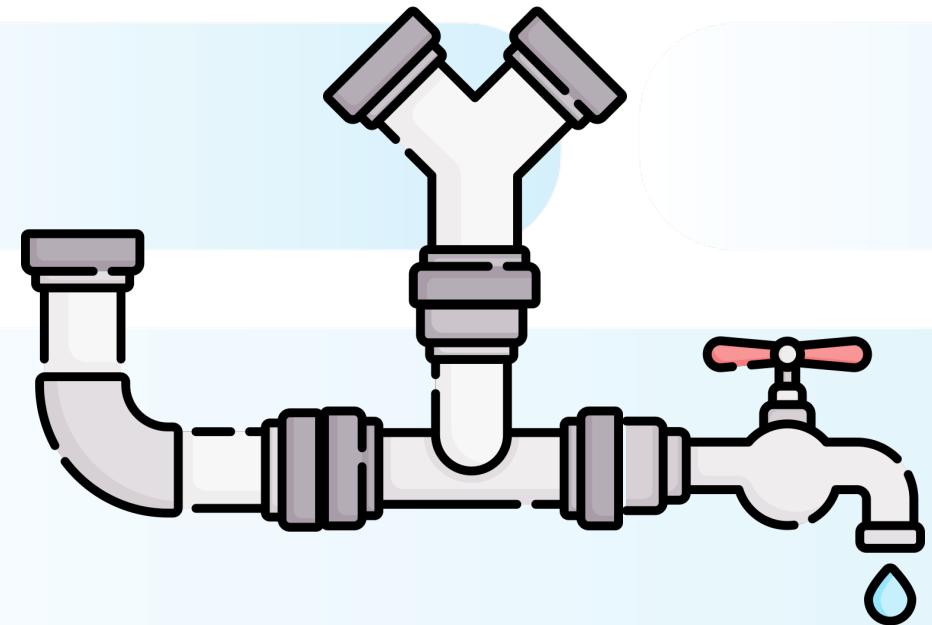


Machine Learning Pipelines



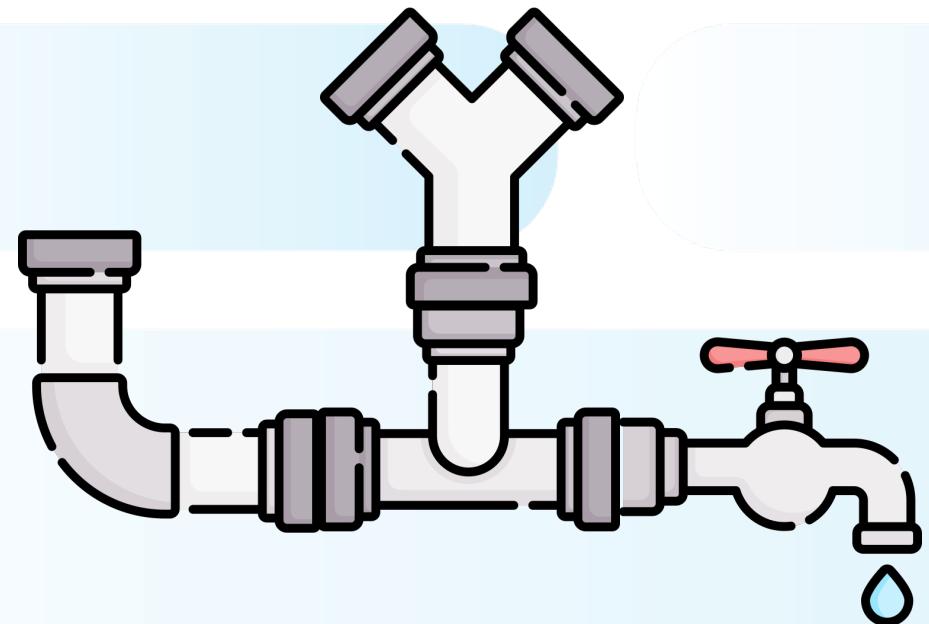
¿Qué es un *pipeline*?

