

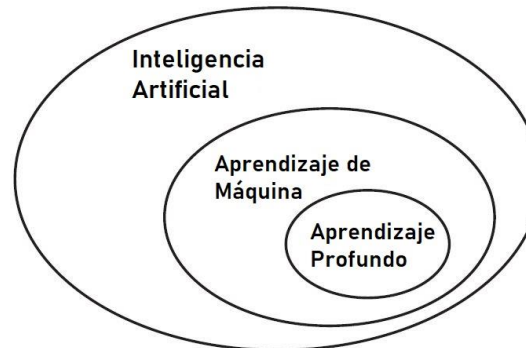
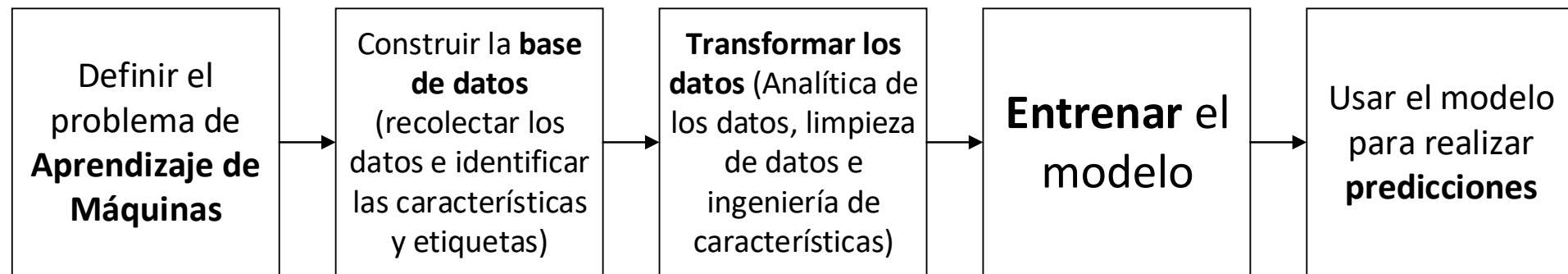


Aprendizaje de Máquinas

CURSO GRUPO BANCOLOMBIA

Universidad Nacional de Colombia

El Proceso de Aprendizaje de Máquina

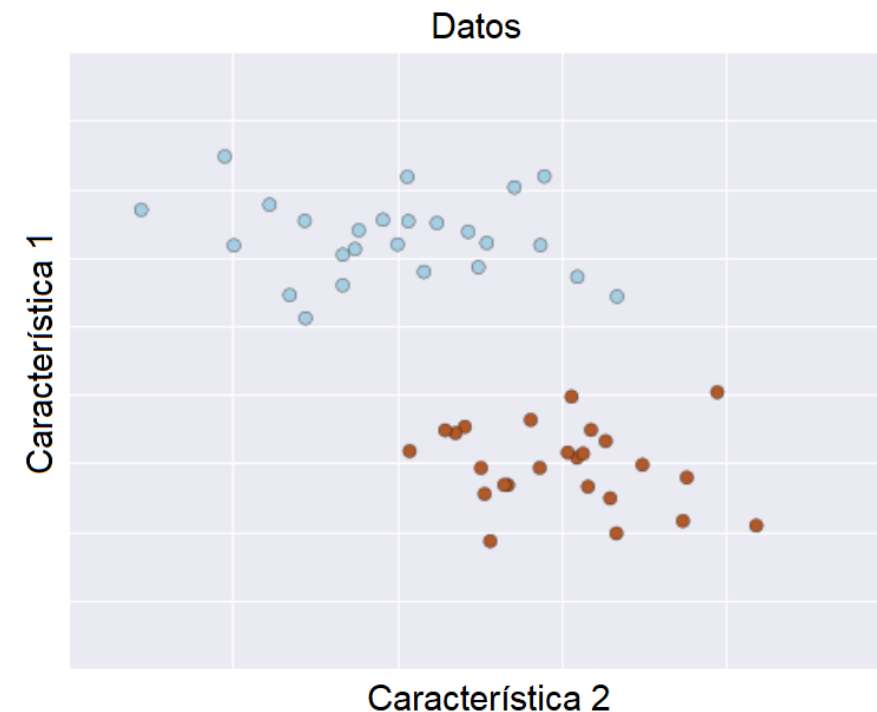


Tomado de: Deep Learning with Python, Francois Chollet, 2018

Aprendizaje de Máquina: Aprendizaje Supervisado

El objetivo del **aprendizaje supervisado** es identificar un modelo que describa el comportamiento de los datos.

