



Universidad
Tecmilenio®





Inteligencia artificial a través del aprendizaje profundo

Aprendizaje profundo en producción

Aprendizaje por refuerzo profundo



Los modelos de aprendizaje automático (ML, por sus siglas en inglés) han proporcionado diferentes habilidades y capacidades a los sistemas de cómputo, y aprender de manera automática sin la necesidad de programar es una de ellas.

En la medida que se implementan modelos de aprendizaje en distintos entornos, toma importancia desarrollar metodologías que permitan describir las etapas y su duración para poder compartirse a los involucrados en un proyecto de esta naturaleza. La manera en la que funciona un sistema de aprendizaje se puede describir a partir de su ciclo de vida, cuyo objetivo principal es encontrar una solución a un problema o proyecto.

Por otro lado, hay una rama del aprendizaje automático: el **aprendizaje por refuerzo profundo**, y es, en esencia, aprendizaje a partir de interacciones y retroalimentaciones. Combina redes neuronales con aprendizaje por refuerzo para impulsar a los sistemas a cumplir sus objetivos.

En el año 2012, DeepMind mostró al mundo una red neuronal profunda que aprendió a jugar Atari con gran habilidad (Puigdomènech et al., 2020). Esta red, llamada Deep Q-Network (DQN), fue la primera aplicación exitosa a gran escala del aprendizaje por refuerzo, utilizando redes neuronales profundas.





Implantación de los ciclos de vida de los modelos de aprendizaje

El ciclo de vida de un modelo de aprendizaje profundo (AP) es un marco de referencia que captura el proceso iterativo de desarrollo de la solución de aprendizaje profundo para un problema en específico. No es lineal, ya que a medida que se avanza a lo largo de las etapas, se puede descubrir más información que provoque regresar a etapas previas y realizar cambios.

El aprendizaje automático (y por lo tanto el aprendizaje profundo), está fuertemente relacionado con datos, tanto, que no existe aprendizaje automático si no se cuenta con una cantidad de datos considerable de donde aprender. Hoy día, la cantidad de datos disponible crece de forma exponencial, lo que facilita de cierta manera el proceso de aprendizaje. Su conexión los algoritmos de cómputo se hace por medio de elementos matemáticos que enseñan a las computadoras a aprender. **Sin matemáticas no habría aprendizaje automático ni profundo, por tanto, no habría inteligencia artificial.**





En el caso del ciclo de vida del aprendizaje automático, los elementos son los siguientes:



Entendimiento del problema

- En esta etapa se incluye la definición y el alcance del proyecto.

Datos

- Definir los datos a utilizar y establecer un punto de referencia.
- Etiquetar y organizar los datos.

Modelado

- Seleccionar y entrenar un modelo.
- Realizar el análisis de los resultados.

Despliegue del modelo

- Despliegue en producción.
- Monitoreo y mantenimiento del sistema.



A diferencia de ML, en el aprendizaje profundo los atributos de los datos no son dados por humanos, sino que se aprenden de manera automática con los datos de entrada; más aun, los atributos son complejos y tienen jerarquía a lo largo de la red neuronal, por lo que se requiere de menos conocimiento del dominio y del problema. Por otro lado, este tipo de modelos son difíciles de explicar, de tal manera que el proceso de mejora (sobre todo el cambio de estructura y sus parámetros) es una tarea de búsqueda empírica.

