# El Aprendizaje por Refuerzo en el IoT

# Javier Corredor

# jacorredorca@sena.edu.co

# Centro Metalmecanico – Distrito Capital - SENA

Introducción: ¿Qué es?

La inteligencia artificial ha estudiado como dotar a la máquina de aptitudes cognitivas que los seres humanos poseen y que han permitido con éxito a través de la evolución permanecer como especie, teniendo que resolver problemas complejos que representan la vida o la muerte. Una de estas aptitudes es el aprendizaje. El aprendizaje se define como el proceso a través del cual se modifican y adquieren habilidades, destrezas, conocimientos, conductas y valores («Aprendizaje», 2020). Para el propósito de este documento el término agente hace referencia a ese dispositivo, máquina, aplicación de software o robot al cual se quiere dotar de la capacidad de aprendizaje.

Es claro que, aunque hay grandes avances en la inteligencia artificial, aun no se consigue transferir la capacidad de aprendizaje a una máquina tal como los humanos o animales disponen. Se ha logrado gracias al poder computacional que disponemos, manejar gran cantidad de datos que permiten emular ciertas aptitudes tales como clasificar objetos, encontrar patrones en estos datos (minería de datos) o controlar de forma autónoma carros de tal forma que pueden reaccionar y generalizar algunos de estos comportamientos a situaciones no esperadas (o programadas).

En la inteligencia artificial se han explorado tres estrategias de aprendizaje: el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo. El primero, **aprendizaje supervisado**, trata de mostrar datos previamente clasificados al agente para que este en un proceso llamado entrenamiento, aprenda el patrón de clasificación y luego al mostrar datos nuevos el agente pueda clasificar de forma correcta, de esta forma se puede observar una característica del aprendizaje como es la generalización de conceptos. Por otro lado, el **aprendizaje no supervisado** trata de que el agente, sin tener datos previamente clasificados trate de encontrar relaciones entre los datos, esto se conoce como *Clustering* (agrupamiento), y también representa una idea de lo que es el aprendizaje dado que a partir de unos hechos se trata de relacionarlos. Por último, está el paradigma de **aprendizaje por refuerzo** donde a partir de la interacción del agente con su entorno y a una función de desempeño (recompensa) el agente aprende estrategias que le permite resolver óptimamente el problema planteado.

¿Cómo Funciona?

El aprendizaje supervisado de forma general se implementa gracias al concepto de las redes neuronales artificiales, recientemente se disponen de algoritmos eficientes (Hinton, 2007) que permiten entrenar las redes de muchas capas permitiendo resolver la clasificación de patrones con muchas características de entradas (a.k.a Deep Learning). El aprendizaje no supervisado se implementa entre otros por medio del algoritmo k-means (JinHuaXu & HongLiu, 2010), el cual tiene diversas modificaciones que permiten clasificar patrones a partir de métricas de que tan cercanos están los datos en dimensiones que conforman hiper-espacios de características. Finalmente el aprendizaje por refuerzo hace el uso de la teoría de la probabilidad, los procesos de decisión de Markov e ideas tomadas de la programación dinámica (Sutton & Barto, 2018).

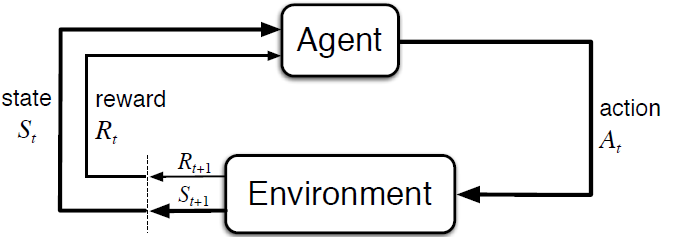


Figura 1 Interacción entorno-agente en el aprendizaje por refuerzo (Sutton & Barto, 2018)

El agente interactúa con el entorno (ver Figura 1) a través de tomar la decisión de realizar uno y otra acción (). Por su parte el entorno se transforma a un estado () y a su vez se tiene una métrica de desempeño del agente con relación a este estado conocida como recompensa (). El objetivo es que el agente maximice la recompensa. Dependiendo del problema, su dinámica, puede ser que para maximizar la recompensa se deban escoger acciones que produzcan bajas recompensas para poder al final obtener una recompensa máxima. Dado que explorar cada posible acción puede ser imposible porque el espacio de búsqueda sería inmanejable, se ha propuesto el modelo de procesos de decisión de Markov y soluciones iterativas considerando la programación dinámica para encontrar la estrategia optima (política) que debe asumir el agente para maximizar la recompensa.

