

Fundamentos del Aprendizaje en la Inteligencia Artificial

1

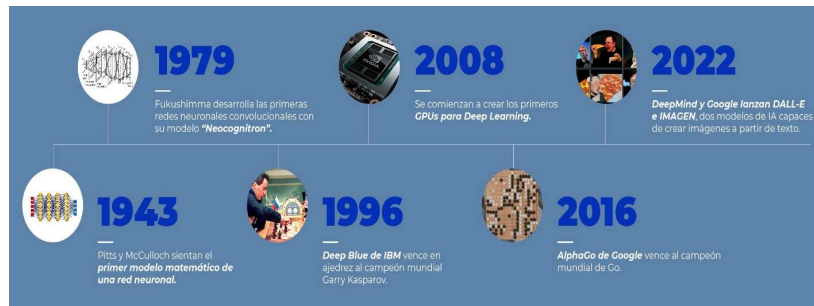
Profesor: Ricardo Rodríguez Bustinza
Email: robust@uni.edu.pe

Historia de la Inteligencia Artificial

- Se puede decir que la Inteligencia Artificial empezó a surgir en 1943 cuando **Warren McCulloch y Walter Pitts** propusieron un **modelo de neurona del cerebro humano y animal**, una abstracción que proporcionó una representación simbólica de la actividad cerebral.
- Podríamos situar el origen de los modelos conexionistas con la **definición de la neurona formal** dada por McCulloch y Pitts en **1943** como un **dispositivo binario con varias entradas y salidas**.
- Más adelante, **Norbert Wiener** elaboró **con estas ideas** junto con otras, dentro del mismo campo, lo que se llamó “**cibernética**”. De aquí nacería, sobre los años 50, la Inteligencia Artificial.

Perspectiva Histórica

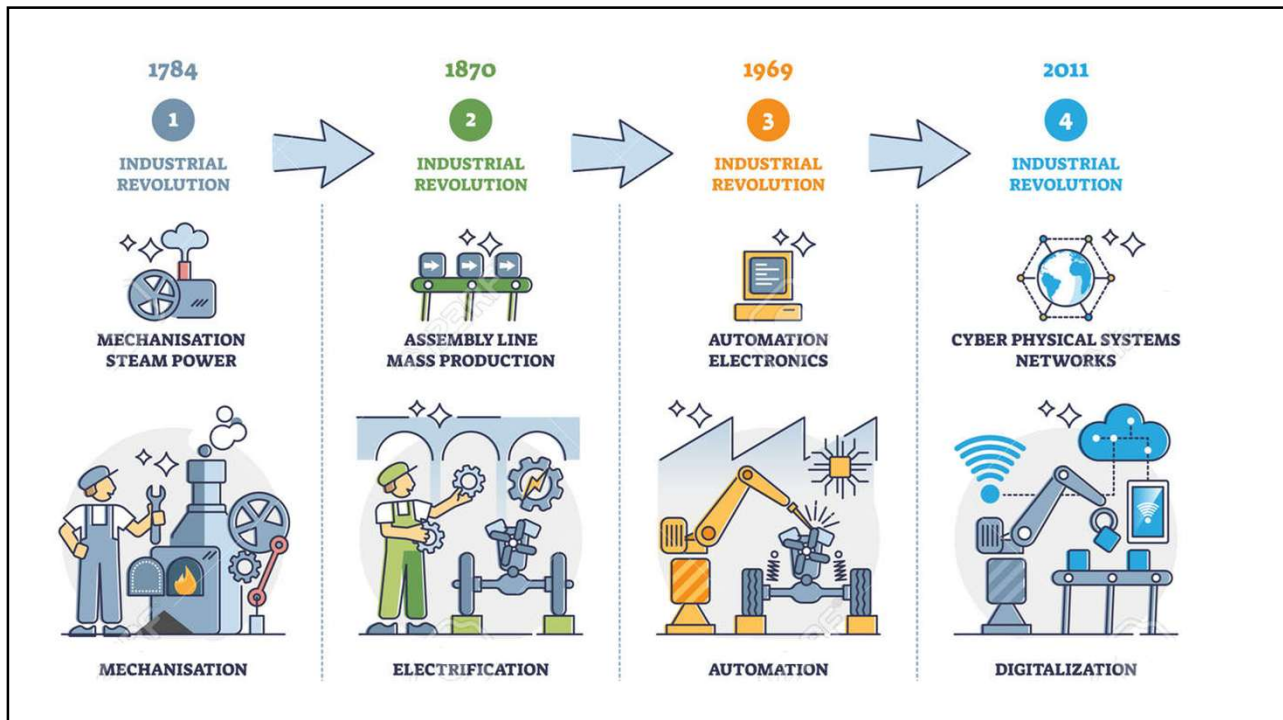
- 1940s: Hebb, McCulloch and Pitts
 - Mecanismo de aprendizaje en neuronas biológicas
 - Redes de elementos 'neuronales' pueden hacer cómputo aritmético
- 1950s: Rosenblatt, Widrow and Hoff
 - Primeras redes artificiales (RNA) prácticas y reglas de aprendizaje



Perspectiva Histórica

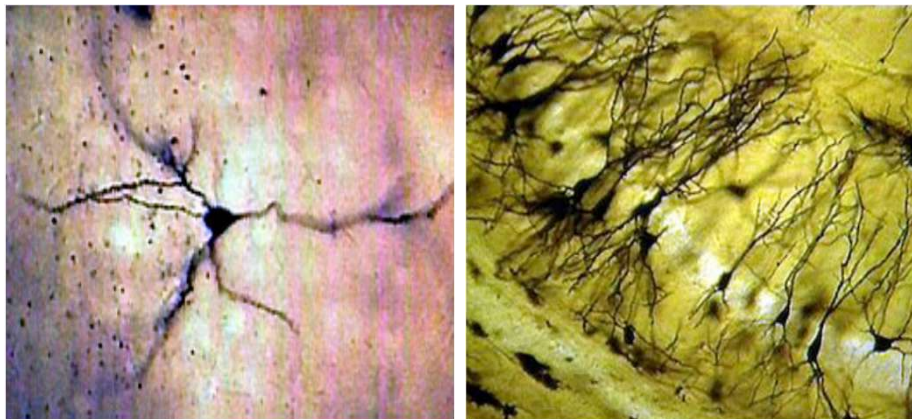
- 1960s: Minsky and Papert
 - Demostraron las limitaciones de las RNA existentes, no se estaban desarrollando nuevos algoritmos de aprendizaje.
- 1970s: Amari, Anderson, Fukushima, Grossberg, Kohonen
 - Continúan trabajando en el tema pero a pasos mas lentos.
- 1980s: Grossberg, Hopfield, Kohonen, Rumelhart
 - Desarrollos importantes causan un resurgimiento.



