

Introducción a Machine Learning

Libro complementario

- Utilizaremos “Introduction to Statistical Learning” de Gareth James como libro complementario.
- Está disponible gratuitamente en línea, podemos conseguirlo en:

[Introduction to Statistical Learning - University of Southern California](http://www-bcf.usc.edu/~gareth/ISL/) ✓

www-bcf.usc.edu/~gareth/ISL/ ▼ Traducir esta página

Home, Download the book PDF (corrected 7th printing). Statistical Learning MOOC covering the entire ISL book offered by Trevor Hastie and Rob Tibshirani.


[Data Sets and Figures](#) · [R Code for Labs](#) · [Get the Book](#) · [About this Book](#)

Springer Texts in Statistics

Gareth James
Daniela Witten
Trevor Hastie
Robert Tibshirani

An Introduction to Statistical Learning

with Applications in R

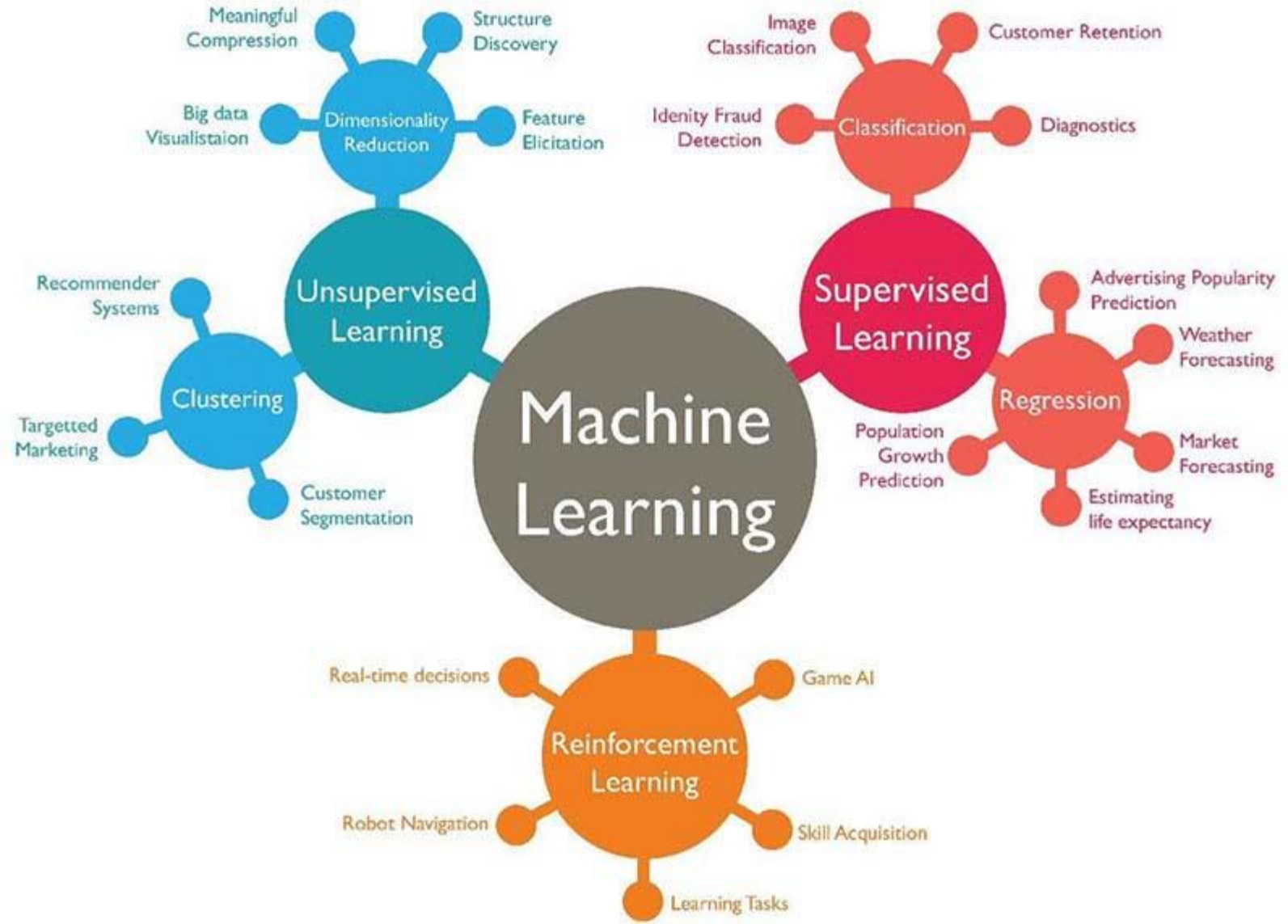
