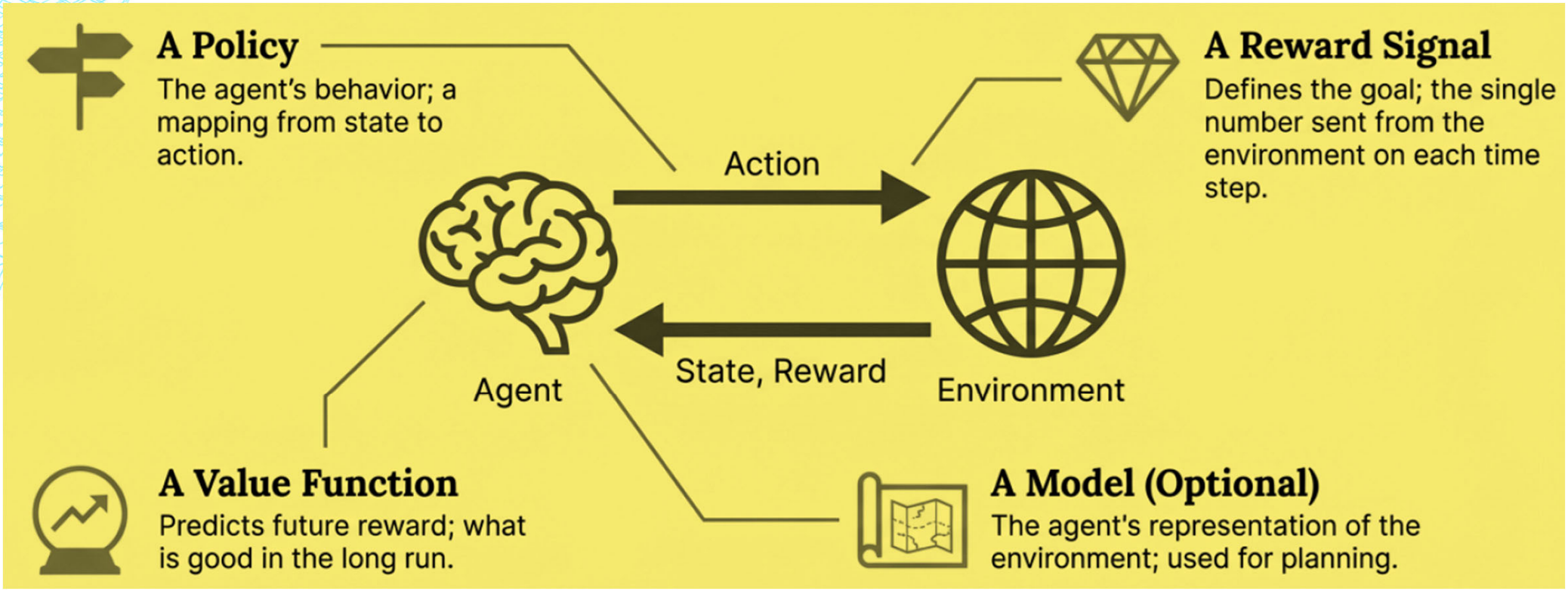
1. The agent executes an action, $\boldsymbol{A_t}$.
2. The environment responds, transitioning to a new state, $\boldsymbol{S_{t+1}}$.
3. The environment provides a scalar reward, $\boldsymbol{R_{t+1}}$.

This simple loop captures the essential features of the problem: cause and effect, uncertainty, and the existence of explicit goals.

1.3 Interacción Agente - Entorno



1.3 Interacción Agente - Entorno



TECNOLÓGICO NACIONAL DE MÉXICO

1.3 Interacción Agente - Entorno

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Policy (π)

A mapping from perceived states to actions. It defines the agent's behavior. Corresponds to stimulus-response rules.

Reward Signal (R)

A single number sent from the environment at each time step. It defines the goal; the agent's sole objective is to maximize the total reward over the long run.