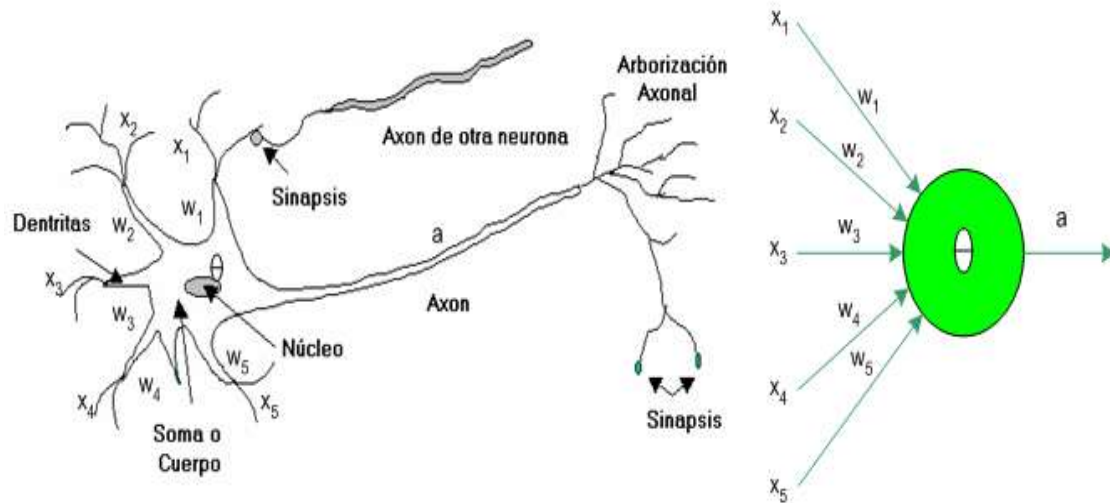


Neurona Biológica

- Se estima que hay 26,000 millones de neuronas en el cerebro humano y en un área de 1 mm^2 hay aproximadamente 50,000. Ellas se comunican a través de una red de conexiones sinápticas teniendo una densidad de aproximadamente 10^4 sinapsis por neurona. Su procesamiento en paralelo hacen posible las habilidades del cerebro.



Partes de la Neurona Biológica

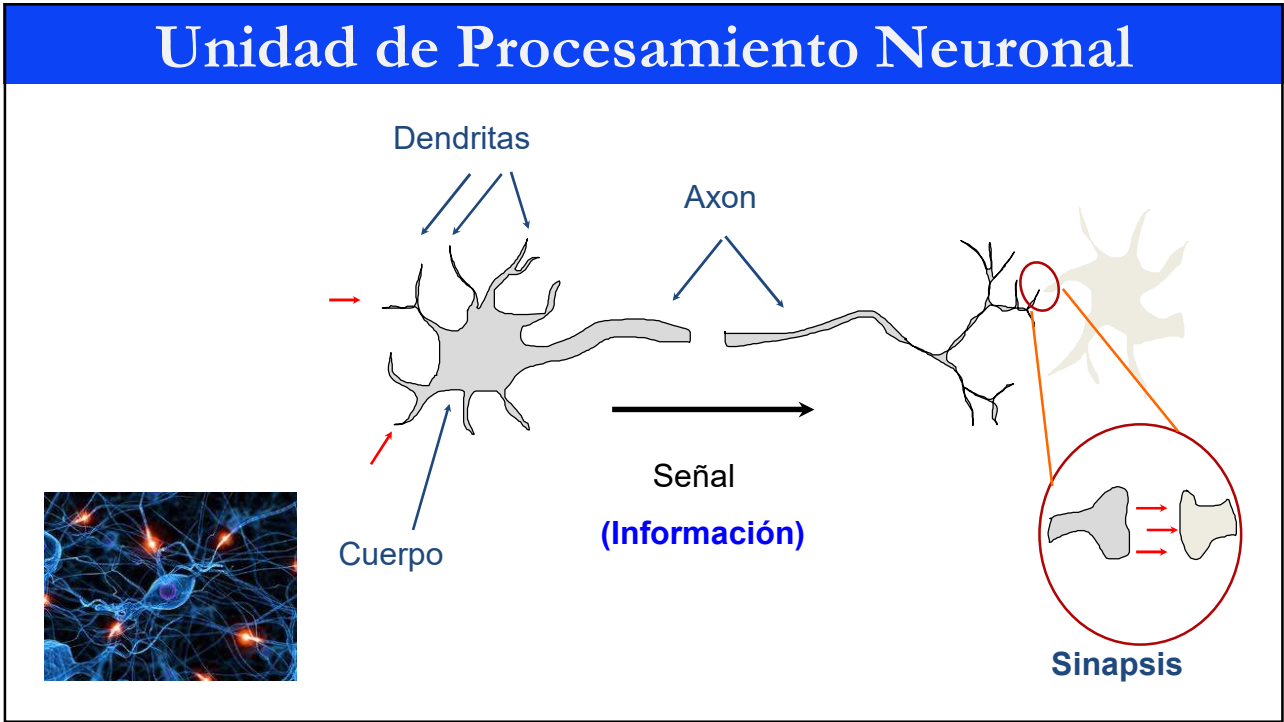
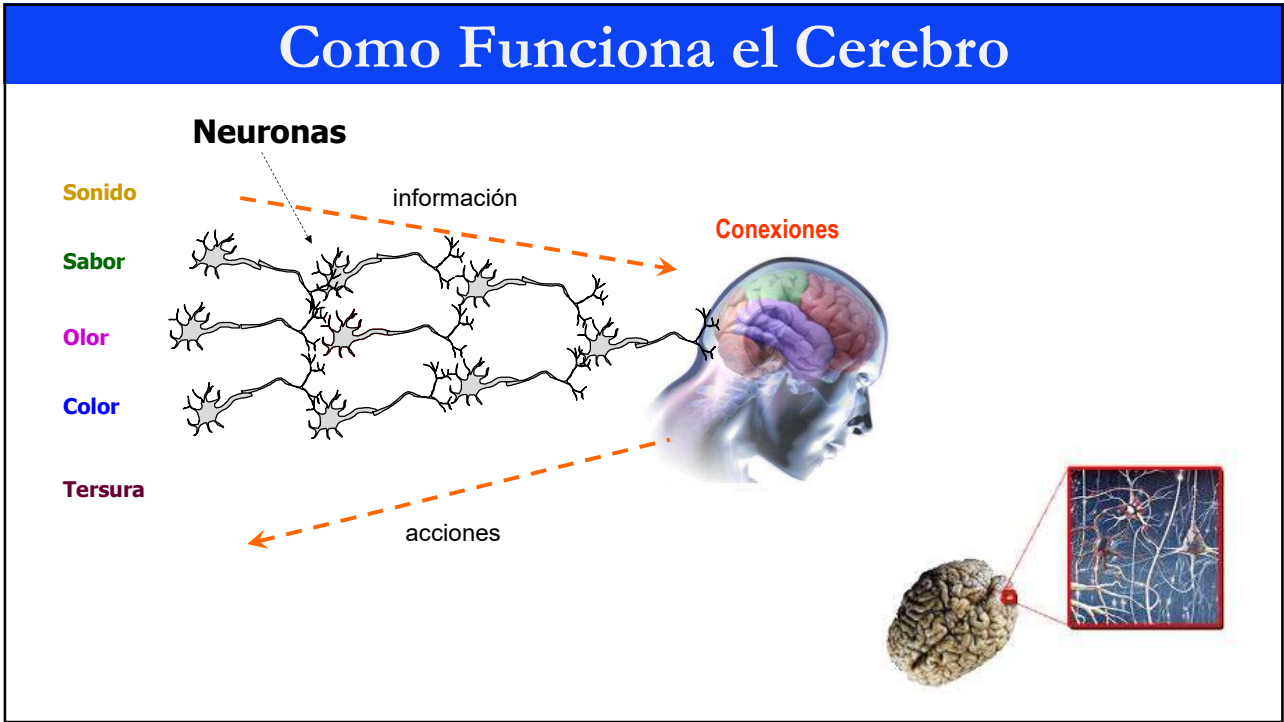


Cerebro Humano

- El cerebro humano tiene varias características deseables para un sistema artificial, por ese motivo sirve como patrón para su estudio.

