



Machine Learning



Machine Learning

Técnicas empleadas para que las máquinas **aprendan** de forma autónoma.

Aprender -> Generalizar comportamientos a partir de una información suministrada en forma de ejemplos.



Machine Learning

Clasificación de las técnicas de Machine Learning según el tipo de aprendizaje:

- **Aprendizaje supervisado:** Los datos traen relacionado un objetivo
- **Aprendizaje No supervisado:** Los datos no traen relacionado ningún objetivo



Machine Learning

Aprendizaje supervisado

- Clasificación
- Regresión (predicción)

Aprendizaje No supervisado

- Agrupación



Machine Learning

Aprendizaje supervisado

- Generalized Linear Models
- Nearest Neighbors (KNN)
- Support Vector Machines (SVM)
- Naive Bayes
- Decision Trees
- Neural network models (supervised)
- ...

Aprendizaje No supervisado

- Gaussian mixture models
- Manifold learning
- Clustering
- Density Estimation
- Covariance estimation
- Neural network models (unsupervised)
- ...



Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Método de clasificación supervisada que sirve para estimar la función de densidad $F(x/C_j)$ de las predictoras x por cada clase C_j .

Estima el valor de la función de densidad de probabilidad o directamente la probabilidad a posteriori de que un elemento x pertenezca a la clase C_j a partir de la información proporcionada por un conjunto de datos de entrenamiento.

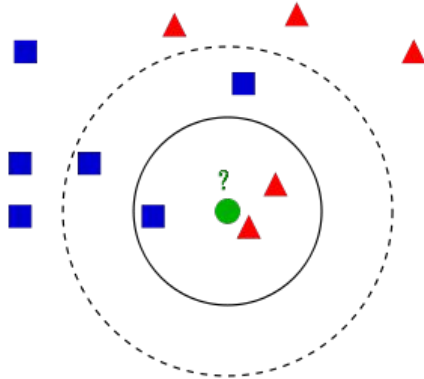
Es una técnica que utiliza aprendizaje vago, no entusiasta -> no existe una generalización previa.



Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Sistema para clasificar futuros deudores de un banco:



■ Cliente que Pagó

▲ Cliente que Incumplió

?
● Futuro Deudor (a clasificar)

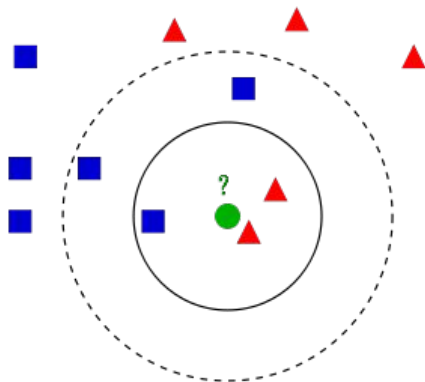


Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Variables que se correlacionan con el cumplimiento en el pago:

- Edad
- Estado civil
- Número de hijos
- Nivel de estudio
- ...



De acuerdo a estas variables ...

- ($k = 3$) De los 3 clientes más parecidos: la mayoría pagó o no?
- ($k = 5$) De los 5 clientes más parecidos: la mayoría pagó o no?



Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Algoritmo:

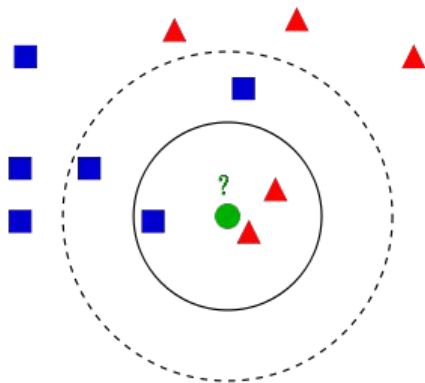
Conjunto de entrenamiento:

Vectores p-dimensionales (p-> # de atributos)

$$x_i = (x_{1i}, x_{2i}, \dots, x_{pi}) \in X$$

Métrica para calcular cercanía de vecinos - Distancia euclidiana:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^p (x_{ri} - x_{rj})^2}$$





Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Ejemplo: Clasificador de especies de plantas Iris



```
from sklearn import datasets
```

```
iris = datasets.load_iris()
```

```
type(iris)
```

```
dir(iris)
```

```
iris.DESCR
```

```
iris.feature_names
```

```
iris.target_names
```

```
iris.data
```

```
iris.target
```

Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Ejemplo: Clasificador de especies de plantas Iris



```
from sklearn.model_selection import train_test_split
```

```
x_train, x_test, y_train, y_test = train_test_split(iris.data, iris.target)
```

```
x_train
```

```
y_train
```

```
x_test
```

```
y_test
```



Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Ejemplo: Clasificador de especies de plantas Iris



```
from sklearn.neighbors import KNeighborsClassifier #Importamos al método
```

```
knn = KNeighborsClassifier(n_neighbors=5) #Declaramos el método
```

```
knn.fit(x_train, y_train) #Ajustamos el método
```

```
knn.score(x_test, y_test) # El porcentaje de acertamiento del método
```

```
knn.predict(x_test)
```

```
Y_test
```

```
knn.predict(x_test[2:3][:4])
```

```
y_test[2:3][:4]
```



