

¿Qué es? Aprendizaje automático

Machine Learning



Dr. Edgar Altszyler

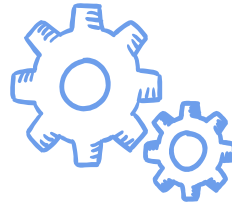
17/10/2018

Tarea:

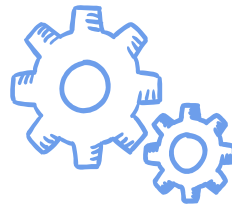
Dada la ***descripción*** de un vino,
¿puedo estimar su ***calidad***?



Queremos un programa que ...



Alta

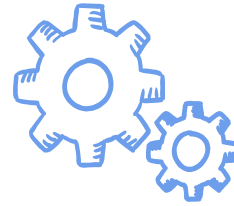


Baja

Queremos un programa que ...



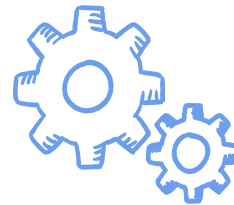
DESCRIPCIÓN



Alta



DESCRIPCIÓN

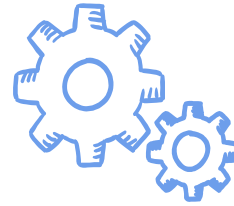


Baja

Queremos un programa que ...



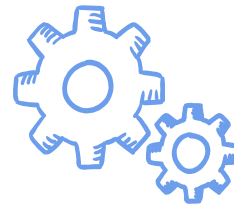
DESCRIPCIÓN	
Densidad	0.992
pH	3.5
Alcohol	12.2



Alta



DESCRIPCIÓN	
Densidad	0.997
pH	2.9
Alcohol	10.5



Baja

Queremos un programa que ...



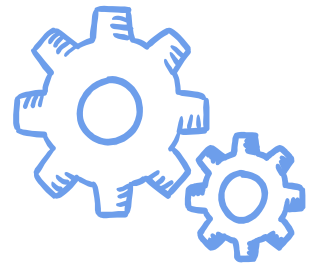
DESCRIPCIÓN	
Densidad	0.992
pH	3.5
Alcohol	12.2



DESCRIPCIÓN	
Densidad	0.997
pH	2.9
Alcohol	10.5



Sistemas expertos

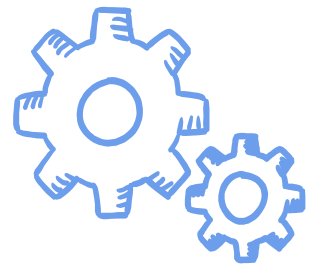


Clasificador

Experta/o del área



Sistemas expertos

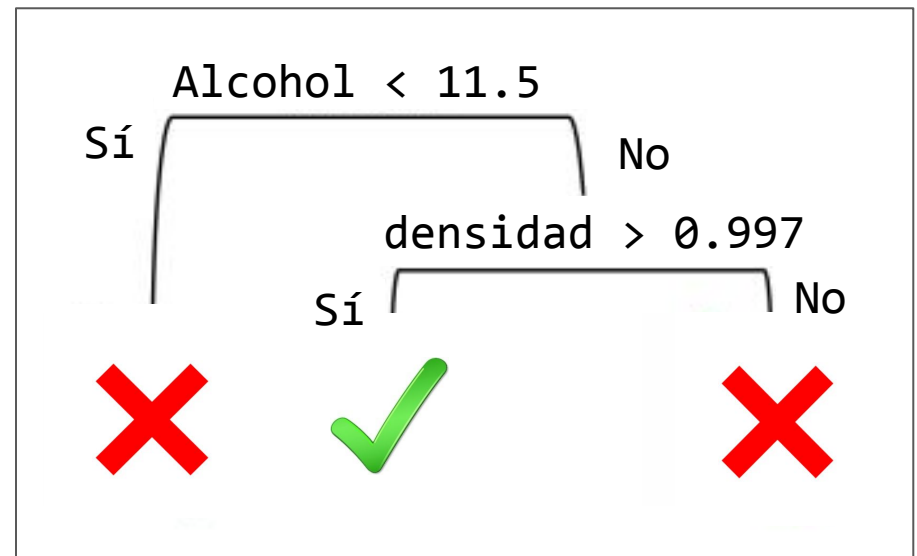


Clasificador

Experta/o del área



Reglas



Machine Learning

Aprender a partir de **ejemplos y experiencia ...**

Machine Learning

Primer paso

Juntar datos

Machine Learning

Primer paso

Juntar datos



Segundo paso

Entrenar un
clasificador

Machine Learning

Primer paso

Juntar datos

Segundo paso

Entrenar un
clasificador

Tercer paso

Hacer
predicciones



Machine Learning

Primer paso

Juntar datos

Segundo paso

Entrenar un
clasificador

Tercer paso

Hacer
predicciones



Recolección de los Datos



Recolección de los Datos



pH	densidad	Alcohol	Calidad
3.4	0.992	9.2	Alta
3.1	0.997	12.5	Alta
3.2	0.999	13.1	Baja
3.7	0.991	10.8	Alta
3.2	1.005	11.3	Baja
...