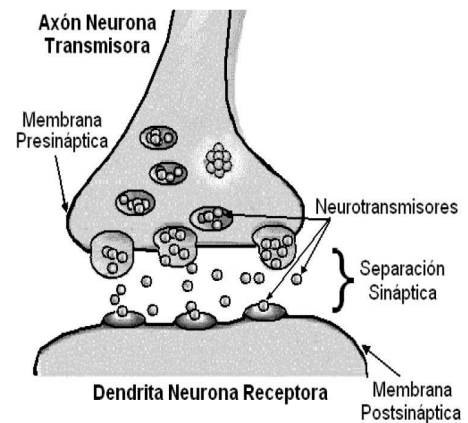
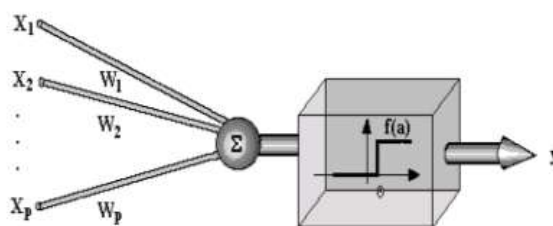


- Es robusto y tolerante a fallas, diariamente mueren neuronas sin afectar su desempeño.
- Es flexible, se ajusta a nuevos ambientes de aprendizaje no hay que programarlo.
- Puede manejar información difusa, con ruido o inconsistente.
- Es altamente paralelo y es pequeño compacto y consume poca energía.

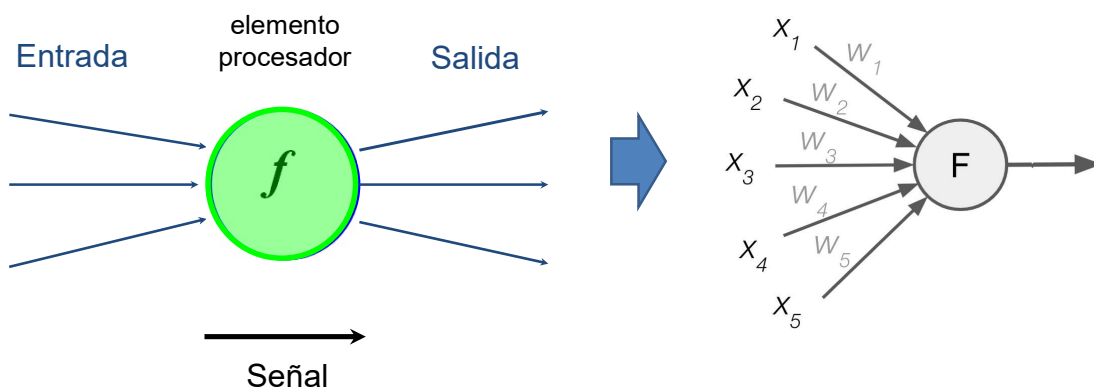


Elementos Básicos de una RNA

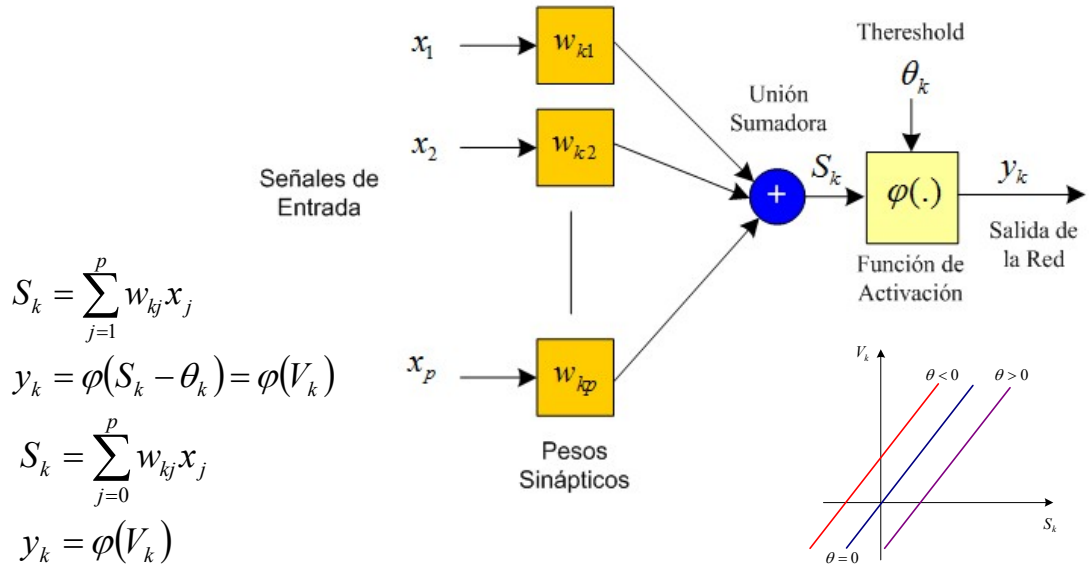
1. Las interligaciones de la neurona es caracterizada por un peso (+ para peso excitatorio y - para peso inhibitorio).
2. Sumador de las señales de entrada a la manera de un combinador lineal.
3. Función de activación para limitar la amplitud (salida) de una neurona. $[0, 1]$ o $[-1, 1]$.



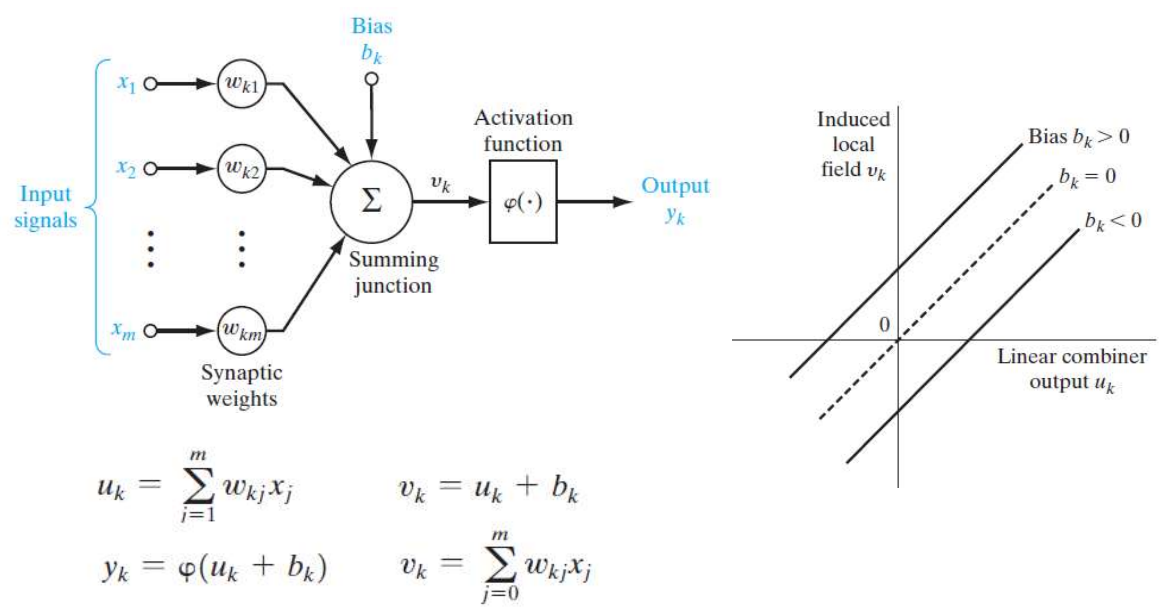
Unidad de Procesamiento



Topología Neuronal (1943)