Machine Learning

Aprendizaje supervisado - Nearest Neighbors (KNN)

Ejemplo: Clasificador de especies de plantas Iris



```
from sklearn.neighbors import RadiusNeighborsClassifier
```

```
knn_r = RadiusNeighborsClassifier(radius=5)
```

```
knn_r.fit(x_train, y_train) #Ajustamos el método
```

```
knn_r.score(x_test, y_test) # El porcentaje de acertamiento del método
```

```
knn_r.predict(x_test)
```

```
x_test
```

```
knn_r.predict(x_test[2:3][:4])
```

```
y_test[2:3][:4]
```



Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

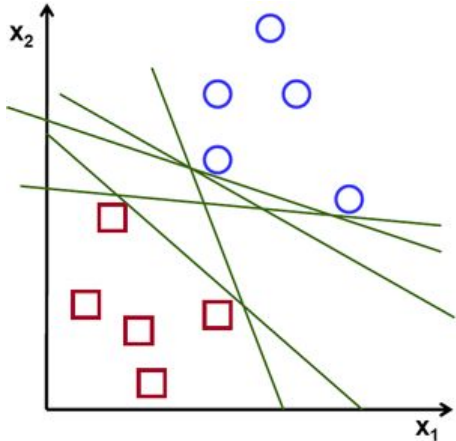
Técnica de aprendizaje supervisado creada por Vladimir Vapnik que permite realizar tareas de clasificación y de regresión mediante la **creación de hiperplanos que separan los datos de entrada**.

Cuando los datos no son separables linealmente se utilizan **funciones kernels que llevan los datos a dimensiones superiores** donde si es posible dicha separación.



Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

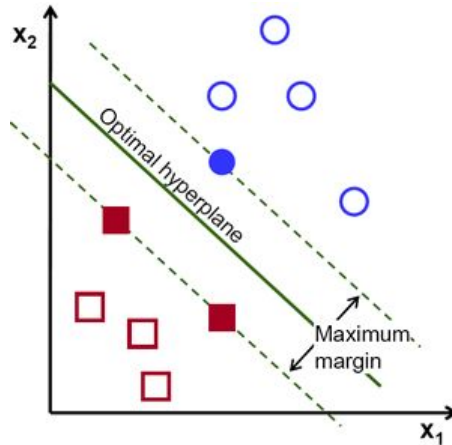


Cuál Hiperplano?



Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

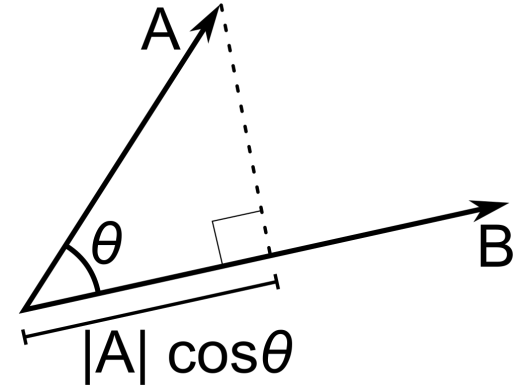
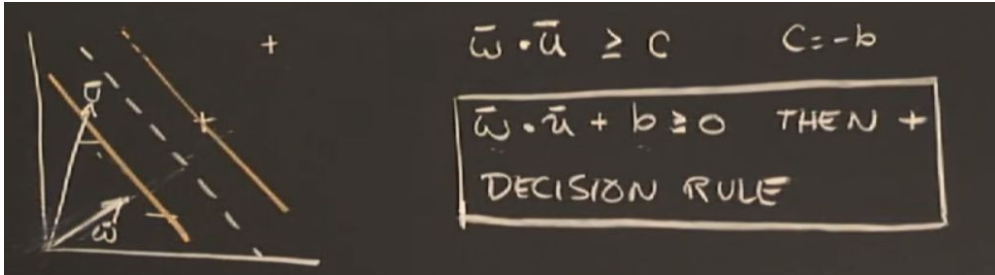


El que presente mayor margen entre los datos más cercanos .

Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

Que es un Hiperplano y cómo separa un plano en dos?

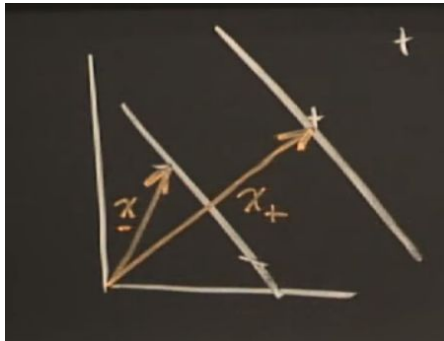




Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

Hiperplano límites (aceras)

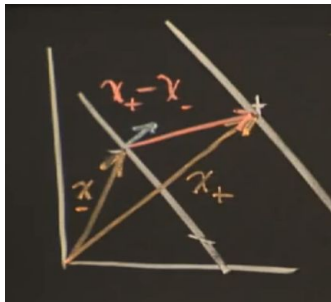


$$\begin{aligned}\bar{\omega} \cdot \bar{x}_+ + b &\geq 1 \\ \bar{\omega} \cdot \bar{x}_- + b &\leq -1\end{aligned}$$

Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

Distancia entre aceras



$$\bar{w} \cdot \bar{x}_+ + b \geq 1$$
$$\bar{w} \cdot \bar{x}_- + b \leq -1$$

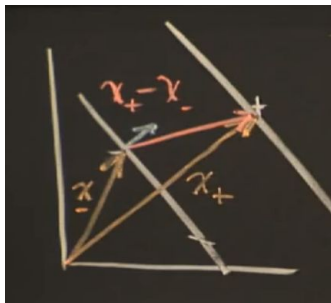
$$(\bar{x}_+ - \bar{x}_-) \cdot \frac{\bar{w}}{\|\bar{w}\|} = \frac{2}{\|\bar{w}\|}$$

WIDTH

Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

Distancia entre aceras



$$\begin{aligned}\bar{\omega} \cdot \bar{x}_+ + b &\geq 1 \\ \bar{\omega} \cdot \bar{x}_- + b &\leq -1\end{aligned}$$

$$(\bar{x}_+ - \bar{x}_-) \cdot \frac{\bar{\omega}}{\|\bar{\omega}\|} = \frac{2}{\|\bar{\omega}\|}$$

WIDTH

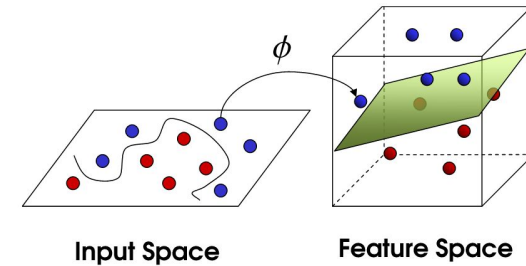
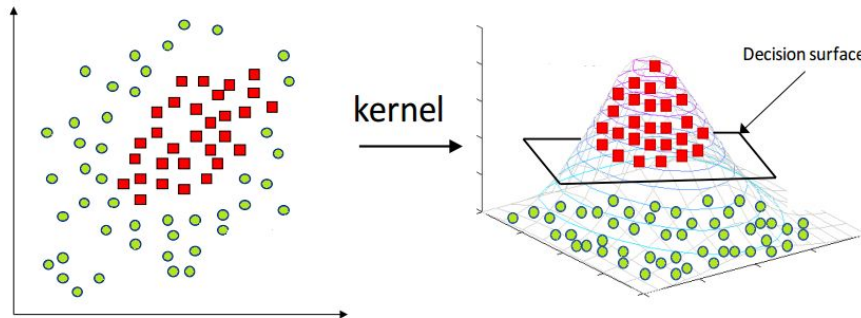
Maximizar



Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

Cuando los datos no son separables linealmente





Machine Learning

Aprendizaje supervisado - Support Vector Machine (SVM)

Kernels típicos:

- linear: $\langle x, x' \rangle$.
- polynomial: $(\gamma \langle x, x' \rangle + r)^d$, where d is specified by parameter `degree`, r by `coef0`.
- rbf: $\exp(-\gamma \|x - x'\|^2)$, where γ is specified by parameter `gamma`, must be greater than 0.
- sigmoid $\tanh(\gamma \langle x, x' \rangle + r)$, where r is specified by `coef0`.



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

Los árboles de decisión son un **método de aprendizaje supervisado no paramétrico** utilizado para la clasificación y la regresión.

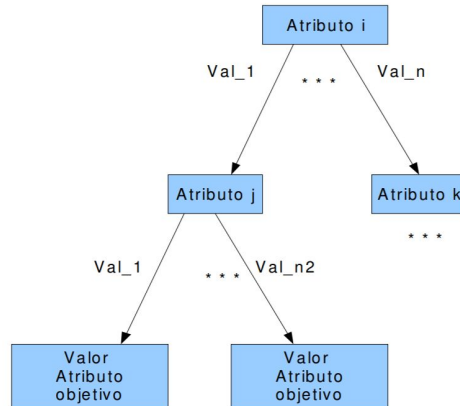
Su función es crear un modelo que **prediga** el valor de una variable objetivo mediante el **aprendizaje de reglas simples de decisión inferidas** a partir de las **características de los datos**.



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

Reglas -> Secuencia de condiciones -> Representan en un grafo



| | Atributo 1 | Atributo 2 | ... | Atributo n | Valor Objetivo |
|----------------------------|---------------|---------------|------------|---------------|-------------------|
| <i>Nombres Atributos</i> → | A | B | ... | N | CLASE |
| patrón 1 | a_1 | b_1 | ... | n_1 | Y_1 |
| patrón 2 | a_2 | b_2 | ... | n_2 | Y_2 |
| ... | ... | ... | ... | ... | ... |
| patrón m | a_m | b_m | ... | n_m | Y_m |