- Flujo de información de un punto a otro
- Transformaciones en el proceso
- Dividido en pasos que comparten información entre ellos



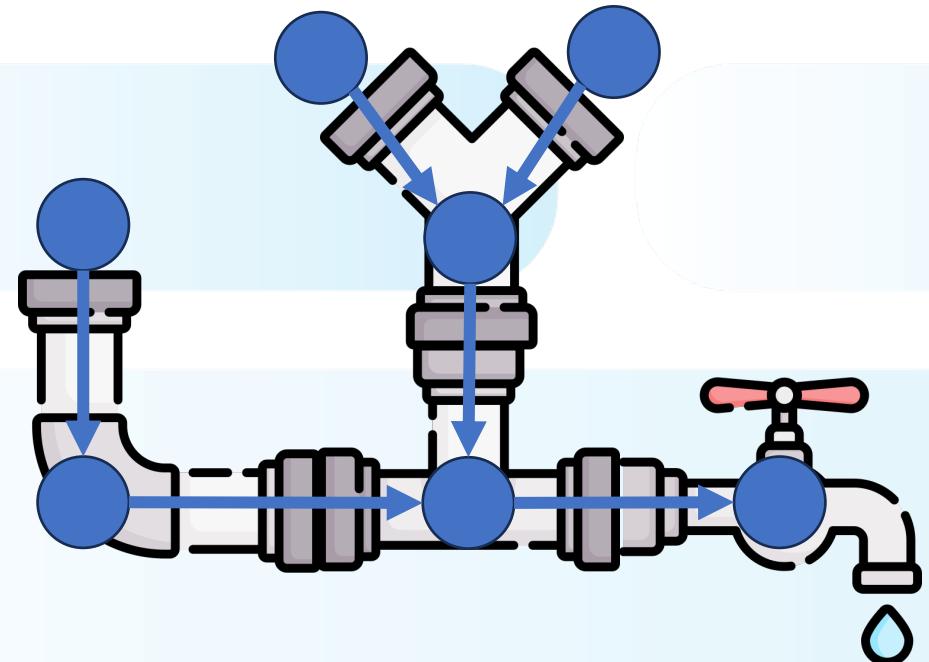
¿Qué es un *DAG*?

- ▶ Grafo Dirigido Acíclico (Directed Acyclic Graph)
- ▶ Siguen una secuencia sin ciclos