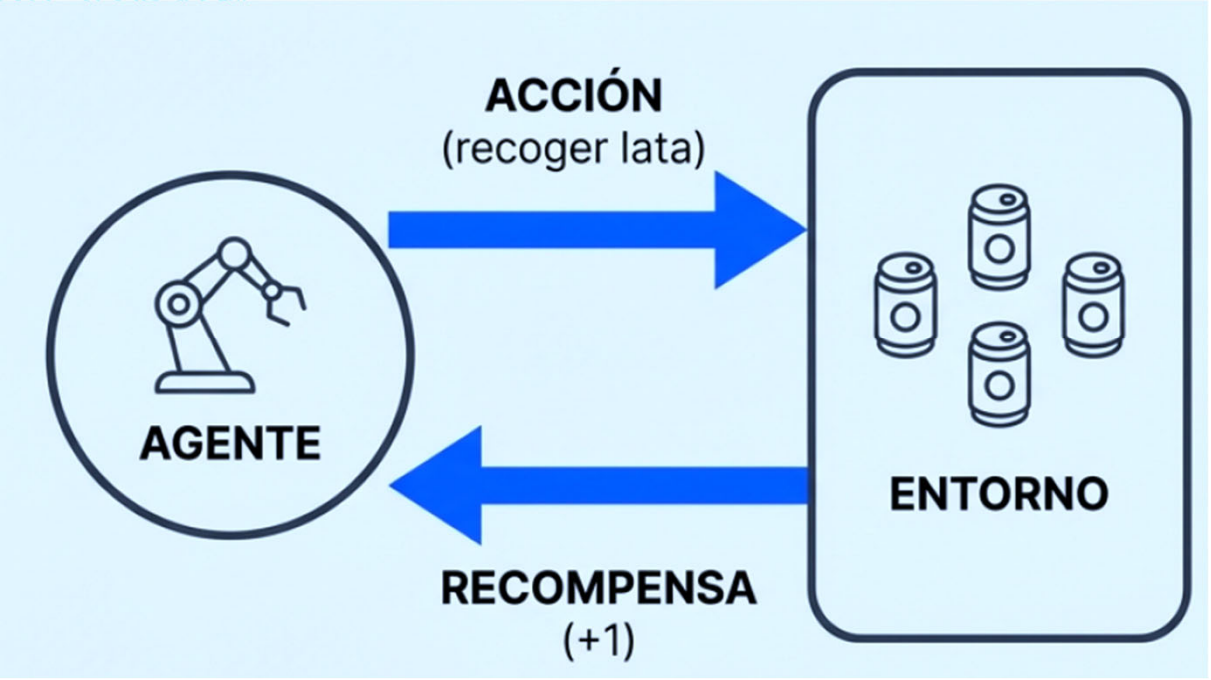
Value Function (V)

The total amount of reward an agent can expect to accumulate over the future, starting from a given state. It indicates the *long-term* desirability of states.

Model (Optional)

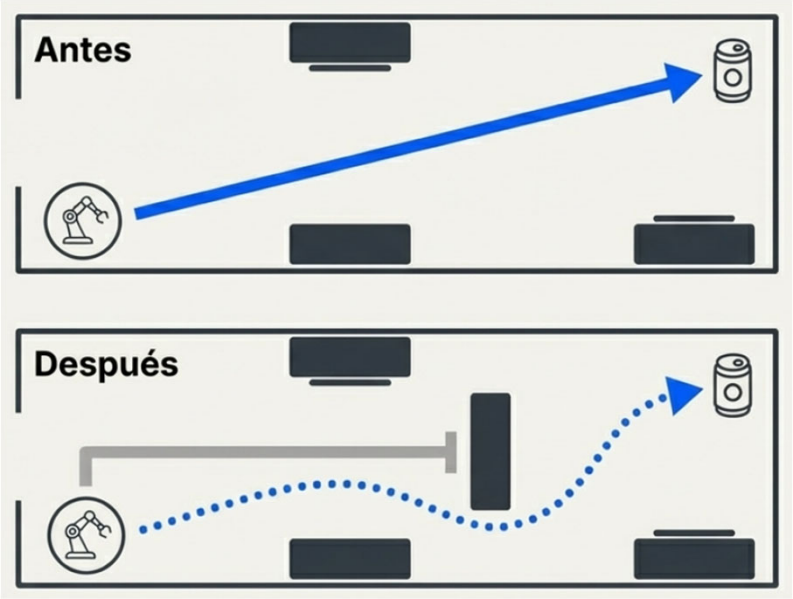
Mimics the behavior of the environment, predicting the next state and reward. Used for planning. Methods using a model are *model-based*; methods without one are *model-free*.