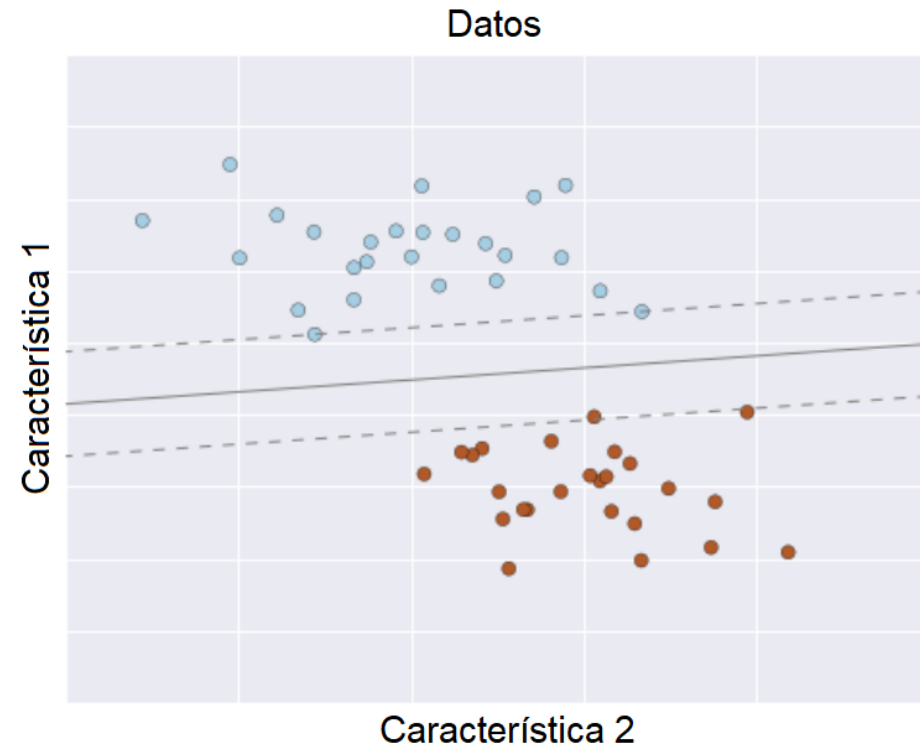
- En esta imagen hay datos bidimensionales, representados por la posición (x, y) en el plano.
- Hay dos clases de datos, representados por el color del punto (azul y rojo).
- A partir de estos datos se espera encontrar un modelo que permita decidir, para un nuevo dato, si pertenece al grupo azul o al grupo rojo.