Modelo No Lineal



Modelo No Lineal

Fixed input $x_0 = +1$

Inputs

Synaptic weights (including bias)

Summing junction

Activation function

Output y_k

$y_k = \varphi(v_k)$

$x_0 = +1$

$w_{k0} = b_k$

$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases}$

$y_k = \begin{cases} 1 & \text{if } v_k \geq 0 \\ 0 & \text{if } v_k < 0 \end{cases}$

$v_k = \sum_{j=1}^m w_{kj}x_j + b_k$

El Perceptron

- Frank Rosenblatt desarrolla una prueba de convergencia en 1962 y definió el rango de problemas para los que su algoritmo aseguraba una solución. Además propuso a los 'Perceptrons' como herramienta computacional.

neurona

x_1

x_2

w_1

w_2

Σ

a

$F(a)$

y

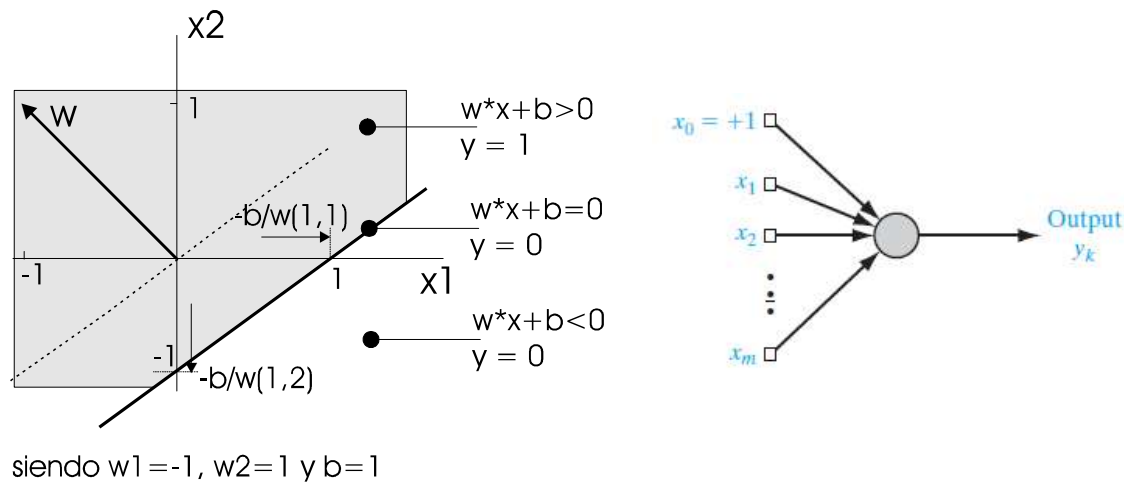
$y = F(a)$

$y = 1, \text{ si } x_1w_1 + x_2w_2 \geq \theta$

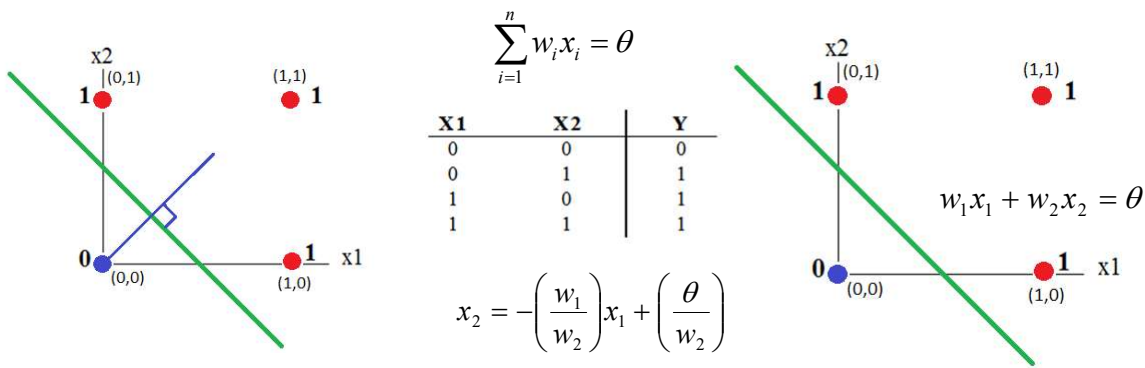
$y = 0, \text{ si } x_1w_1 + x_2w_2 < \theta$

Separabilidad Lineal

- El Perceptron tiene la capacidad de clasificar vectores de entrada dividiendo el espacio de entrada en dos regiones mediante una hiperplano. Si el espacio de entrada no es separable de esta manera, un Perceptron no encontrará solución.