El aprendizaje por refuerzo se presenta como el nuevo paradigma que combinado con las otras estrategias de aprendizaje puede resolver problemas complejos, de hecho, la combinación de redes neuronales y conceptos del aprendizaje por refuerzo se encuentran en el denominado Deep Reinforcement Learning.

Aplicación y ejemplos de herramientas

Con relación a las herramientas disponibles para implementar este tipo de algoritmos y su potencial de uso en la nube ver Tabla 1.

Tabla 1 Relación de utilidades de Reinforcement Learning

|  |  |  |  |
| --- | --- | --- | --- |
| **Utilidad** | **Open Source?** | **Link - Consulta** | **Observación** |
| AWS DeepRacer | NO | https://aws.amazon.com/es/deepracer/ | Simulador de carreras 3D basado en la nube, un coche de carreras totalmente autónomo en una escala de 1/18, impulsado por el aprendizaje por refuerzo, y una liga de carreras global. |
| AWS - SageMaker Reinforcement Learning | NO | https://docs.aws.amazon.com/sagemaker/latest/dg/reinforcement-learning.html | Admite TensorFlow y Apache MXNet.  Un kit de herramientas de RL que gestiona la interacción entre el agente y el entorno y proporciona una amplia selección de algoritmos de RL, Intel [Coach](https://nervanasystems.github.io/coach/) y Ray [RLlib](https://docs.ray.io/en/latest/rllib.html). |
| Azure - Personalizer | NO | https://azure.microsoft.com/en-us/services/cognitive-services/personalizer/ | Brinda a los usuarios experiencias relevantes que mejoran con el tiempo, en función de su comportamiento. A diferencia de los motores de recomendación que ofrecen algunas opciones de un gran catálogo, Personalizer presenta el mejor resultado para un usuario, cada vez que interactúa con su aplicación. |
| Azure - Machine Learning | NO | https://techcommunity.microsoft.com/t5/azure-ai/introducing-reinforcement-learning-on-azure-machine-learning/ba-p/1403028 | Escale el aprendizaje por refuerzo a poderosos clústeres de cómputo, soporte escenarios de múltiples agentes, acceda a algoritmos, marcos y entornos de RL de código abierto |
| DeepMind Lab | SI | https://deepmind.com/research/publications/deepmind-lab | DeepMind Lab es una plataforma similar a un juego personalizable en 3D diseñada para la investigación de IA basada en agentes. Se observa desde un punto de vista en primera persona, a través de los ojos del agente simulado. |
| Gym – openAI | SI | https://gym.openai.com/ | Gym es un conjunto de herramientas para desarrollar y comparar algoritmos de aprendizaje por refuerzo. |

El potencial de este tipo de algoritmos en el IoT es inmenso dado que se implementa el concepto de mejorar el desempeño a partir de la interacción con el entorno, por ejemplo, como se plantea en la herramienta de Azure, Personalizer. Allí se promete mejorar la experiencia del usuario considerando el mejor resultado de recomendación de producto para un usuario basado en cada vez que el usuario interactúa con la aplicación. Este enfoque también es importante en control de procesos, dado que el modelo que se requiere construir en algunas aplicaciones se puede mejorar al interactuar el agente con los datos medidos. Especial énfasis en el área de robótica dado que un robot, así como cualquier organismo viviente debería tener la capacidad de adaptarse, tomar decisiones que a la larga optimicen su comportamiento.

Conclusión

En el diplomado se revisaron los paradigmas de aprendizaje supervisado y no supervisado, en este documento se extienden estos algoritmos mediante un resumen del concepto aprendizaje por refuerzo. Al realizar una búsqueda de las herramientas disponibles se evidencia que a pesar de que es una aplicación de conceptos bien estudiados en la teoría de la probabilidad, recién se dispone de aplicaciones comerciales como en las nubes de Amazon (AWS) y Azure, además hay desarrollo en aplicaciones de software libre que permiten aplicar estos conceptos desde un enfoque académico.

# Referencias

Aprendizaje. (2020). En *Wikipedia, la enciclopedia libre*. https://es.wikipedia.org/w/index.php?title=Aprendizaje&oldid=128804350

Hinton, G. E. (2007). Learning multiple layers of representation. *Trends in Cognitive Sciences*, *11*(10), 428-434. https://doi.org/10.1016/j.tics.2007.09.004

JinHuaXu, & HongLiu. (2010). Web user clustering analysis based on KMeans algorithm. *2010 International Conference on Information, Networking and Automation (ICINA)*, *2*, V2-6-V2-9. https://doi.org/10.1109/ICINA.2010.5636772

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. 2nd Edition, 352pg.

Corredor, Javier <https://github.com/jcorredorc/iot_eafit.git>, 2020