Tomado de: Python Data Science Handbook, VanderPlas, 2017

Aprendizaje Supervisado

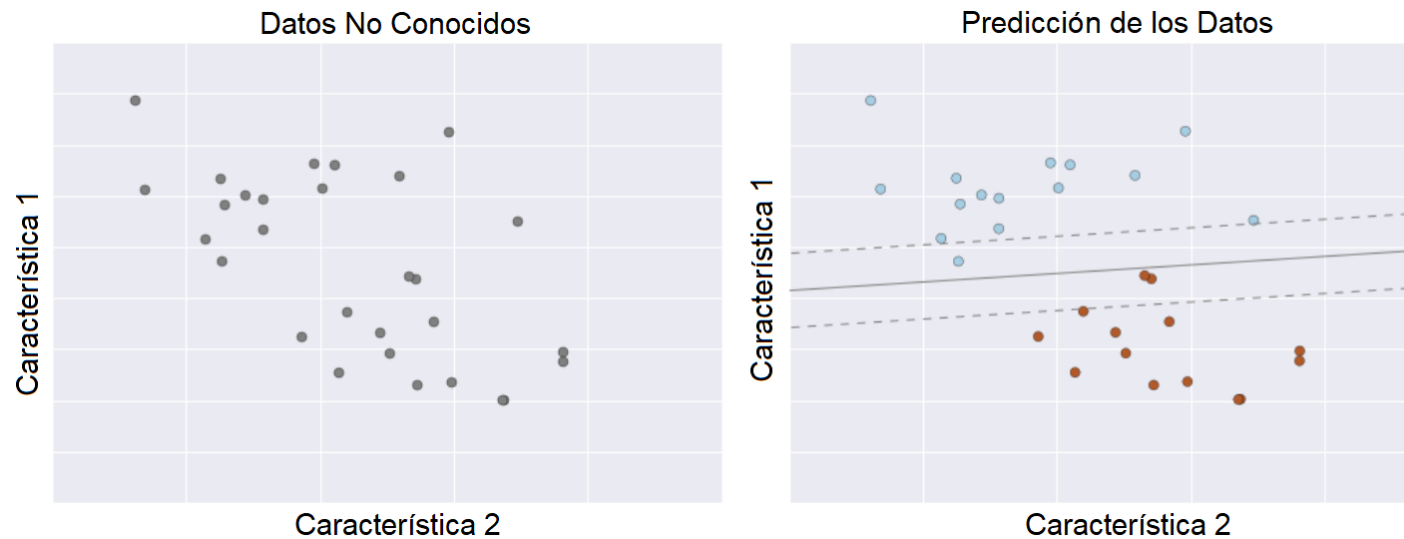
- Existen diferentes tipos de modelos, uno de ellos es asumir que hay una línea que separa las dos clases. Así, cada lado de la línea representa si un dato pertenece a la clase azul o a la clase roja.
- Los parámetros óptimos del modelo, se encuentran a partir de los datos existentes en el proceso llamado **Entrenamiento**.



Tomado de: Python Data Science Handbook, VanderPlas, 2017

Aprendizaje Supervisado

Ahora que el modelo está entrenado (es decir, se encontró la línea óptima que separa las clases), se pueden evaluar nuevos datos no conocidos y clasificarlos en los grupos azul o rojo según corresponda. Esta etapa es conocida como **Predicción**.

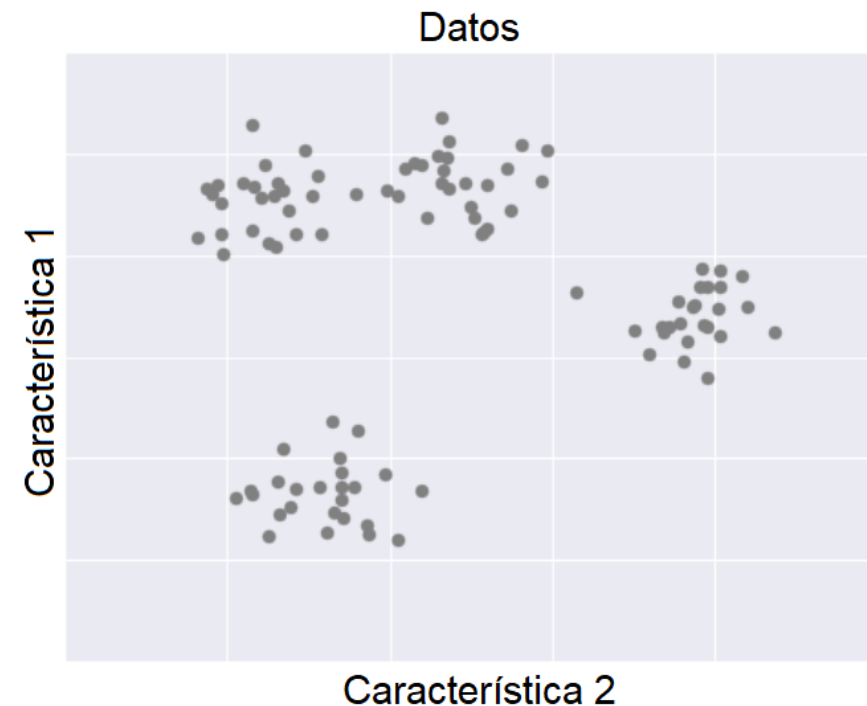


Tomado de: Python Data Science Handbook, VanderPlas, 2017

Aprendizaje de Máquina: Aprendizaje No Supervisado

El objetivo del **aprendizaje no supervisado** es identificar un modelo que encuentre relaciones entre los datos.