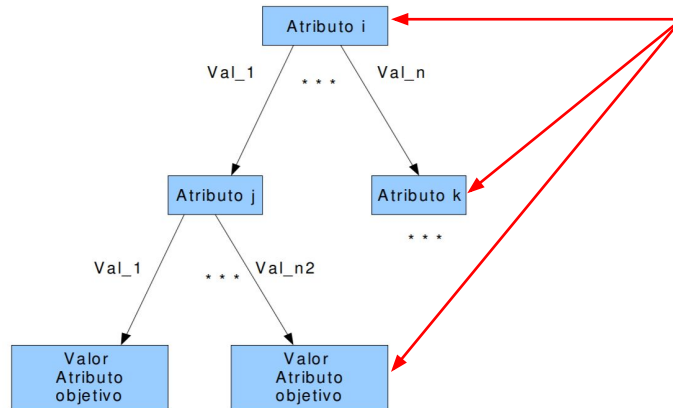
Conjunto de datos de entrenamiento



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

Reglas -> Secuencia de condiciones -> Representan en un grafo



Cómo seleccionar el orden de los atributos?

| | Atributo 1 | Atributo 2 | ... | Atributo n | Valor Objetivo |
|----------------------------|---------------|---------------|------------|---------------|-------------------|
| <i>Nombres Atributos</i> → | A | B | ... | N | CLASE |
| <i>patrón 1</i> | a_1 | b_1 | ... | n_1 | Y_1 |
| <i>patrón 2</i> | a_2 | b_2 | ... | n_2 | Y_2 |
| ... | ... | ... | ... | ... | ... |
| <i>patrón m</i> | a_m | b_m | ... | n_m | Y_m |

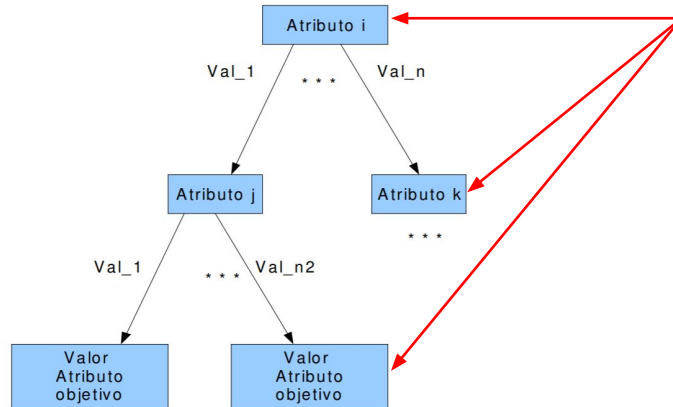
Conjunto de datos de entrenamiento



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

Reglas -> Secuencia de condiciones -> Representan en un grafo



Cómo seleccionar el orden de los atributos?

Teoría de la información

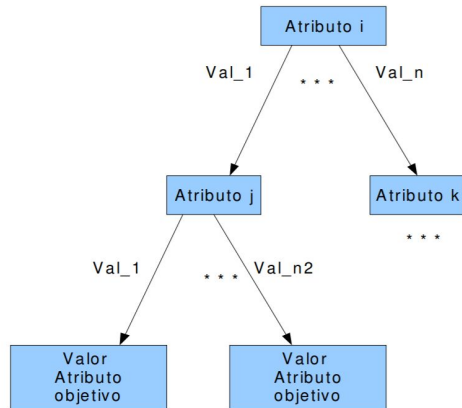
| | Atributo 1 | Atributo 2 | ... | Atributo n | Valor Objetivo |
|---------------------|---------------|---------------|-----|---------------|-------------------|
| Nombres Atributos → | A | B | ... | N | CLASE |
| patrón 1 | a_1 | b_1 | ... | n_1 | Y_1 |
| patrón 2 | a_2 | b_2 | ... | n_2 | Y_2 |
| ... | ... | ... | ... | ... | ... |
| patrón m | a_m | b_m | ... | n_m | Y_m |

Conjunto de datos de entrenamiento



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)



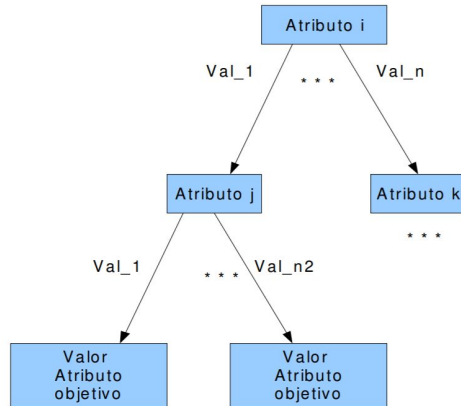
Teoría de la información

Cantidad de información -> Entropía -> Ganancia de información



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)



Teoría de la información

Cantidad de información -> Entropía -> Ganancia de información

$$\log_2 \left(\frac{1}{p(i)} \right)$$

$$\sum_{i=1}^n p(i) \log_2 \left(\frac{1}{p(i)} \right)$$

$$Entropía(S) - \sum_{v \in A} \frac{|S_v|}{|S|} Entropía(S_v)$$



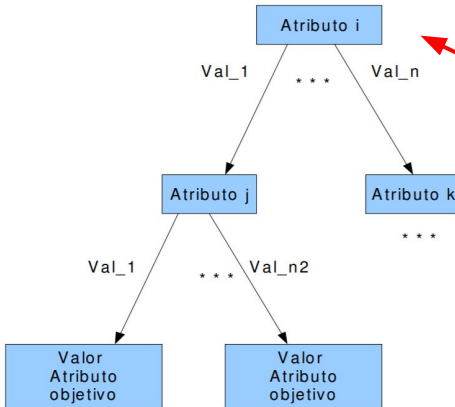
Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