Recolección de los Datos



Instancias
(ejemplos)

Features			Etiquetas
pH	densidad	Alcohol	Calidad
3.4	0.992	9.2	Alta
3.1	0.997	12.5	Alta
3.2	0.999	13.1	Baja
3.7	0.991	10.8	Alta
3.2	1.005	11.3	Baja
...

Machine Learning

Primer paso

Juntar datos



Segundo paso

Entrenar un
clasificador



Tercer paso

Hacer
predicciones

Entrenar un clasificador

Un **clasificador** puede ser pensado como una caja de reglas...

Entrenar un clasificador

Un **clasificador** puede ser pensado como una caja de reglas...

Vacía al principio



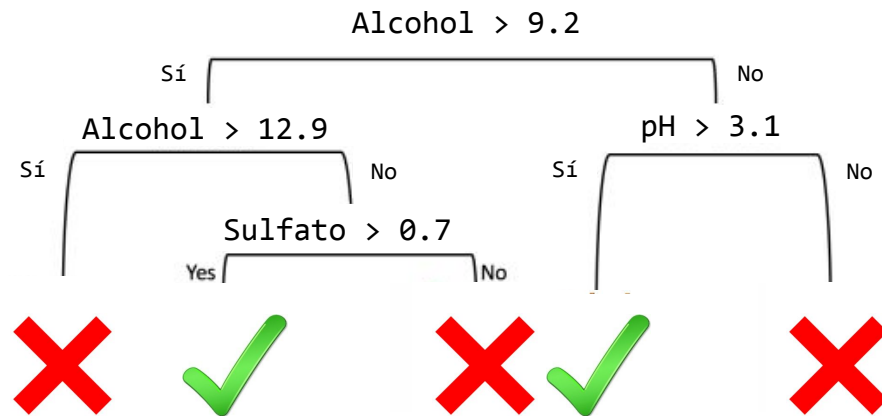
pH	dens.	Alcohol	Calidad
3.4	0.992	9.2	Alta
3.1	0.997	12.5	Alta
3.2	0.999	13.1	Baja
3.7	0.991	10.8	Alta
3.2	1.005	11.3	Baja
...