- Para lograr esto, la máquina debe aprender de un conjunto de datos **sin etiqueta**.
- Un caso común es la clusterización o agrupación, donde los datos son agrupados de acuerdo a su similitud.



Tomado de: Python Data Science Handbook, VanderPlas, 2017

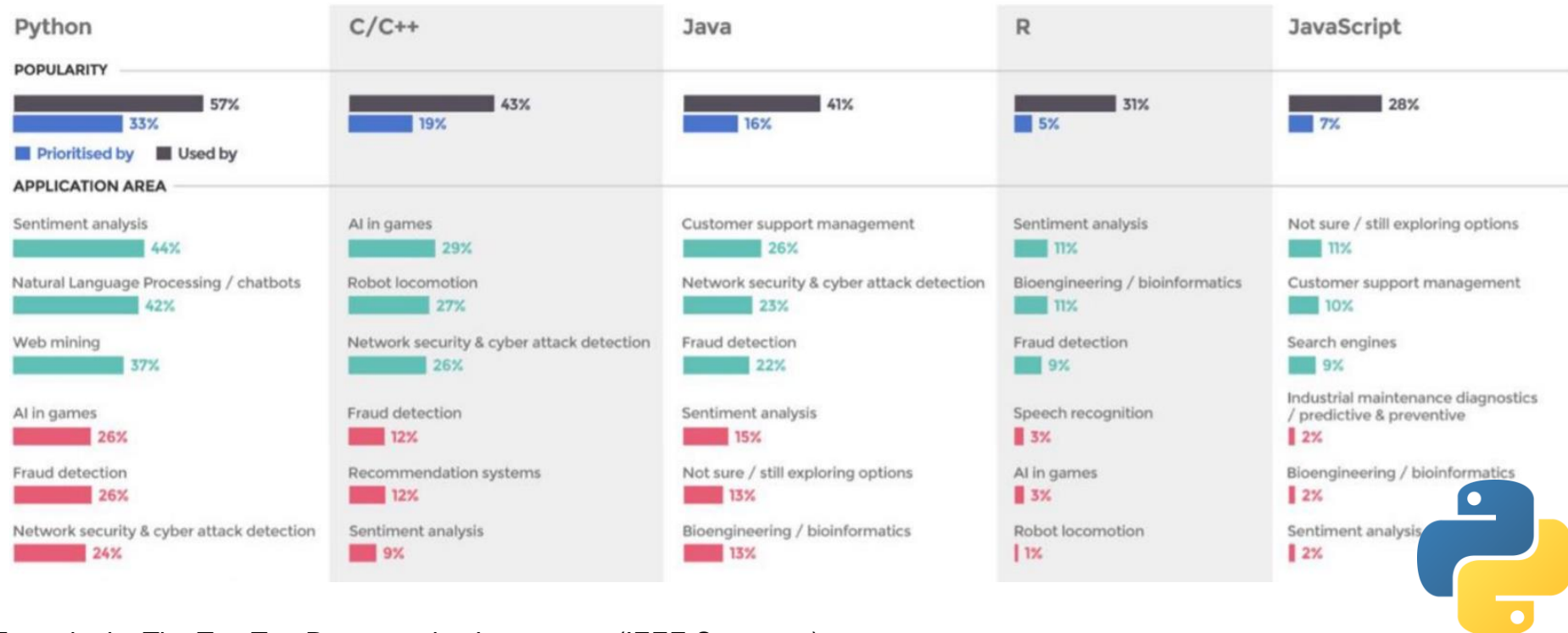
Aprendizaje No Supervisado

- Visualmente se pueden diferenciar algunos grupos.
- El algoritmo debe ser capaz de identificar las relaciones o patrones que existen entre los diferentes datos, sin información previa o históricos, ni etiquetas. Esta es la principal diferencia con el aprendizaje supervisado.



Tomado de: Python Data Science Handbook, VanderPlas, 2017

Herramienta Para el Aprendizaje de Maquinas: Python



Tomado de: The Top Ten Programming Languages (IEEE Spectrum)