1.3 Interacción Agente - Entorno

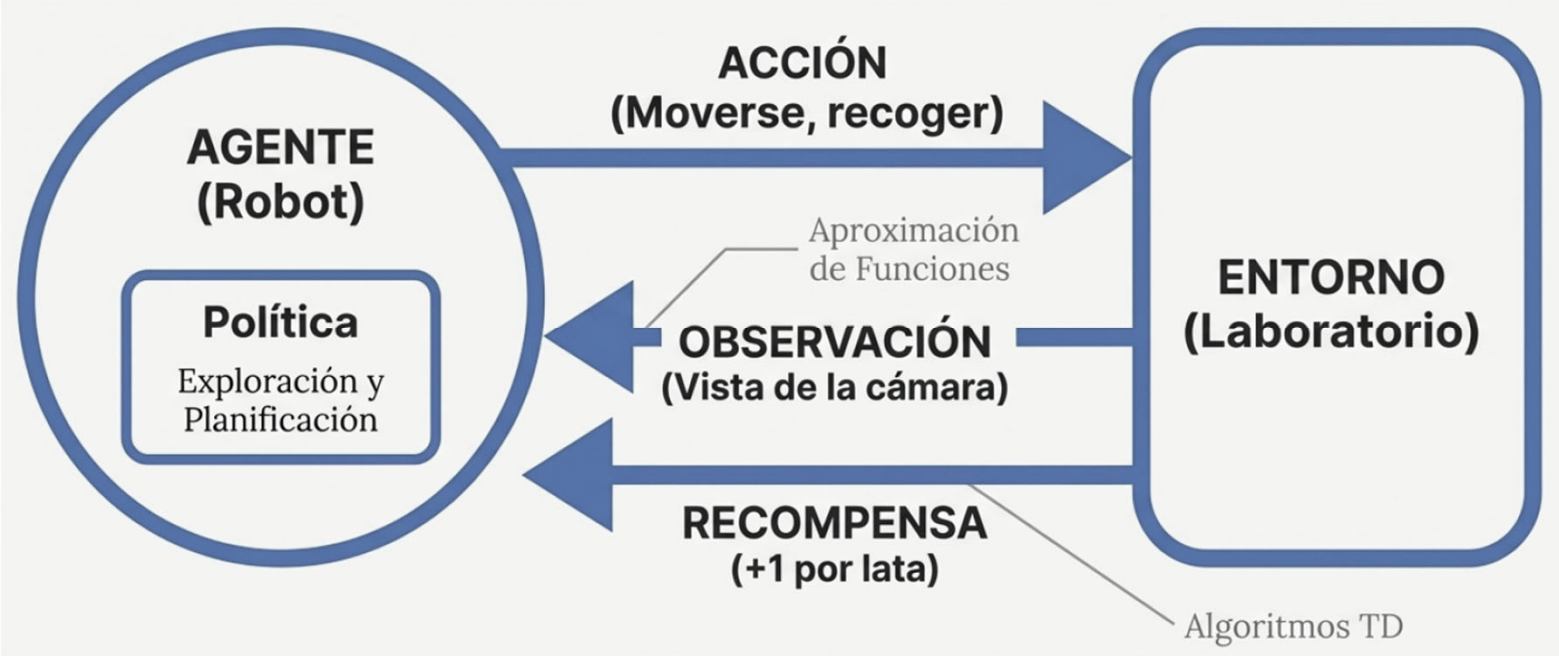


- **La Recompensa**: El número de latas que el robot recoge.
- **El Agente**: Aprende por sí mismo cómo maximizar esa recompensa.