Lleno de reglas



Lleno de reglas



Machine Learning

Primer paso

Juntar datos

Segundo paso

Entrenar
clasificador

Tercer paso

Hacer
predicciones



Hacer Predicciones

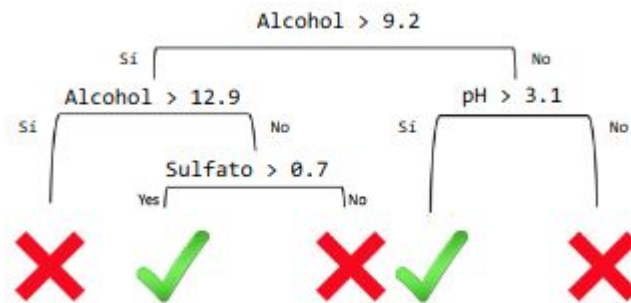


Hacer Predicciones

DESCRIPCIÓN	
Densidad	0.992
pH	3.5
Alcohol	12.2

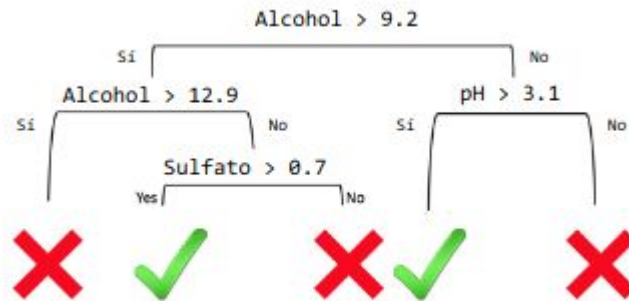
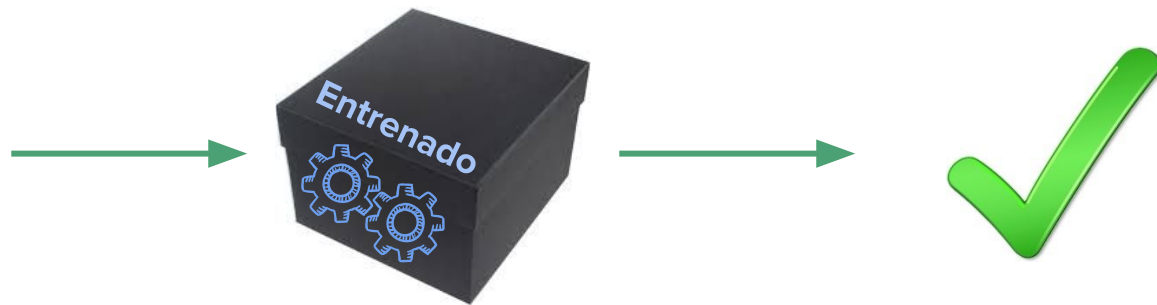
Hacer Predicciones

DESCRIPCIÓN	
Densidad	0.992
pH	3.5
Alcohol	12.2



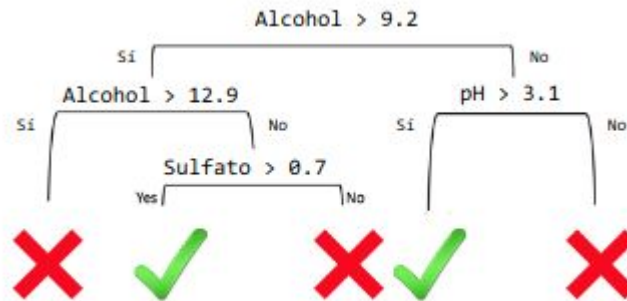
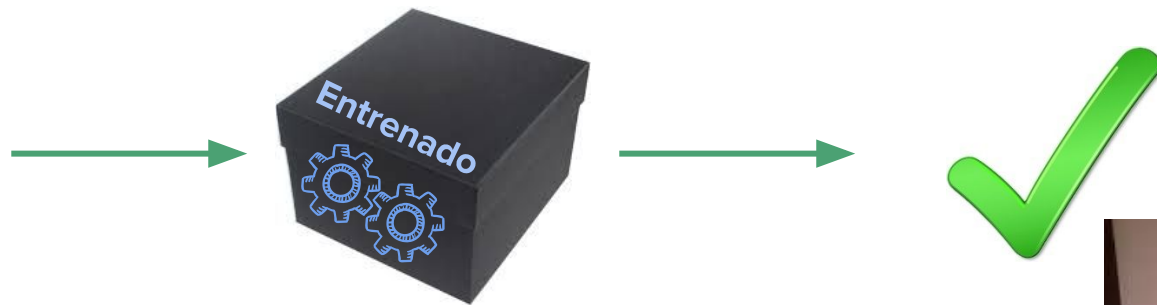
Hacer Predicciones

DESCRIPCIÓN	
Densidad	0.992
pH	3.5
Alcohol	12.2



Hacer Predicciones

DESCRIPCIÓN	
Densidad	0.992
pH	3.5
Alcohol	12.2



Machine Learning

¿Qué es lo que realmente hacen los clasificadores?

¿Qué diferencia hay entre los features de los vinos de calidad Alta y Baja ?

Volatile Acidity	1.1
sulphates	0.6
pH	3.0
...	...

Volatile Acidity	0.51
sulphates	0.45
pH	3.3
...	...

Volatile Acidity	0.85
sulphates	0.82
pH	3.1
...	...