 Springer

Libro Complementario

- Los estudiantes que quieran la teoría matemática deben hacer las lecturas sugeridas.
- Los estudiantes que solo quieren aplicar los modelos y están más interesados en las aplicaciones de Python pueden simplemente enfocarse más en estos materiales.

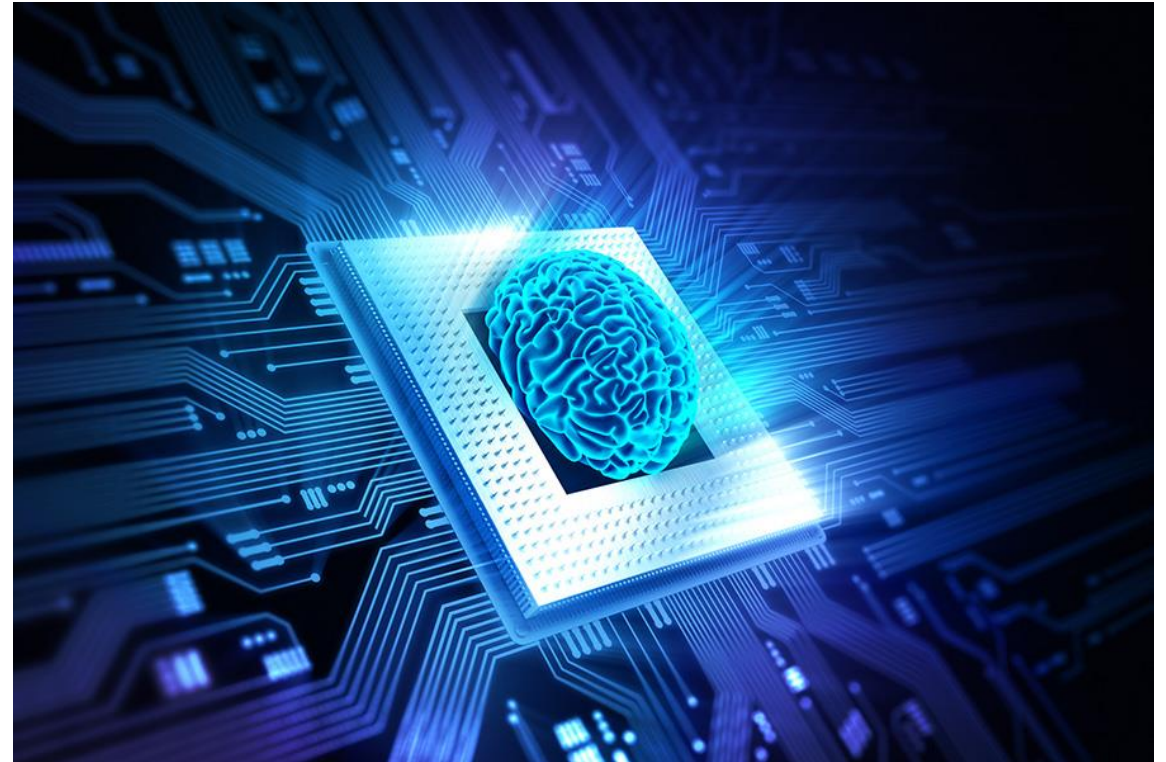
Libro Complementario

- Lea los Capítulos 1 y 2 si quiere obtener una mejor comprensión general antes de continuar con estos materiales.



¿Qué es Machine Learning o Aprendizaje Automático?

- El aprendizaje automático es un método de análisis de datos que automatiza la creación de modelos analíticos.
- Mediante el uso de algoritmos que aprenden iterativamente de los datos, el aprendizaje automático permite que las computadoras encuentren información oculta sin tener que programar explícitamente dónde buscar.