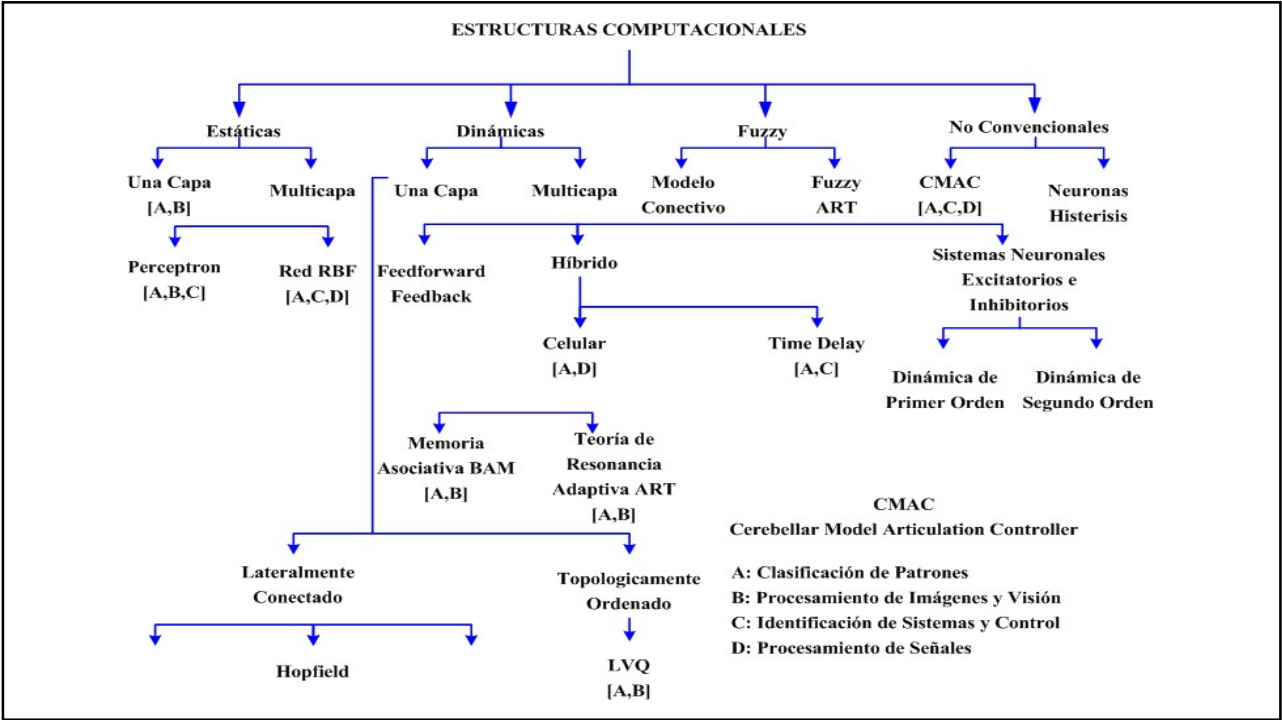
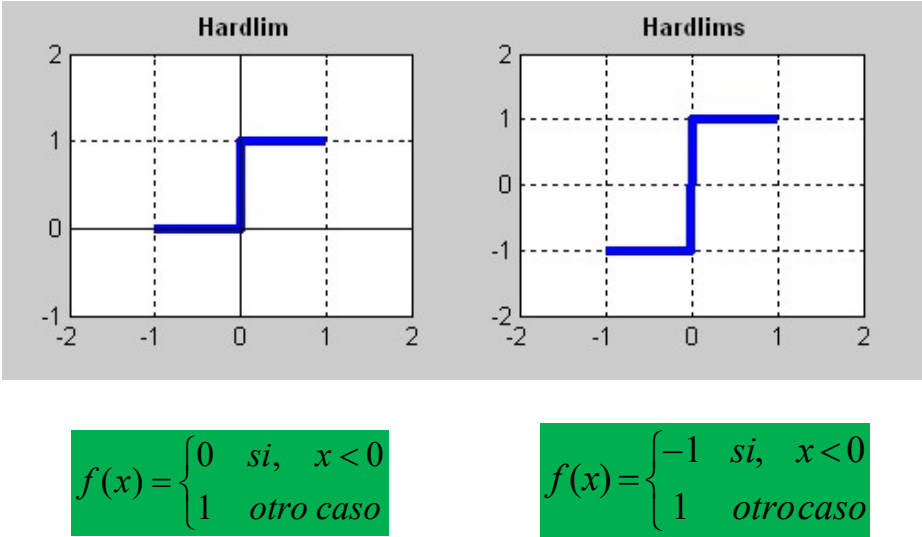


¿Qué Tipos de Problemas Resuelve?



- Un hiperplano es un objeto de dimensión $n - 1$ que actúa en un espacio de dimensión n .
- En general un Perceptron de n entradas puede ejecutar cualquier función que esté determinada por un hiperplano que corte un espacio de dimensión n . ¿Implicaciones?

Función de Activación



Tipos de Aprendizaje

■ Aprendizaje Supervisado

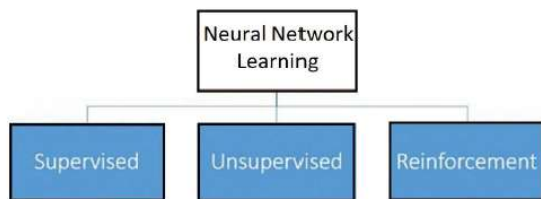
- A la red se le provee un conjunto de ejemplos que representen la relación de entradas/salida deseada

■ Aprendizaje por Refuerzo

- A la red sólo se le da una indicación sobre el grado de desempeño de la red.

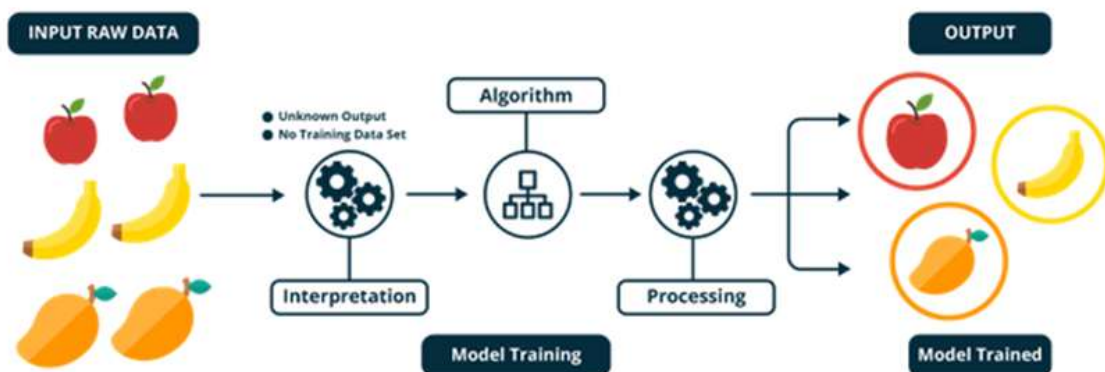
■ Aprendizaje No Supervisado

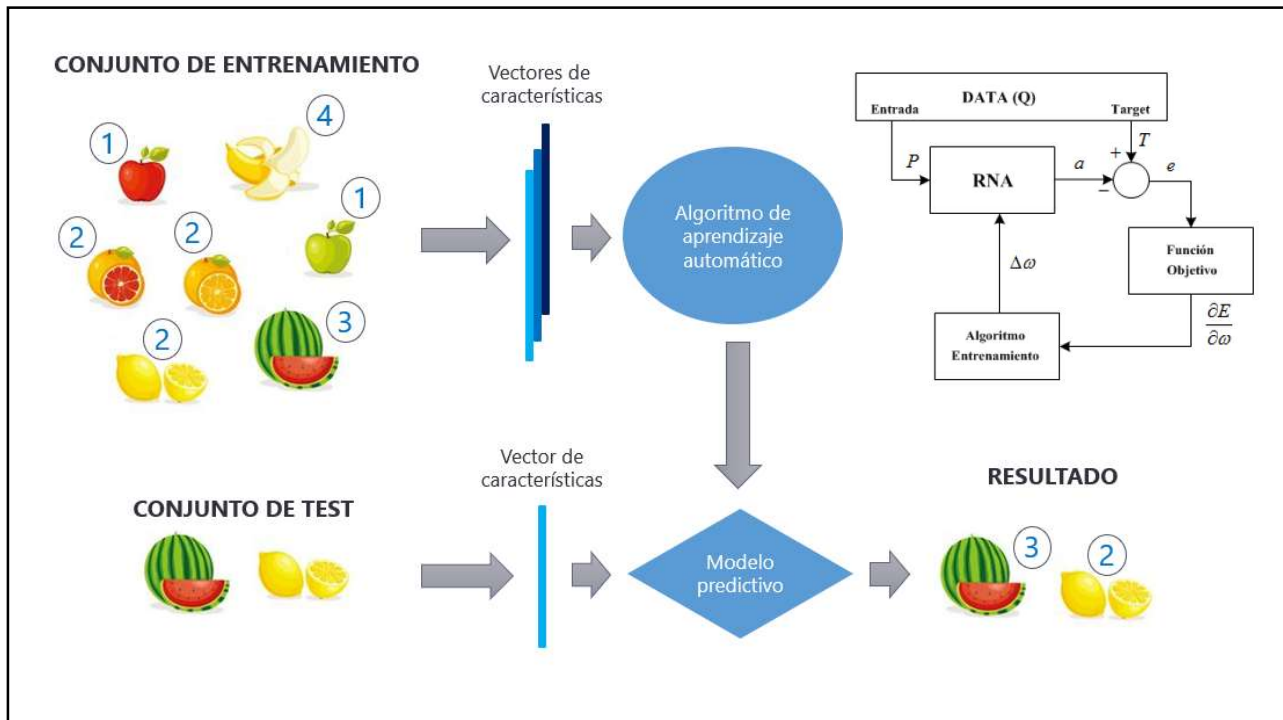
- La única información disponible para la red son las entradas. La red aprende a formar categorías a partir de un análisis de las entradas. La red se auto organiza (aprende sola).