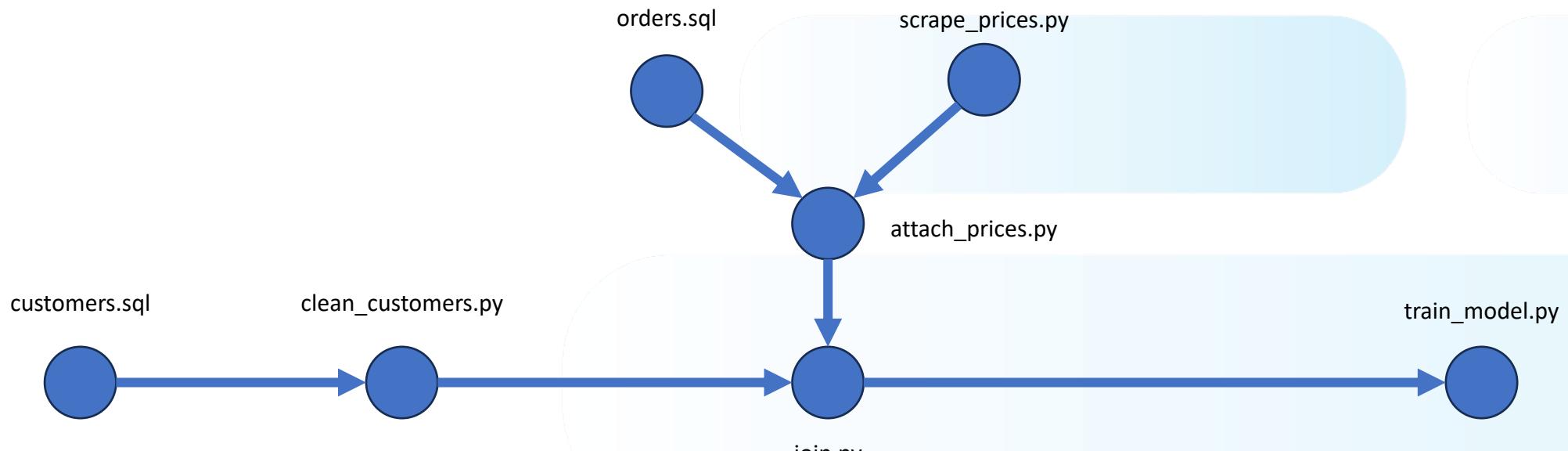


¿Qué es un *DAG*?

- ▶ Grafo Dirigido Acíclico (Directed Acyclic Graph)
- ▶ Siguen una secuencia sin ciclos



¿Qué es un *DAG*?

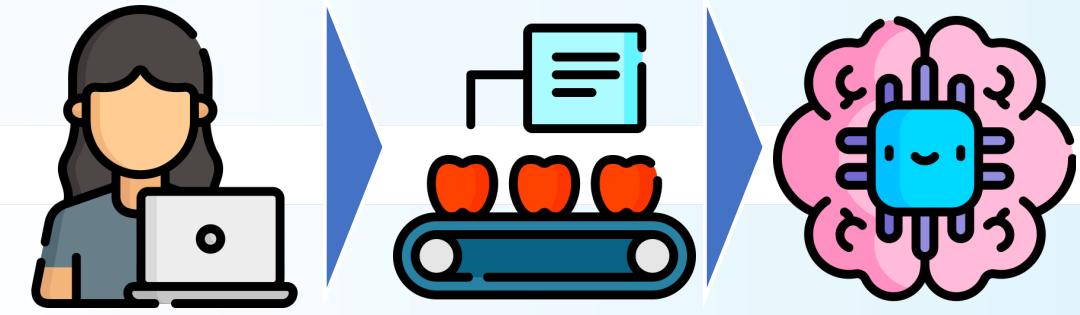
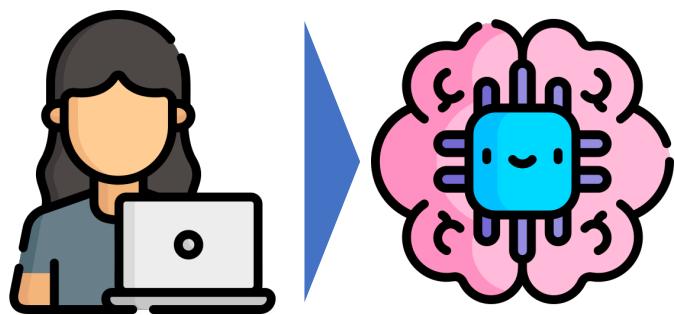


El equipo de machine learning como fábrica de modelos



No crees modelos...

crea código que cree modelos



Partes básicas de un pipeline



Ejemplo básico de un pipeline, no es necesariamente una estructura que tengas que seguir forzosamente



Partes básicas de un pipeline



Solamente leer los datos de las fuentes necesarias – sin transformaciones significativas



Partes básicas de un pipeline



Validación de la información recién consumida, verificando que los datos tengan el esquema y distribuciones esperadas



Partes básicas de un pipeline



Pre-procesamiento para convertir nuestros datos en algo que el modelo de ML pueda consumir para ser entrenado



Partes básicas de un pipeline



Entrenamiento del modelo usando los datos que acabamos de transformar



Partes básicas de un pipeline



Validación del desempeño y métricas del modelo, revisión de sesgos y aseguramiento de calidad



Partes básicas de un pipeline



Monitoreo de datos de entrada, salidas y en general del desempeño del modelo.



Conclusión



Un pipeline es parte del ciclo de vida de un modelo de ML, pero no es todo el ciclo de vida