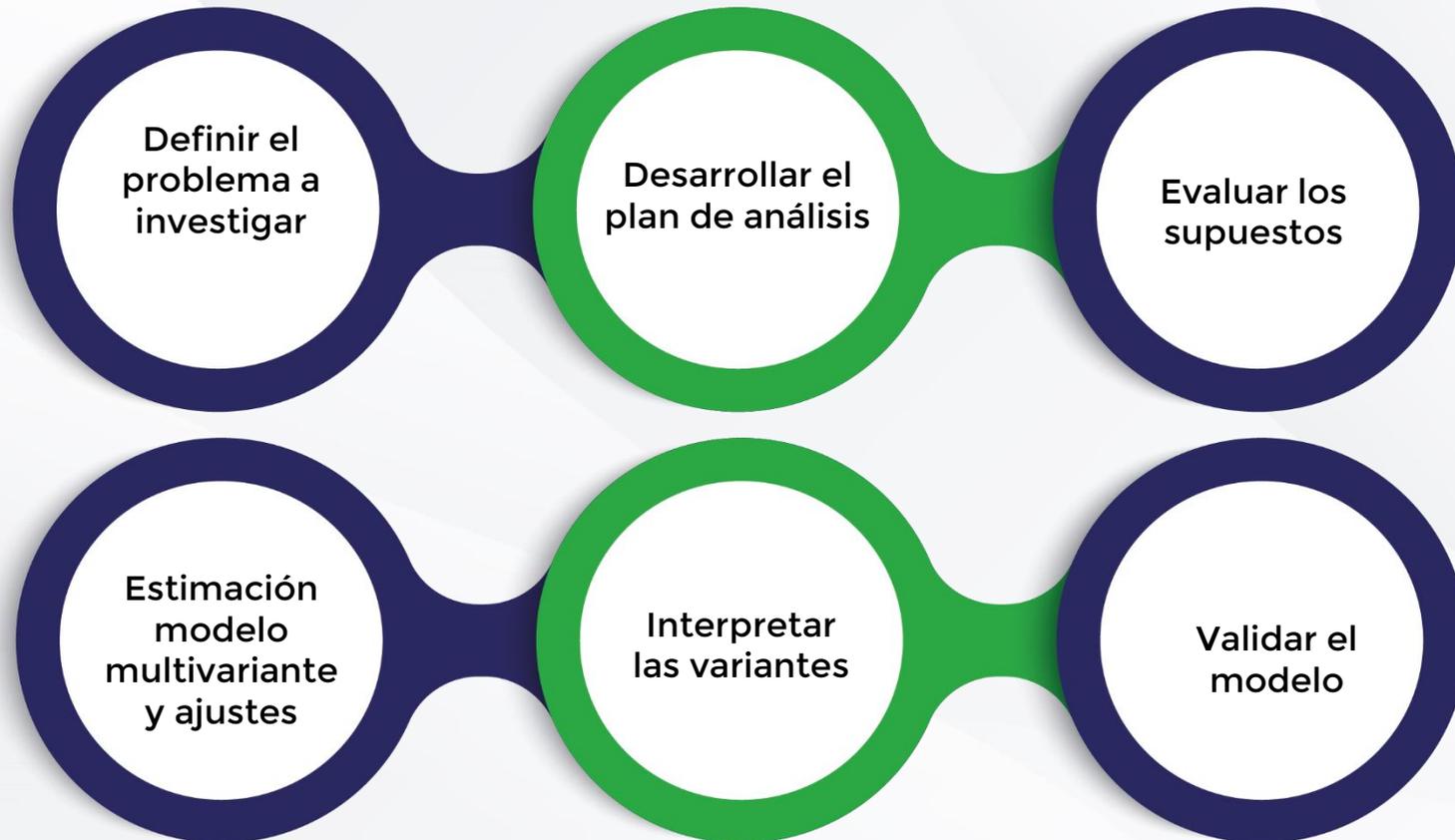


En la figura anterior, se muestra el ciclo de vida típico de un modelo de aprendizaje profundo que puede separarse en dos grandes etapas: la de experimentación y la de despliegue. Una vez que se define la tarea de predicción, la etapa de experimentación inicia utilizando modelos conocidos y que funcionan bien en dominios con tareas similares, se especifican los datos de entrenamiento y la función de pérdida (costo o error) en la salida para que con esta información, de forma repetida, se ajusten las operaciones y conexiones, así como los parámetros de la red neuronal profunda, se entrena y evalúa el desempeño del modelo y así sucesivamente, hasta que no exista mejora en la precisión del resultado. En la fase de despliegue, todos los recursos se enfocan en poner al modelo en un ambiente de producción.





Gestión de los ciclos de vida de los modelos de aprendizaje profundo



Existen distintos esfuerzos para abordar los diferentes aspectos en la gestión del ciclo de vida de un modelo de aprendizaje profundo, en la academia, por ejemplo, ModelDB (Vartak et al., 2016), ModelHub (Miao et al., 2017) y ProvDB (Miao y Deshpande, 2018).