1.3 Interacción Agente - Entorno



1.3 Interacción Agente - Entorno



TECNOLÓGICO NACIONAL DE MÉXICO

1.3 Interacción Agente - Entorno

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

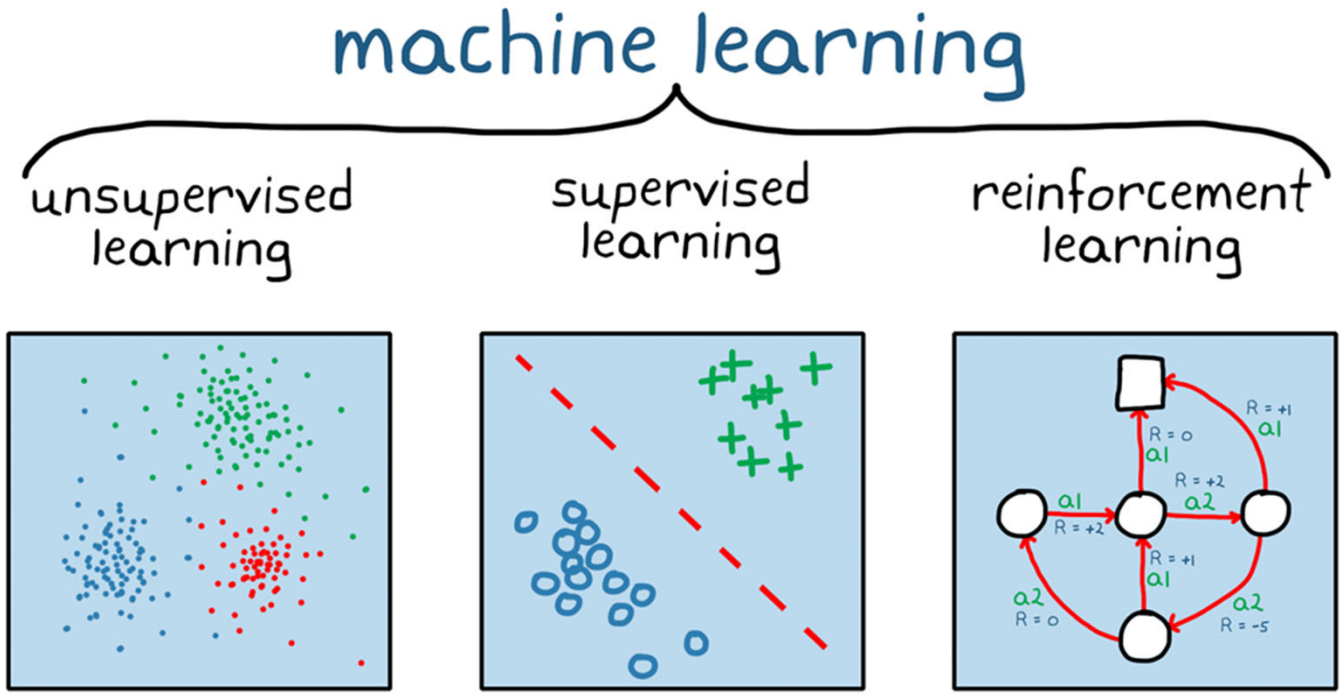
Interesting Sites

https://la.mathworks.com/products/reinforcement-learning.html

https://matlabacademy.mathworks.com/es/details/reinforcement-learning-onramp/reinforcementlearning