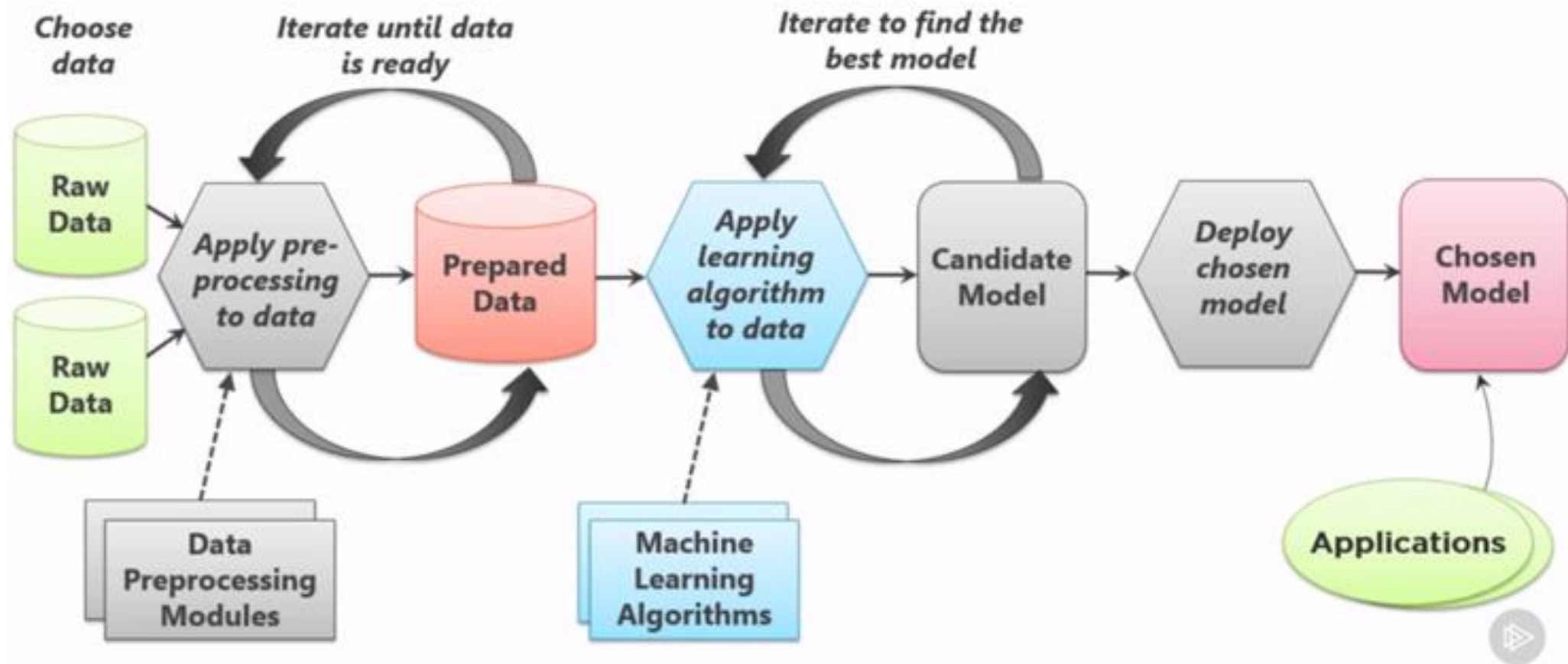


¿Para qué se usa?

- Detección de fraude.
- Resultados de búsqueda web.
- Anuncios en tiempo real en páginas web
- Calificación de crédito y las mejores ofertas siguientes.
- Predicción de fallas de equipos.
- Nuevos modelos de precios.
- Detección de intrusión de red.
- Motores de recomendación
- Segmentación del cliente
- Análisis de sentimiento de texto
- Predecir la rotación de clientes
- Reconocimiento de patrones e imágenes.
- Filtrado de spam de correo electrónico.
- Modelado financiero