Aprendizaje de funciones

El aprendizaje automático (AP) proporciona métodos automatizados que pueden detectar patrones en los datos, con el objetivo de resolver alguna tarea específica. Dependiendo de la forma en la que resuelve los problemas, se puede clasificar en tres tipos:

Aprendizaje supervisado (AS): utilizado en tareas de reconocimiento de datos no estructurados y etiquetados, como clasificación o regresión.

Aprendizaje no supervisado (ANS): este tipo de aprendizaje se utiliza para inferir distribuciones, reducción de la dimensionalidad, agrupamiento e inferencia de variables latentes de datos no etiquetados.

Aprendizaje por refuerzo (AR): usado en problemas de toma de decisiones, donde un agente interactúa con un entorno, percibiendo estados y tomando acciones de acuerdo con una política que le permita maximizar una recompensa.





El **AR** es un tipo de aprendizaje dirigido por objetivos a partir de la interacción con el entorno. A la entidad que aprende no se le indica qué acciones tomar, sino que debe descubrir por sí misma cuáles son las acciones que producen la mayor recompensa, su objetivo, a prueba y error; tales acciones afectan no solo a su recompensa inmediata, sino a las próximas, dado que las acciones actuales pueden determinar situaciones futuras.

Incluye diferentes conceptos que deben ser comprendidos claramente (Graesser y Keng, 2019):

Agente o actor: es una entidad que ejecuta acciones. El algoritmo es el agente.

Acción: es un "movimiento" realizado por el agente.

Entorno: es el mundo donde el agente actúa y que reacciona a tales acciones en forma de estados y recompensas.

Estado: es una situación concreta en la que el agente se encuentra, un lugar específico en un instante específico.

Recompensa: es la forma de medir el éxito o el fracaso de las acciones tomadas por el agente en un estado determinado. Pueden ser inmediatas o futuras.

Política: es la estrategia utilizada por el agente para determinar la siguiente acción a tomar con base en su estado actual. Permite un mapeo entre estados y acciones (las que den mayor recompensa).

Valor: es el rendimiento esperado a largo plazo del estado actual bajo cierta política.

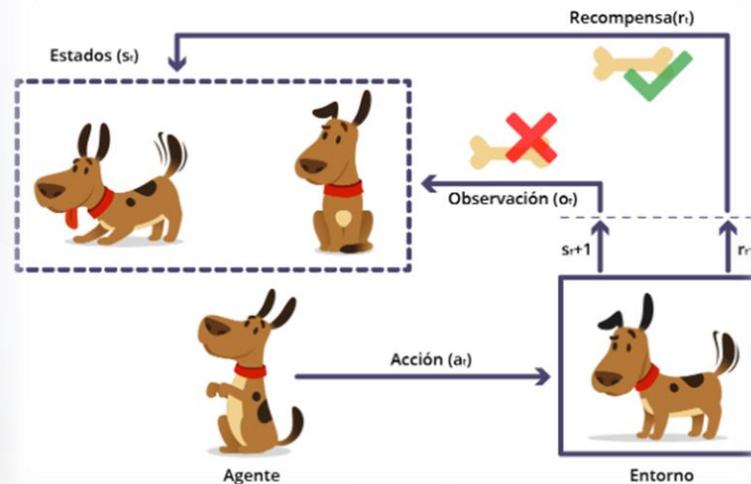
Valor Q: es el rendimiento esperado a largo plazo de una acción realizada en un estado actual bajo cierta política.

A diferencia de los otros tipos de aprendizaje, el AR no tiene supervisores, sino recompensas; la retroalimentación no es instantánea, el componente temporal es importante, ya que las acciones son secuenciales y las acciones realizadas por el agente pueden afectar la información subsecuente que reciba.





En la siguiente figura se muestra la analogía del aprendizaje por refuerzo con el escenario de entrenar a una mascota para realizar cierto truco. En este ejemplo, al perro no se le dice lo que debe hacer, solo se le premia si hace lo correcto o se le castiga en caso contrario. En cada instante (o paso) tiene que recordar qué fue lo que hizo para recibir una recompensa o castigo (problema de asignación de crédito).