https://la.mathworks.com/discovery/reinforcement-learning.html

1.4 Un tercer paradigma del aprendizaje de máquinas



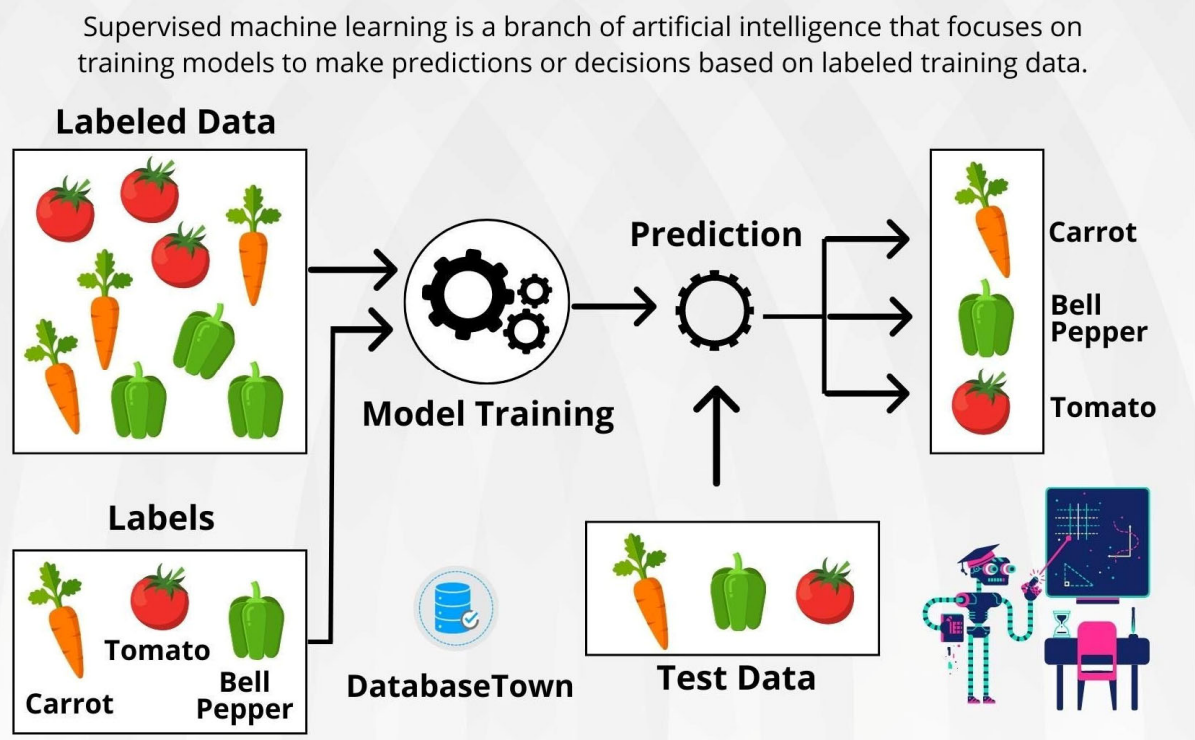
TECNOLÓGICO NACIONAL DE MÉXICO

1.4 Un tercer paradigma del aprendizaje de máquinas

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Supervised Learning (SL)

Goal: Learn a mapping from inputs to outputs. Generalize to new data.
Data: A training set of labeled examples provided by a supervisor.
Feedback: "Here is the situation, and here is the correct action/label."
Types of SL: SL tasks can be broadly divided into classification and regression problems.
Common Algorithms for Classification: Linear classifiers, SVM, Decision Trees, k-nearest neighbor (KNN), Logistic regression and Random Forest, Naïve Bayes and NN.
Common Algorithms for Regression: Linear Regression, Lasso Regression, Ridge Regression and Polynomial Regression.