Aprendizaje Supervisado

El **aprendizaje supervisado** es una técnica para deducir una función a partir de datos de entrenamiento. Los datos de entrenamiento consisten de pares de objetos (normalmente vectores): una componente del par son los datos de entrada y el otro, los resultados deseados.





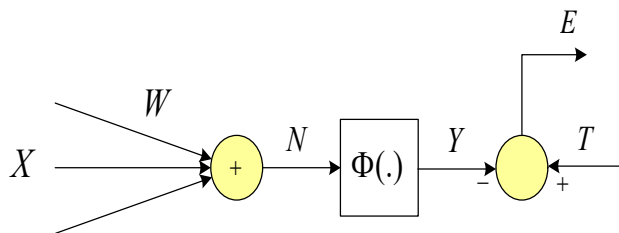
Aprendizaje de una RNA

Secuencia de eventos que ocurren durante el aprendizaje:

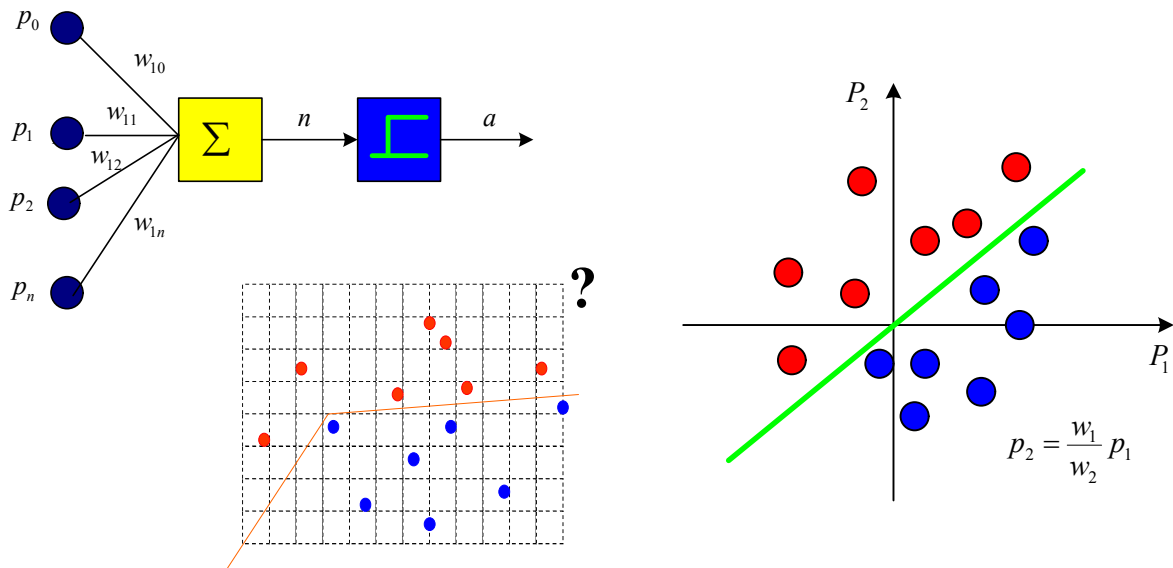
1. La **estimulación** de la RNA por un ambiente
2. La RNA sufre **cambios** en sus parámetros libres como resultado de dicha estimulación.

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t+1)$$

3. La RNA **responde de manera diferente** al ambiente debido a los cambios que ocurrieron en su estructura interna



Clasificación de Patrones



Aplicaciones

- **Espacio aéreo**

- Pilotos automáticos de alto desempeño, simulaciones y predicciones de trayectoria de vuelo, sistemas de control de vuelo, detección de fallas en componentes de la nave.



- **Bancos**

- Lectores de documentos, evaluadores de asignación de crédito, identificador de firmas.



- **Electrónica**

- Predicción de secuencias de códigos, control de procesos, análisis de fallas de circuitos, visión de máquina, síntesis de voz, modelado no lineal.



- **Robótica**

- Control de trayectorias, control de manipuladores, sistemas de visión. Robótica móvil y la navegación autónoma.

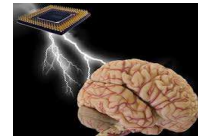


- **Voz**

- Reconocimiento de voz, compresión de voz, sintetizadores de texto a voz.

- **Telecomunicaciones**

- Compresión de datos e imágenes, servicios automáticos de información, traducción de lenguaje hablado en tiempo real.



- **Transportación**

- Sistemas ruteadores, diagnóstico de motores, tiempos y movimientos.

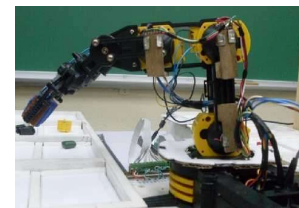
- **Seguridad**

- Reconocimiento de rostros, identificación y acceso de personas



- **Manufactura**

- Control de procesos de manufactura, análisis y diseño de productos, diagnóstico de máquinas y procesos, identificación de partes en tiempo real, sistemas de inspección de calidad, predicción de fin de proceso, análisis de mantenimiento de máquinas, modelado de sistemas dinámicos.



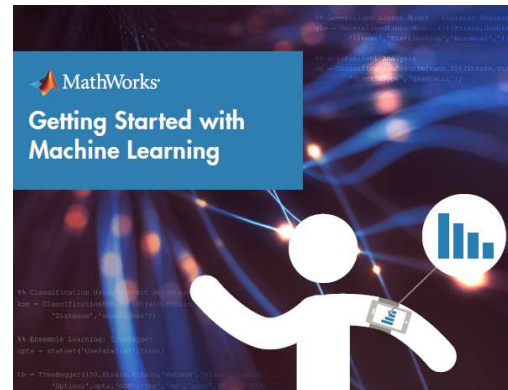
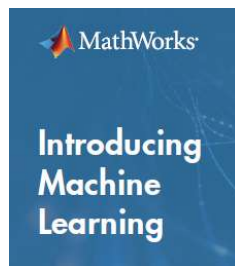
- **Medicina**

- Detección de cáncer mamario o en la piel, análisis de EEG y ECG, diseño de prótesis, optimización de tiempos de trasplante, reducción de gastos en hospitales.