Teoría de la información

El atributo que tenga mayor Ganancia de información

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|Sv|}{|S|} \text{Entropia}(Sv)$$





Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

| Variable | Ganancia de información (Variable) |
|--------------------------|---|
| Puntualidad y asistencia | $1.56 - [4/7(1.5) + 3/7(1.58)] = 0.025$ |
| Participación | $1.56 - [2/7(1.0) + 3/7(1.58) + 2/7(1.0)] = 0.311$ |
| Aprovechamiento | $1.56 - [1/7(0) + 2/7(1.0) + 3/7(0.92) + 1/7(0)] = 0.880$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|Sv|}{|S|} \text{Entropia}(Sv)$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

| Variable | Ganancia de información (Variable) |
|--------------------------|---|
| Puntualidad y asistencia | $1.56 - [4/7(1.5) + 3/7(1.58)] = 0.025$ |
| Participación | $1.56 - [2/7(1.0) + 3/7(1.58) + 2/7(1.0)] = 0.311$ |
| Aprovechamiento | $1.56 - [1/7(0) + 2/7(1.0) + 3/7(0.92) + 1/7(0)] = 0.880$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

$$\text{Clase} \begin{cases} \text{Exento} & = 2/7 \\ \text{Final} & = 3/7 \\ \text{Extraordinario} & = 2/7 \end{cases}$$

$$I(2/7, 3/7, 2/7) = -2/7 \log_2 2/7 - 3/7 \log_2 3/7 - 2/7 \log_2 2/7 = 1.56$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

| Variable | Ganancia de información (Variable) |
|--------------------------|---|
| Puntualidad y asistencia | $1.56 - [4/7(1.5) + 3/7(1.58)] = 0.025$ |
| Participación | $1.56 - [2/7(1.0) + 3/7(1.58) + 2/7(1.0)] = 0.311$ |
| Aprovechamiento | $1.56 - [1/7(0) + 2/7(1.0) + 3/7(0.92) + 1/7(0)] = 0.880$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

• **Puntualidad y Asistencia** = {No asiste, Asiste}

PyA = No asiste

$$\text{No Asiste} \begin{cases} \text{Exento} & = 1/4 \\ \text{Final} & = 2/4 \\ \text{Extraordinario} & = 1/4 \end{cases}$$

$$I(\text{Py A} = \text{No Asiste}) = I(1/4, 2/4, 1/4)$$

$$= -1/4 \log_2 1/4 - 2/4 \log_2 2/4 - 1/4 \log_2 1/4 = 1.50$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

| Variable | Ganancia de información (Variable) |
|--------------------------|---|
| Puntualidad y asistencia | $1.56 - [4/7(1.5) + 3/7(1.58)] = 0.025$ |
| Participación | $1.56 - [2/7(1.0) + 3/7(1.58) + 2/7(1.0)] = 0.311$ |
| Aprovechamiento | $1.56 - [1/7(0) + 2/7(1.0) + 3/7(0.92) + 1/7(0)] = 0.880$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

• Puntualidad y Asistencia = {No asiste, Asiste}

PyA = Asiste

$$\text{Asiste} \begin{cases} \text{Exento} & = 1/3 \\ \text{Final} & = 1/3 \\ \text{Extraordinario} & = 1/3 \end{cases}$$

$$I(\text{PyA} = \text{Asiste}) = I(1/3, 1/3, 1/3) \\ = -1/3 \log_2 1/3 - 1/3 \log_2 1/3 - 1/3 \log_2 1/3 = 1.58$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

| Variable | Ganancia de información (Variable) |
|--------------------------|---|
| Puntualidad y asistencia | $1.56 - [4/7(1.5) + 3/7(1.58)] = 0.025$ |
| Participación | $1.56 - [2/7(1.0) + 3/7(1.58) + 2/7(1.0)] = 0.311$ |
| Aprovechamiento | $1.56 - [1/7(0) + 2/7(1.0) + 3/7(0.92) + 1/7(0)] = 0.880$ |





Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)



| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

Entonces para "Aprovechamiento = Bueno" se tiene:

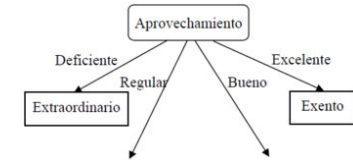
| Variable | Ganancia de información (Variable) |
|--------------------------|--|
| Puntualidad y asistencia | $0.92 - [2/3(0) + 1/3(0)] = 0.92$ |
| Participación | $0.92 - [1/3(0) + 1/3(0) + 1/3(0)] = 0.92$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|Sv|}{|S|} \text{Entropia}(Sv)$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)



| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

Entonces para "Aprovechamiento = Bueno" se tiene:

| Variable | Ganancia de información (Variable) |
|--------------------------|--|
| Puntualidad y asistencia | $0.92 - [2/3(0) + 1/3(0)] = 0.92$ |
| Participación | $0.92 - [1/3(0) + 1/3(0) + 1/3(0)] = 0.92$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