Proceso del Aprendizaje Automático

The Machine Learning Process



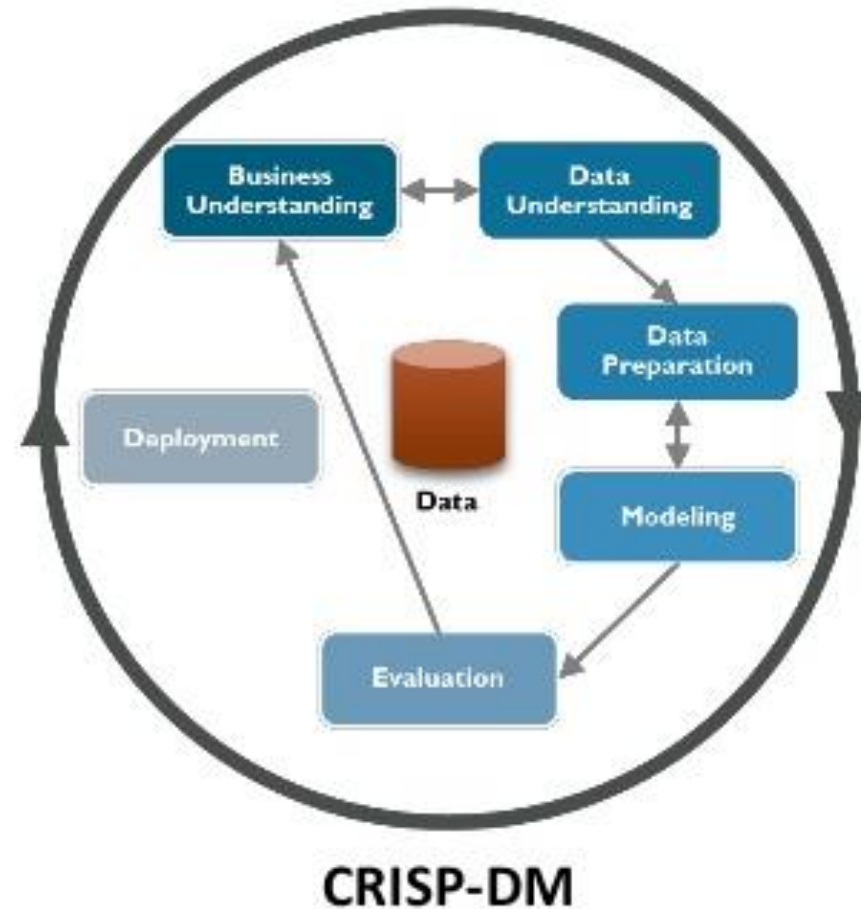
Proceso del Aprendizaje Automático

1. Data Engineering – 80%

- Data extraction
- Data cleaning
- Data transformation
- Data normalization
- Feature extraction

2. Machine Learning – 20%

- Model fitting
- Hyperparameters tuning
- Model evaluation



Aprendizaje Supervisado

- Los algoritmos de aprendizaje supervisados se entrenan usando ejemplos etiquetados, como una entrada donde se conoce el resultado deseado.
- Por ejemplo, unos objetos puede tener puntos de datos etiquetados como "M" (en mal estado) o "B" (buen estado).



buen estado



mal estado



¿Cuáles están en buen estado o mal estado?

Aprendizaje Supervisado

- El algoritmo de aprendizaje recibe un conjunto de entradas junto con las correspondientes salidas correctas, y el algoritmo aprende comparando su salida real con las salidas correctas para encontrar errores.
- Luego modifica el modelo en consecuencia.

Aprendizaje Supervisado

- A través de métodos como la clasificación, la regresión, la predicción y el aumento de gradiente, el aprendizaje supervisado usa patrones para predecir los valores de la etiqueta en datos adicionales no etiquetados.
- El aprendizaje supervisado se usa comúnmente en aplicaciones donde los datos históricos predicen eventos futuros probables.

Aprendizaje Supervisado

- A través de métodos como la clasificación, la regresión, la predicción y el aumento de gradiente, el aprendizaje supervisado usa patrones para predecir los valores de la etiqueta en datos adicionales no etiquetados.
- El aprendizaje supervisado se usa comúnmente en aplicaciones donde los datos históricos predicen eventos futuros probables.