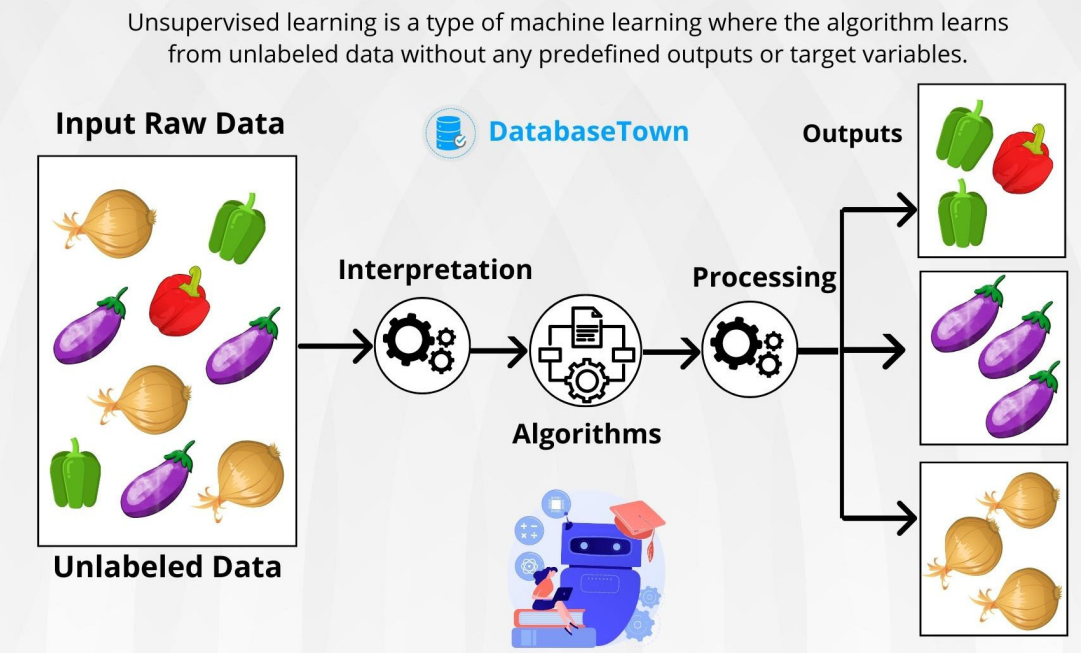


https://www.ibm.com/think/topics/supervised-learning

1.4 Un tercer paradigma del aprendizaje de máquinas

Unsupervised Learning (UL)

Goal: Find hidden structure in data.
Data: Collections of unlabeled data
Feedback: No explicit feedback; the goal is to find patterns inherent in the data.
Common unsupervised learning approaches: Clustering, Association, and Dimensionality reduction.
Common Algorithms: K-means clustering, Hierarchical clustering, Apriori, Principal Component Analysis (PCA), Autoencoders.



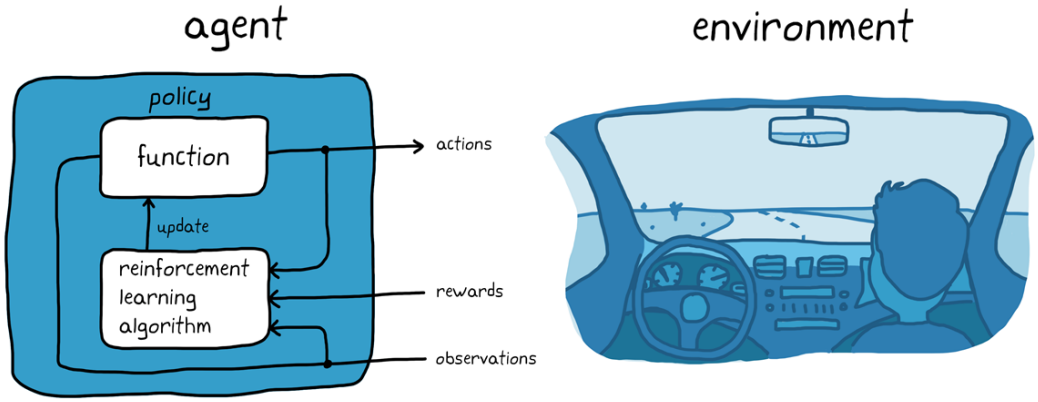
https://www.ibm.com/think/topics/unsupervised-learning#2014952965

1.4 Un tercer paradigma del aprendizaje de máquinas

Reinforcement Learning (RL)

Reinforcement learning is a machine learning technique in which a computer agent learns to perform a task through repeated trial-and-error interactions with a dynamic environment. This learning approach allows the agent to make a series of decisions that maximize a reward metric for the task completed, without human intervention and without being explicitly programmed to complete the task.

RL does not require a static dataset; instead, it operates in a dynamic environment and learns from collected experiences. Data points or experiences are collected during training through trial-and-error interactions between the environment and a software agent.



https://la.mathworks.com/discovery/reinforcement-learning.html

TECNOLÓGICO NACIONAL DE MÉXICO

EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

1.4 Un tercer paradigma del aprendizaje de máquinas

Evaluation U1