• Información de la clase: $I(C) = 0.92$

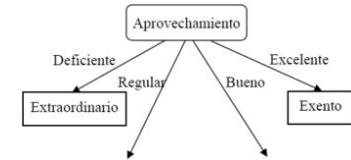
$$\text{Bueno} \begin{cases} \text{Exento} & = 1/3 \\ \text{Final} & = 2/3 \end{cases}$$

$$I(1/3, 2/3) = -1/3 \log_2 1/3 - 2/3 \log_2 2/3 = 0.92$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)



| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

Entonces para "Aprovechamiento = Bueno" se tiene:

| Variable | Ganancia de información (Variable) |
|--------------------------|--|
| Puntualidad y asistencia | $0.92 - [2/3(0) + 1/3(0)] = 0.92$ |
| Participación | $0.92 - [1/3(0) + 1/3(0) + 1/3(0)] = 0.92$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

• **Puntualidad y Asistencia** = {No asiste, Asiste}

PyA = No asiste

Py A = No asiste

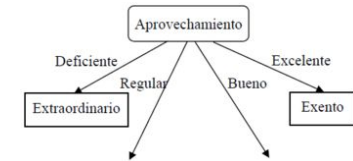
$$\text{No Asiste} = \begin{cases} \text{Exento} & = 0/2 \\ \text{Final} & = 2/2 \end{cases}$$

$$I(\text{Py A} = \text{No Asiste}) = I(0/2, 2/2) \\ = -0/2 \log_2 0/2 - 2/2 \log_2 2/2 = 0.0$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)



| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

Entonces para "Aprovechamiento = Bueno" se tiene:

| Variable | Ganancia de información (Variable) |
|--------------------------|--|
| Puntualidad y asistencia | $0.92 - [2/3(0) + 1/3(0)] = 0.92$ |
| Participación | $0.92 - [1/3(0) + 1/3(0) + 1/3(0)] = 0.92$ |

$$\text{Gan Inf}(S, A) = \text{Entropia}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

• **Puntualidad y Asistencia** = {No asiste, Asiste}

PyA = Asiste

Py A = Asiste

$$\text{Asiste} \begin{cases} \text{Exento} & = 1/1 \\ \text{Final} & = 0/1 \end{cases}$$

$$I(\text{Py A} = \text{Asiste}) = I(1/1, 0/1) \\ = 1/1 \log_2 1/1 - 0/1 \log_2 0/1 = 0.0$$



Machine Learning

Aprendizaje supervisado - Decision Trees (DTs)

| Alumno | ATRIBUTO | | | Nota |
|--------|----------------|---------------|-----------------|----------------|
| | Punt. y asist. | Participación | Aprovechamiento | |
| 1 | No asiste | Media | Excelente | Exento |
| 2 | Asiste | Alta | Bueno | Exento |
| 3 | No asiste | Media | Bueno | Final |
| 4 | No asiste | Baja | Bueno | Final |
| 5 | Asiste | Alta | Regular | Final |
| 6 | Asiste | Baja | Deficiente | Extraordinario |
| 7 | No asiste | Media | Regular | Extraordinario |

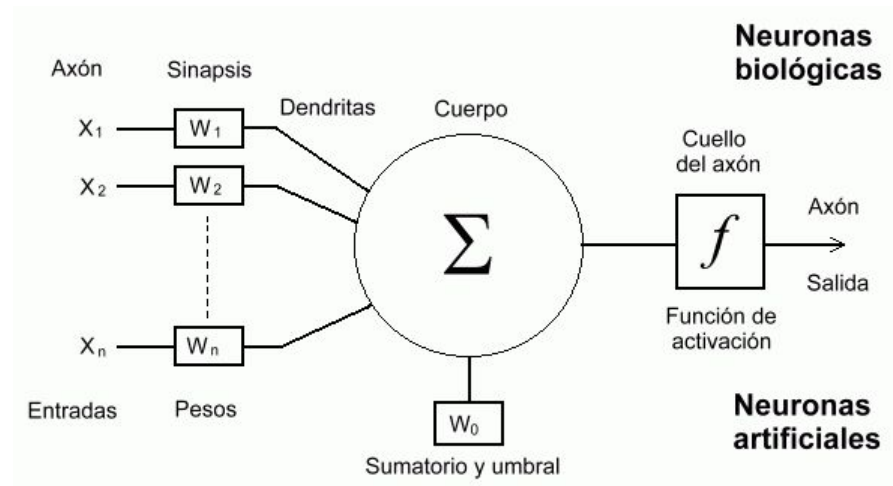
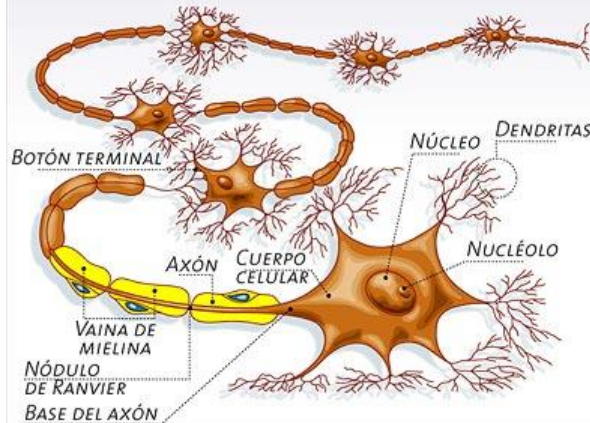


Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)

Neurona

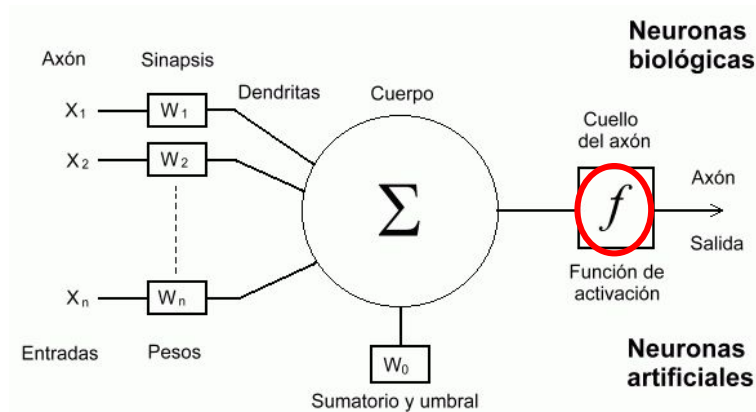
La neurona es la unidad funcional y estructural del sistema nervioso que produce y transmite el impulso nervioso. Se encuentra formada por tres partes: el **cuerpo neuronal** o **soma**; una prolongación larga y poco ramificada llamada **axón**, y otras prolongaciones muy ramificadas alrededor del soma llamadas **dendritas**.





Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)



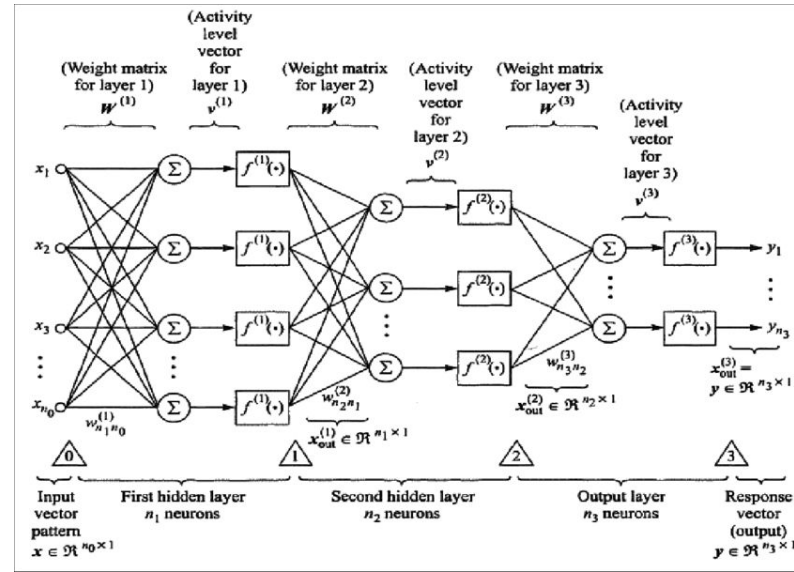
| | Función | Rango | Gráfica |
|------------|--|-----------------------------|---------|
| Identidad | $y = x$ | $[-\infty, +\infty]$ | |
| Escalón | $y = \text{sign}(x)$ $y = H(x)$ | $\{-1, +1\}$ $\{0, +1\}$ | |
| Sigmoidea | $y = \frac{1}{1 + e^{-x}}$ $y = \text{tanh}(x)$ | $[0, +1]$ $[-1, +1]$ | |
| Gaussiana | $y = Ae^{-Bx^2}$ | $[0, +1]$ | |
| Sinusoidal | $y = A \text{sen}(\omega x + \varphi)$ | $[-1, +1]$ | |



Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)