Machine Learning

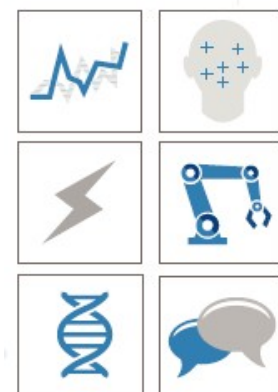
- El aprendizaje automático enseña a las computadoras a realizar las tareas que los humanos y animales aprenden de la experiencia.
- Los algoritmos de aprendizaje automático utilizan métodos computacionales para "aprender" la información directamente de los datos sin depender de una ecuación predeterminada como modelo.



Aplicaciones del Mundo Real

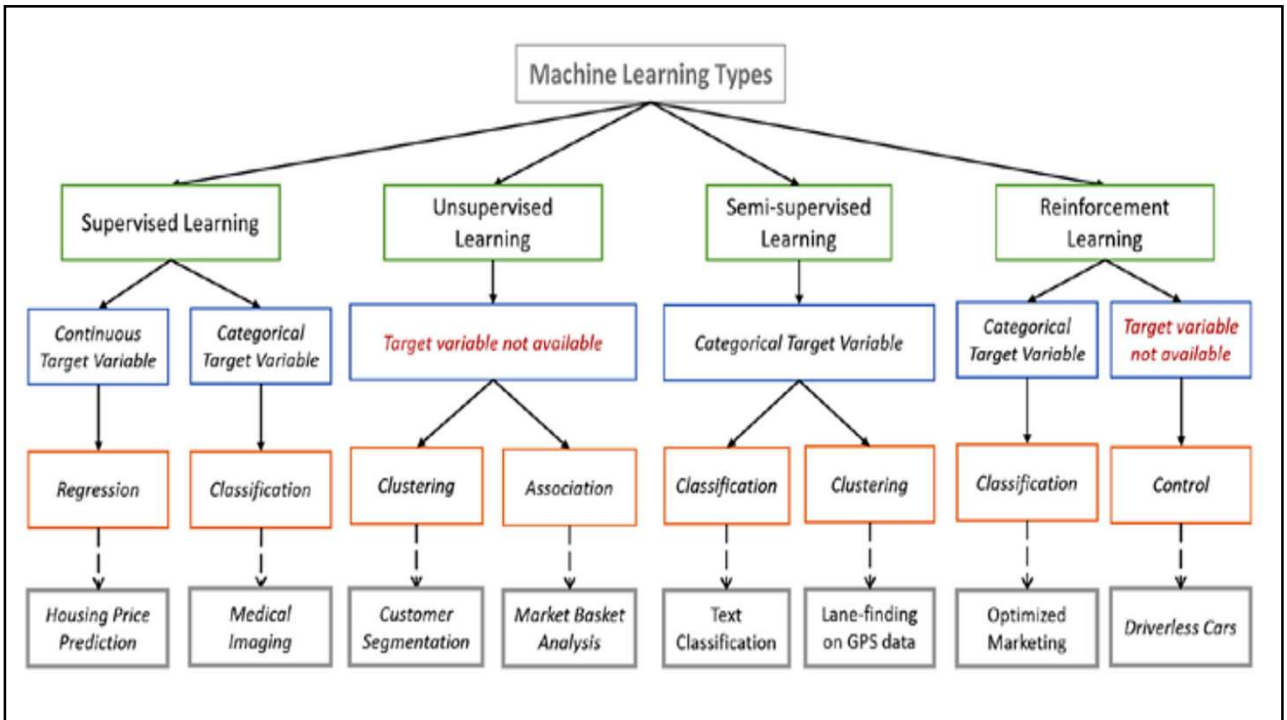
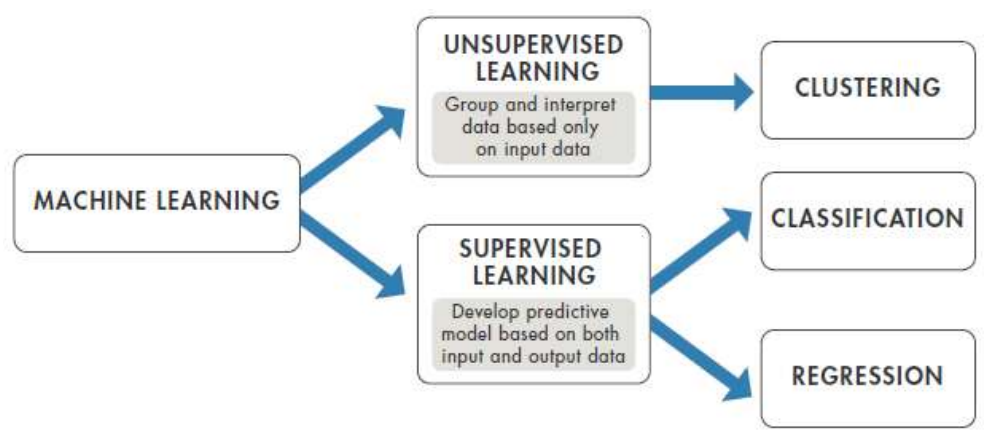
Con el aumento de los grandes datos, el aprendizaje especial importancia para resolver problemas en áreas:

- **Finanzas computacionales**, para crédito puntuación y algoritmo de negociación.
- **Procesamiento**, imágenes y visión por ordenador, para reconocimiento facial, detección de movimiento y detección de objetos.
- **Biología computacional**, para detección del tumor, descubrimiento de fármacos.
- **Industria**, aplicado al parque automotor, aeroespacial y fabricación, para mantenimiento predictivo.
- **Procesamiento**, es una tecnología de machine learning que brinda a las computadoras la capacidad de interpretar, manipular y comprender el lenguaje humano.



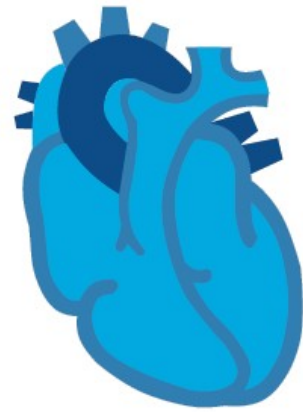
Como Trabaja una Maquina de Aprendizaje

El aprendizaje automático utiliza dos tipos de técnicas: aprendizaje supervisado que entrena un modelo sobre datos de entrada y salida que puede predecir los resultados futuros y el aprendizaje sin supervisión, que encuentra patrones ocultos o estructuras intrínsecas en los datos de entrada.

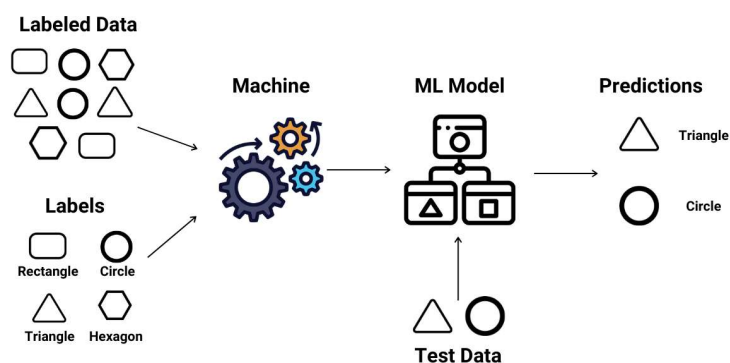


Aprendizaje Supervisado

- El objetivo de una maquina con aprendizaje supervisado es construir un modelo que hace predicciones basadas en evidencia en presencia de incertidumbres.
- Un algoritmo de aprendizaje supervisado toma un conjunto conocido de datos de entrada y respuestas conocidas a los datos (salida) y modelo para generar predicciones razonables para la respuesta a nuevos datos.
- El aprendizaje supervisado utiliza técnicas de clasificación y regresión para desarrollar modelos predictivos.