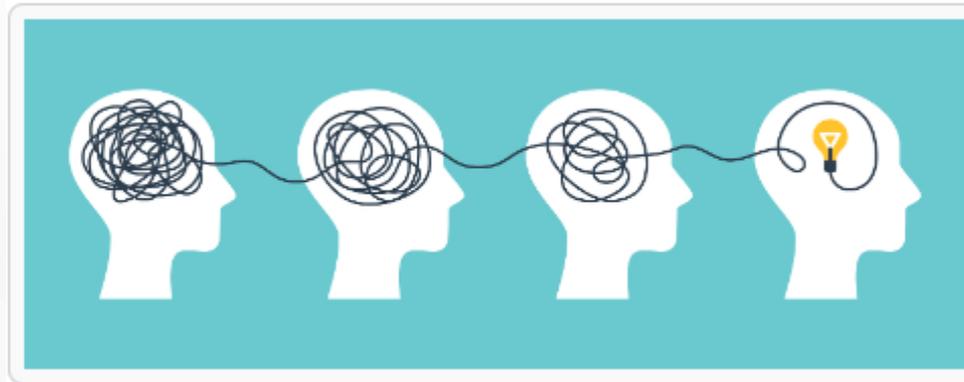
El problema general del AR se formaliza como un proceso de control estocástico de tiempo discreto, donde un agente interactúa con su entorno: a partir de una observación inicial O_0 , el agente comienza en un estado específico dentro de su entorno s_0 . En cada instante de tiempo t , el agente debe tomar una acción que tiene tres consecuencias: obtiene una recompensa r_t , hace una observación O_{t+1} y sufre una transición de estado a s_{t+1} .





Proceso de decisión de Markov

El proceso de decisión de Markov (PDM) define un marco de referencia matemático para resolver un problema de AR (François-Lavet et al., 2018) y requiere de la propiedad de Markov, que establece que el futuro depende solo del presente y no del pasado.



El proceso de Markov consiste en una secuencia de estados que obedecen a la propiedad de Markov. Cuando un problema de AR satisface la propiedad de Markov, es decir, el futuro depende solo del estado actual y de la acción tomada y no del pasado, entonces el problema se formula como un proceso de decisión de Markov.

Un proceso de decisión de Markov se define a partir de una tupla de cinco elementos (S, A, R, p, γ):

Un conjunto finito de todos los estados posibles, S .

Un conjunto finito de acciones, A .

Una función de recompensa, R .

Una función de transición, p .

Un factor de descuento, γ .





Responde de manera clara:

¿Cuál es la diferencia entre el ciclo de vida del aprendizaje automático y el ciclo de vida de un modelo de aprendizaje profundo?

¿Qué implica la gestión de los ciclos de vida de los modelos de aprendizaje profundo?

¿Cuáles son las diferencias entre los otros tipos de aprendizaje automático y el aprendizaje por refuerzo?

Identifica y menciona los elementos de un proceso de decisión de Markov



1216658919





Sin las herramientas y los procedimientos adecuados que permitan abordar los desafíos de la gestión del ciclo de vida de un modelo de AP, las personas involucradas dedican mucho tiempo y esfuerzo para darle seguimiento a sus experimentos, registrar los valores de los parámetros utilizados, llevar un registro de versiones de los modelos y la visualización de los resultados, por citar algunas tareas adicionales al modelado.

Por otro lado . . .

El AR es un proceso de aprendizaje a prueba y error basado en recompensas y es esencial para construir agentes que no solo puedan percibir e interpretar el mundo, sino también actuar e interactuar con él.





- François-Lavet, V., Henderson, P., Islam, R., Pineau, B., y Pineau, J. (2018). An Introduction to Deep Reinforcement Learning. *Foundations and Trends in Machine Learning*, 11(3-4). Recuperado de <https://arxiv.org/pdf/1811.12560.pdf>
- Graesser, L., y Keng, W. (2019). *Foundations of Deep Reinforcement Learning: Theory and Practice in Python*. Estados Unidos: Addison-Wesley Professional.
- Puigdomènech, A., Piot, B., Kapturowski, S., Sprechmann, P., Vitvitskyi, A., Guo, D., y Blundell, C. (2020). *Agent57: Outperforming the Human Atari Benchmark*. Recuperado de <https://deepmind.com/blog/article/Agent57-Outperforming-the-human-Atari-benchmark>