Multi-layer Perceptron



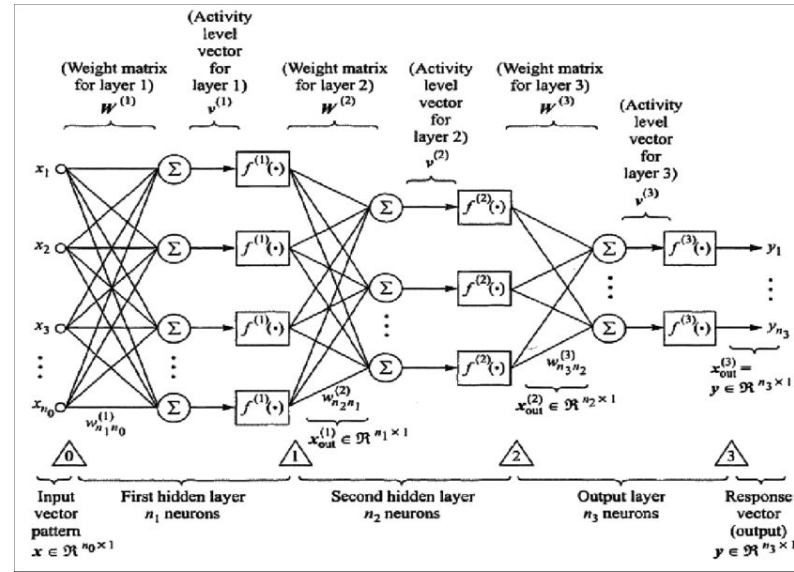


Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)

Multi-layer Perceptron

Cómo aprende?





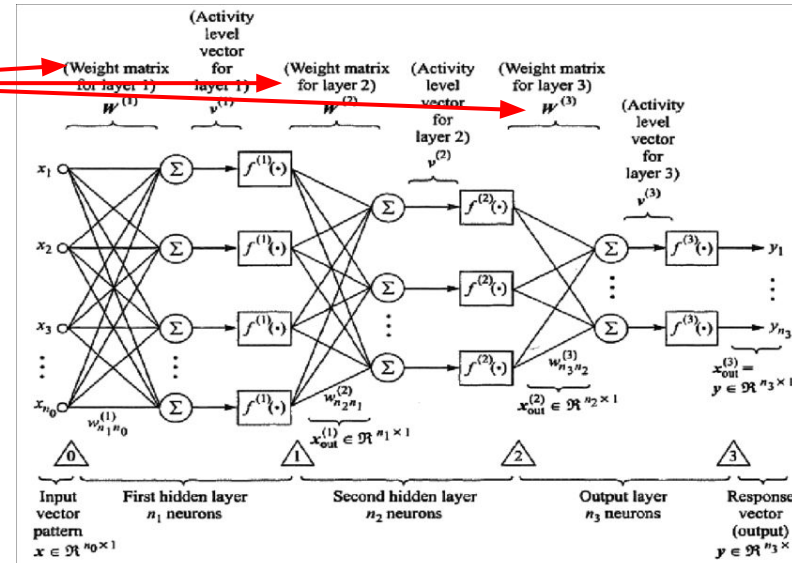
Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)

Multi-layer Perceptron

Cómo aprende?

Sinapsis: Se actualizan los pesos para minimizar el error entre la salida esperada y la salida actual.





Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)

Multi-layer Perceptron

Cómo aprende?

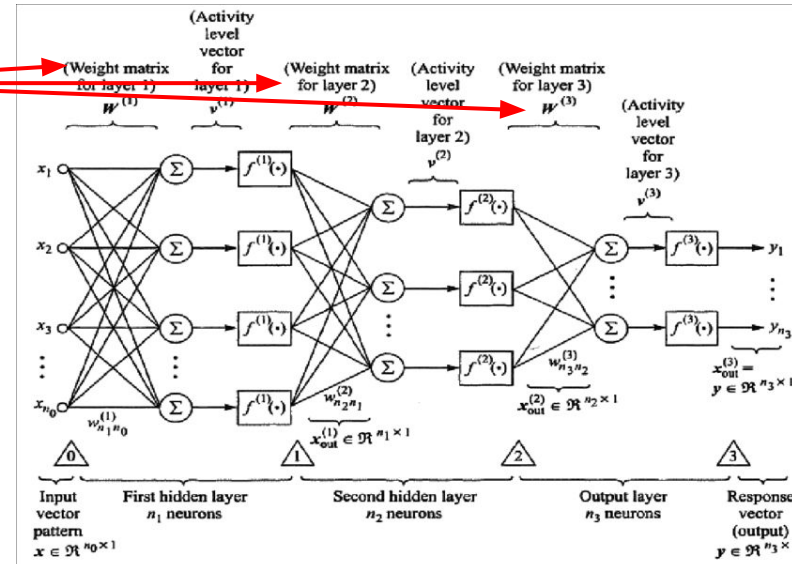
Sinapsis: Se actualizan los pesos para minimizar el error entre la salida esperada y la salida actual.

$$E(y, y') = \frac{1}{2} \|y - y'\|^2$$

$$E = \frac{1}{2n} \sum_x \|(y(x) - y'(x))\|^2$$

$$w \mapsto E(f_N(w, x_1), y_1)$$

$$\frac{\partial E}{\partial y'} = y' - y$$





Machine Learning

Aprendizaje supervisado – Artificial Neural Networks (ANN)

Multi-layer Perceptron

Cómo aprende?

Sinapsis: Se actualizan los pesos para minimizar el error entre la salida esperada y la salida actual.

$$E(y, y') = \frac{1}{2} \|y - y'\|^2$$

$$E = \frac{1}{2n} \sum_x \|(y(x) - y'(x))\|^2$$

$$w \mapsto E(f_N(w, x_1), y_1)$$

$$\frac{\partial E}{\partial y'} = y' - y$$