Alta

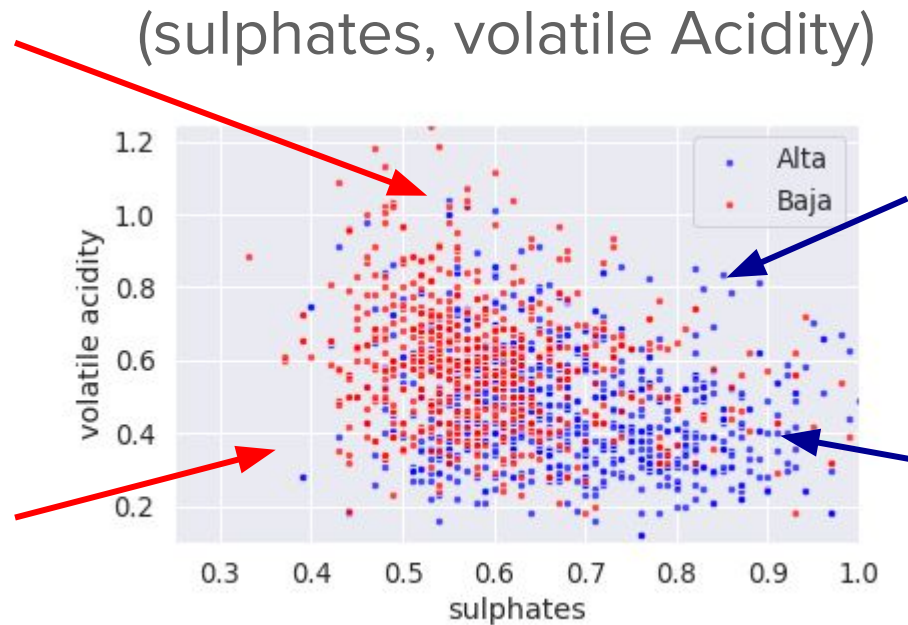
Volatile Acidity	0.53
sulphates	0.88
pH	3.7
...	...

Alta

¿Qué diferencia hay entre los features de los vinos de calidad Alta y Baja ?

Volatile Acidity	1.1
sulphates	0.6
pH	3.0
...	...

Volatile Acidity	0.51
sulphates	0.45
pH	3.3
...	...



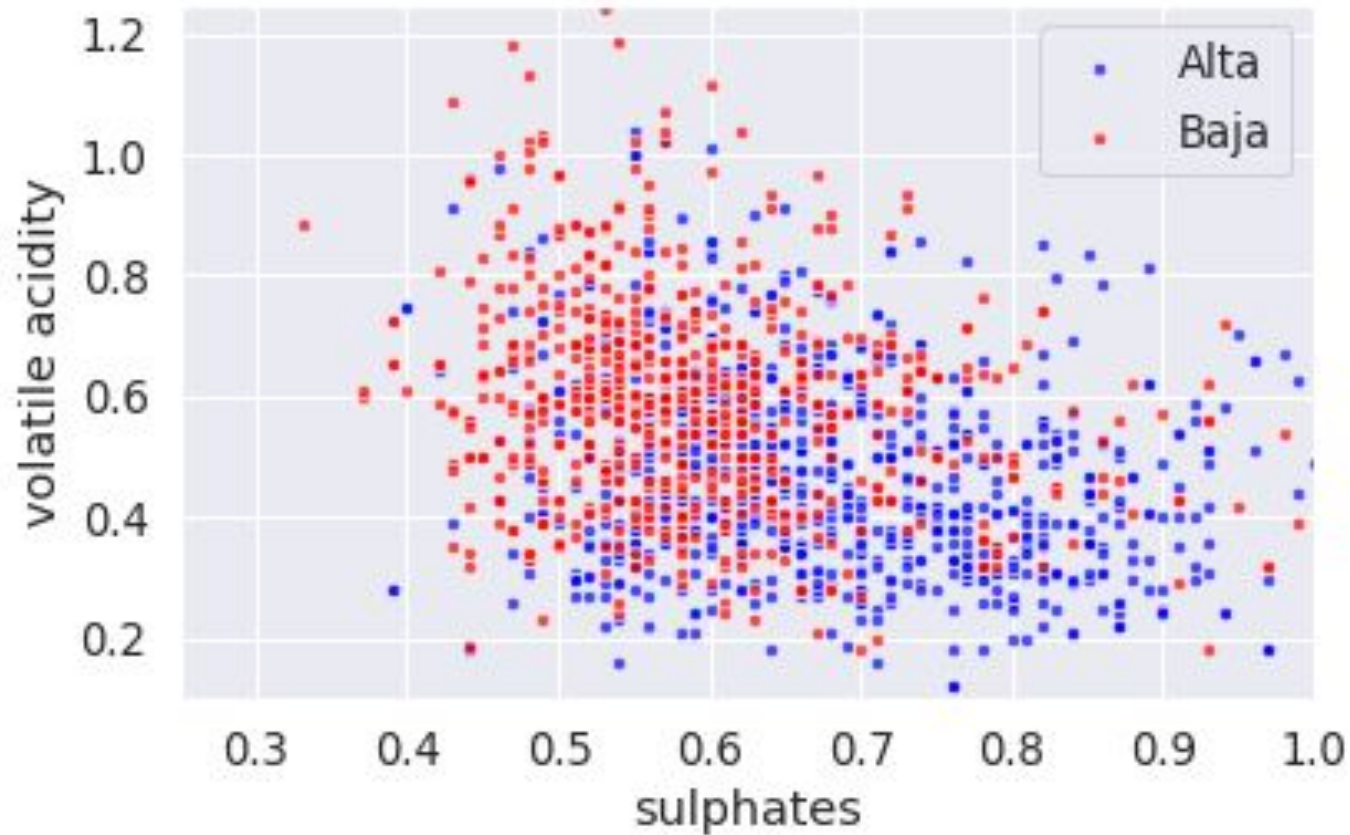
Volatile Acidity	0.85
sulphates	0.82
pH	3.1
...	...