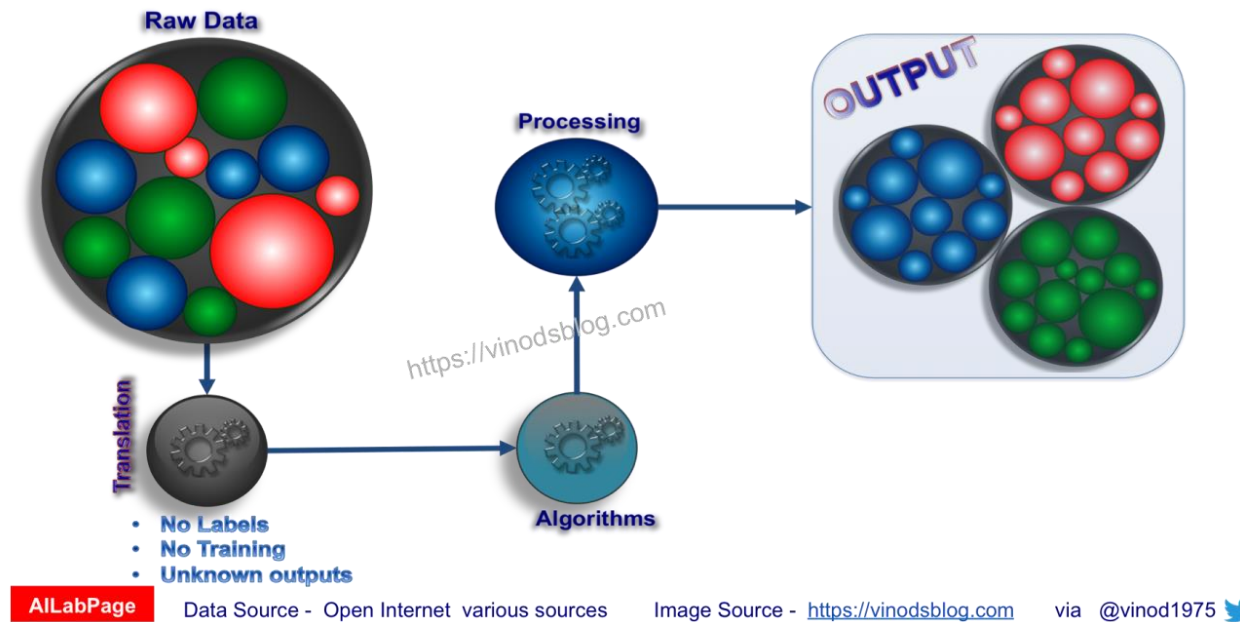
Aprendizaje Supervisado

- Por ejemplo, puede anticipar cuándo es probable que las transacciones con tarjeta de crédito sean fraudulentas o qué cliente de seguros es probable que presente un reclamo.
- O puede intentar predecir el precio de una casa en función de las diferentes características de las casas para las que tenemos datos de precios históricos.



Aprendizaje No Supervisado

- El aprendizaje no supervisado se usa con datos que no tienen etiquetas históricas.
- Al sistema no se le dice la "respuesta correcta". El algoritmo debe descubrir lo que se muestra.
- El objetivo es explorar los datos y encontrar alguna estructura dentro.