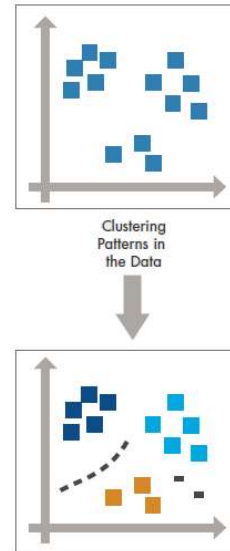


- Las técnicas de clasificación predicen respuestas discretas por ejemplo, si un correo electrónico es genuino o spam, o si un tumor es canceroso o benigno. Se emplean modelos de clasificación de datos de entrada en categorías.
- Las técnicas de regresión predicen respuestas continuas, por ejemplo, cambios en la temperatura o fluctuaciones la demanda de energía. Las aplicaciones típicas incluyen previsión de carga de electricidad y comercio (trading) algorítmico.



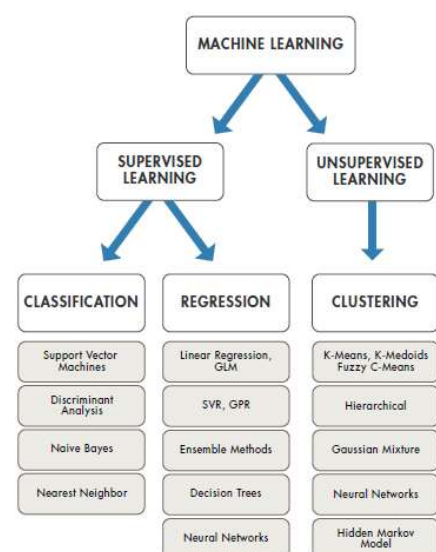
Aprendizaje No Supervisado

- El aprendizaje no supervisado encuentra patrones ocultos o estructuras intrínsecas en los datos. Se utiliza para extraer inferencias de conjuntos de datos que consisten de datos de entrada sin respuestas etiquetadas.
- Clustering es el aprendizaje más común sin supervisión técnica. Se utiliza para el análisis exploratorio de datos para encontrar patrones o agrupaciones en los datos.
- Las aplicaciones para el agrupamiento incluyen análisis de secuencias de genes, investigación de mercado y reconocimiento de objetos.



¿Cómo Decidir que Algoritmo Usar?

- Elegir el algoritmo correcto parece abrumador—allí son docenas de algoritmos de aprendizaje de máquina supervisado y sin supervisión, y cada uno toma un acercamiento diferente al aprendizaje. No hay mejor método o un tamaño único.
- Encontrar el algoritmo es en parte sólo de la prueba y el error, incluso muy experimentado los científicos de datos no afirman si un algoritmo funcionará sin la prueba, sin embargo, la selección de algoritmos también depende del tamaño y tipo de datos con los que está trabajando, las ideas que desea obtener los datos y cómo se utilizarán esos conocimientos.

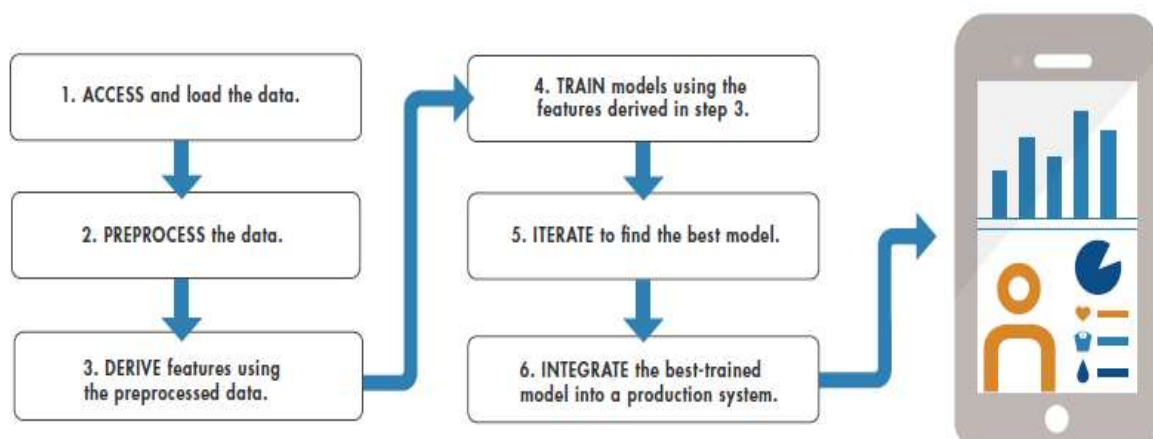


¿Cuándo Usar Machine Learning?

- Considere utilizar machine learning cuando tenga una tarea compleja o problema que involucra una gran cantidad de datos y muchas variables, pero ninguna fórmula o ecuación existente.



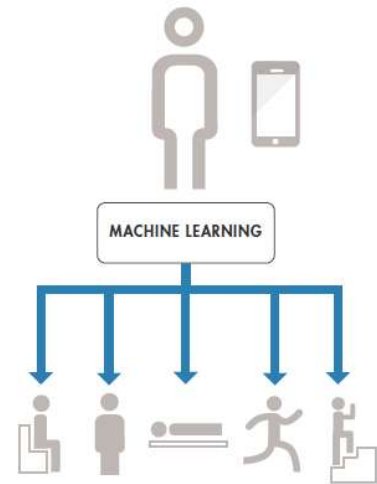
Presentación del Flujo de Trabajo



Formación de un Modelo

Formación de un modelo para clasificar las actividades físicas

- Este ejemplo se basa en una aplicación App de monitorización de la salud por medio del teléfono celular.
- La entrada consta de tres datos de sensores axiales del teléfono, acelerómetro y giroscopio. Las respuestas (o salida) son las actividades realizadas-caminar, estar de pie, correr, subir escaleras, o acostado.
- Queremos utilizar los datos de entrada para formar un modelo de identificar estas actividades. Dado que nuestro objetivo es la clasificación, aplicando el aprendizaje supervisado. El modelo (o clasificador) entrenado se integrará en una aplicación para ayudar a los usuarios a rastrear sus niveles de actividad durante el día.



Construir un Modelo

Supongamos que su objetivo es predecir si algún usuario desconocido disfrutará de algún curso desconocido. Simplemente debe responder “sí” o “no”. Para hacer una conjetura, se le permite preguntas binario sobre el usuario/curso en consideración.

- Usted: ¿Esta el curso bajo la consideración de sistemas?
- Yo: si
- Usted: ¿Ha tomado el estudiante otros cursos de sistemas?
- Yo: si
- Usted: ¿Al estudiante le gusta la mayoría de los cursos anteriores?
- Yo: si
- Usted: Se predice que al estudiante le gusta los cursos.

