

Ahora ya conoces la idea general del machine learning tanto supervisado como no supervisado y la diferencia que hay con el deep learning.

Te lo recuerdo:

El machine learning tiene el foco en la predicción y solucionar el problema en base a crear o entrenar herramientas matemáticas.

Existen dos tipos de machine learning:

- Supervisado: cuando tenemos información de la salida.
- No supervisado: no tenemos información de la salida.

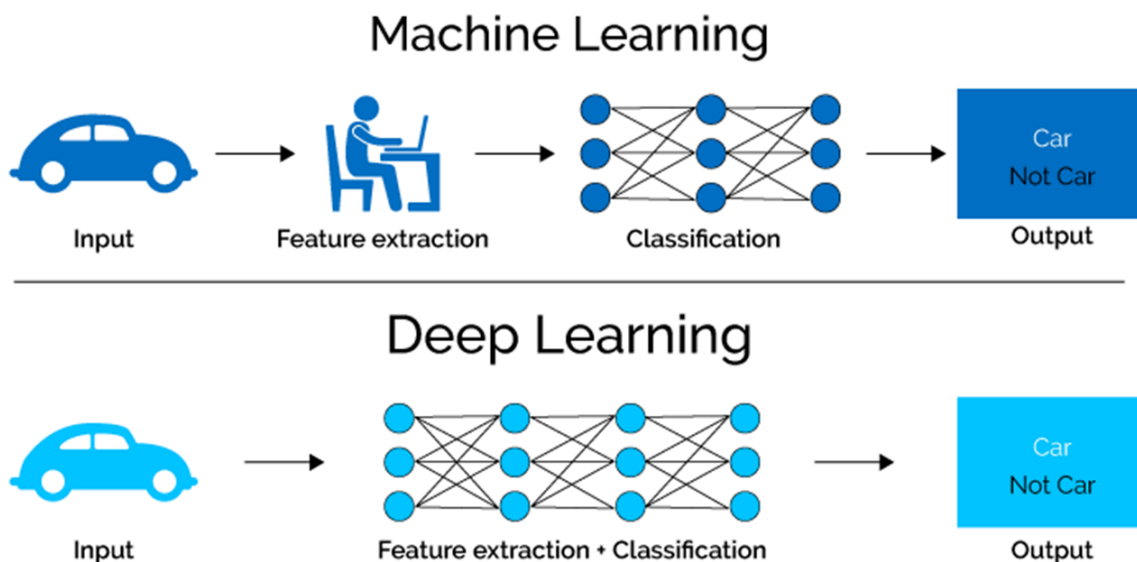
Y podemos solucionar o atacar problemas de:

- Regresión
- Clasificación
- Clustering
- Reducción dimensional
- Reglas de asociación

Ahora bien, en machine learning siempre tenemos que preparar las características o features de entrada del problema.

En cambio, en deep learning es la propia red neuronal convolucional quien se encarga de trabajar de extraer las características.

¡Bien!



En esta hoja de trabajo vamos a trabajar los conceptos que hemos planteado durante la lección para ayudarte a asentar las ideas.

Para ello te planteo lo siguiente:

Te voy a proponer que mires los 7 datasets siguientes:

1. Dataset Fruit: <https://www.kaggle.com/datasets/muratkokludataset/date-fruit-datasets>
2. Dataset Fraud detection: <https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset>
3. Dataset Créditos: <https://www.kaggle.com/datasets/luisenriquesguerrero/creditos-personales-actualizado>
4. Dataset Alcohol Consumption: <https://www.kaggle.com/datasets/pralabhoudel/alcohol-consumption-by-country>
5. Dataset Satisfaction: <https://www.kaggle.com/datasets/teeimahal20/airline-passenger-satisfaction>
6. Dataset Vaccine Articles: <https://www.kaggle.com/datasets/brandonconrady/covid19-vaccine-articles>
7. Dataset Pistacho: <https://www.kaggle.com/datasets/muratkokludataset/pistachio-dataset>

De cada dataset quiero que pienses qué problema de machine learning o deep learning es. Recuerda qué los problemas pueden ser: de clasificación, de regresión, de clustering, de reglas de asociación, de reducción dimensional o de reglas de asociación.

Además dime si lo solucionarías con machine learning o deep learning y qué variable de salida propones para cada uno de los datasets solo en caso de regresión o clasificación.

Dataset	Tipo de problema: clasificación , regresión , clustering, ...	¿Aplicaría machine learning o deep learning?	Variable de salida (sólo si es de regresión o clasificación)
1 Fruit	Clasificación	Ambas • Características : machine learning • Imágenes: deep learning	Tipo de dátil (Barhee, Deglet Nour, Sukkary, Rotab Mozafati, Ruthana, Safawi, and Sagai)
2 Fraud detection	Clasificación	Machine learning	Si es fraude o no
3 Créditos	Clasificación	Machine learning	Si contratará un crédito personal o no
4 Alcohol Consumption	Regresión	Machine learning	Consumo de alcohol por país según su PIB
5 Satisfaction	Clasificación	Machine learning	Satisfacción (satisfecho/a, neutral, insatisfecho/a)
6 Vaccine Articles	NLP	Deep learning	
7 Pistacho	Clasificación	Deep learning	Si es un tipo de pistacho u otro.

Pista: intenta pensar como un problema de entradas - salida.
Te será mucho más fácil entender el problema :)

¡Genial!

Estás empezando a pensar como un científico de datos y definir la estrategia de proyectos de machine learning o deep learning.

¡Celébralo como te guste! Vas por buen camino

1 abrazo y hasta pronto.
Brindo por tus éxitos :)

Jordi Ollé