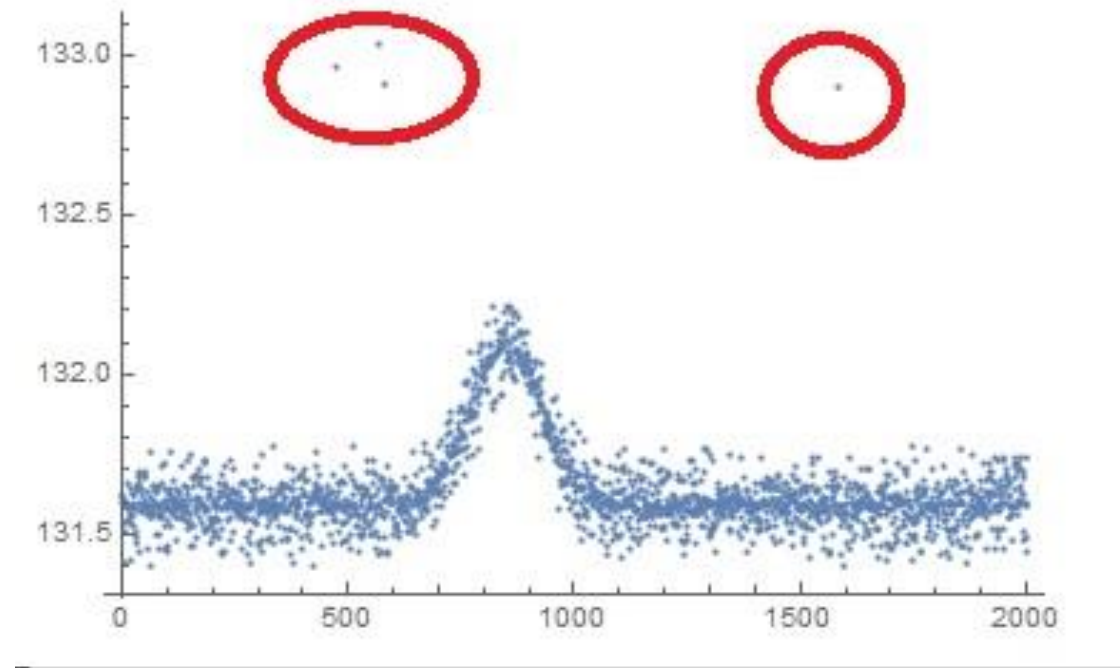
Aprendizaje No Supervisado

- O puede encontrar los principales atributos que separan segmentos de clientes entre sí.
- Las técnicas populares incluyen mapas autoorganizados, mapeo del vecino más cercano, clustering k-means y descomposición de valores singulares.



Aprendizaje No Supervisado

- Estos algoritmos también se utilizan para segmentar temas en un texto, recomendar elementos e identificar valores atípicos de datos.



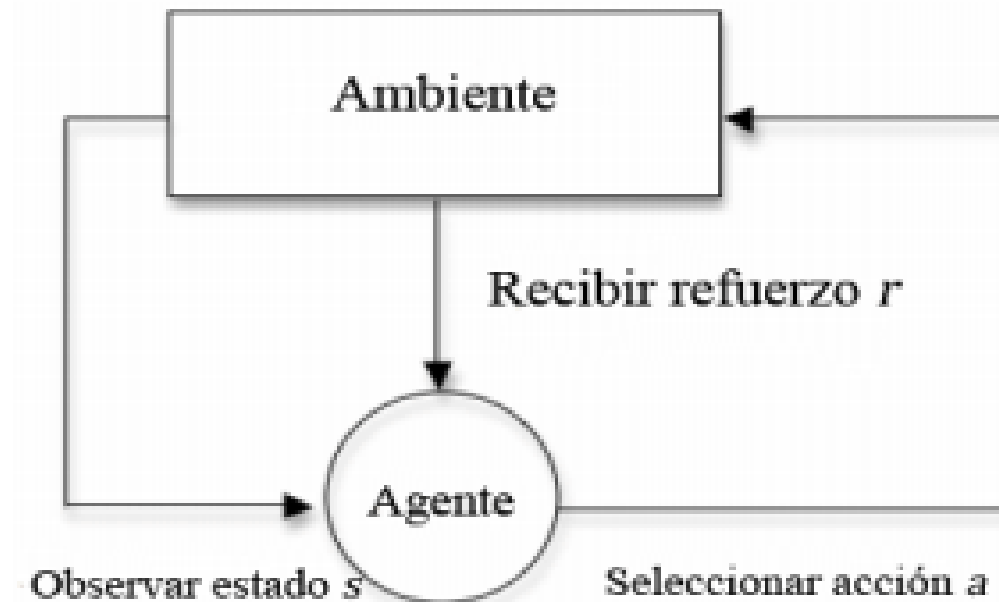
Aprendizaje Reforzado

- El aprendizaje reforzado a menudo se usa para robótica, juegos y navegación.
- Con el aprendizaje reforzado, el algoritmo descubre a través de prueba y error qué acciones rinden las mayores recompensas.



Aprendizaje Reforzado

- Este tipo de aprendizaje tiene tres componentes principales: el agente (el que aprende o el que toma las decisiones), el entorno (todo con lo que el agente interactúa) y las acciones (lo que el agente puede hacer).



Aprendizaje Reforzado

- El objetivo es que el agente elija acciones que maximicen la recompensa esperada durante un período de tiempo determinado.
- El agente alcanzará el objetivo mucho más rápido siguiendo una buena política.
- Entonces, el objetivo en el aprendizaje de refuerzo es aprender la mejor política.