Alta

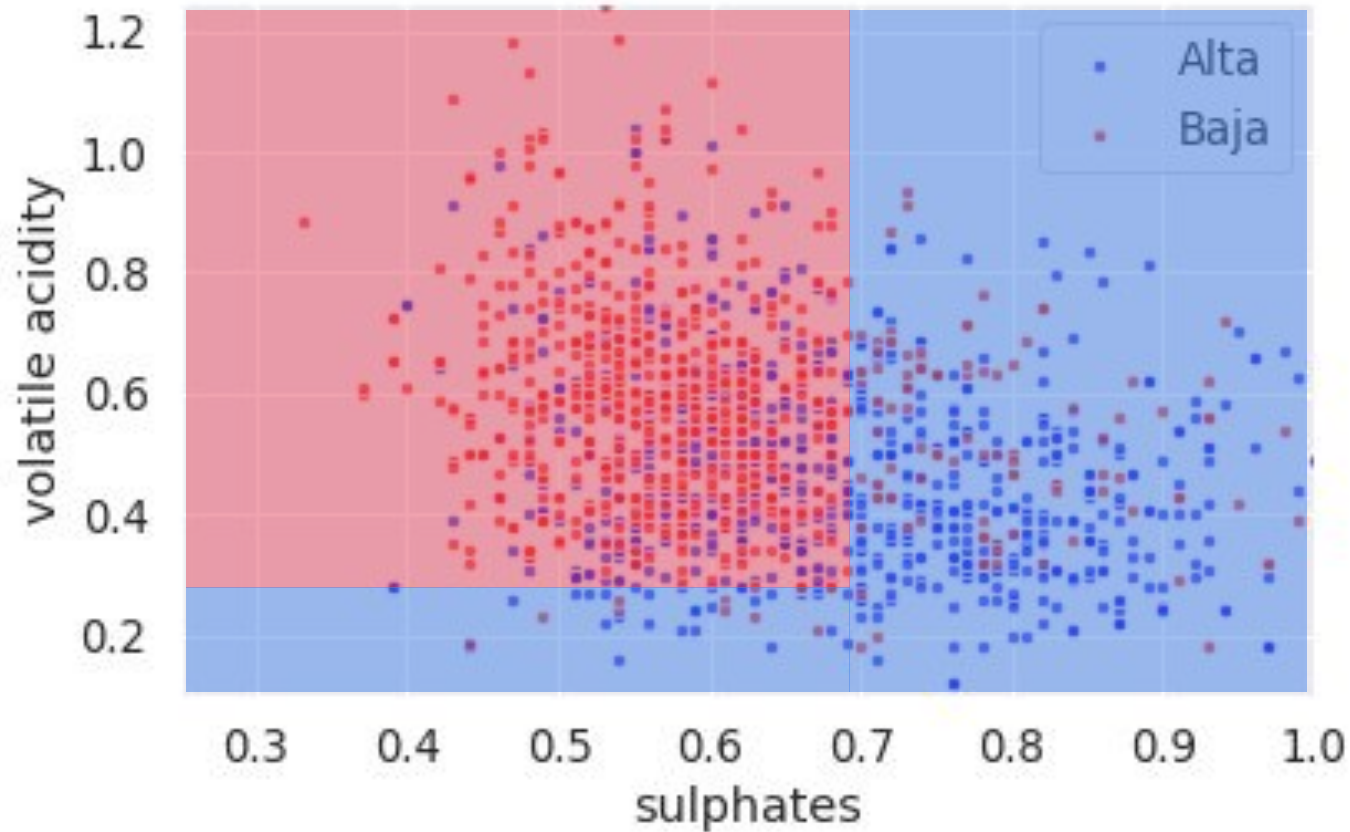
Volatile Acidity	0.53
sulphates	0.88
pH	3.7
...	...

Alta

¿Qué diferencia hay entre los features de los vinos de calidad Alta y Baja ?

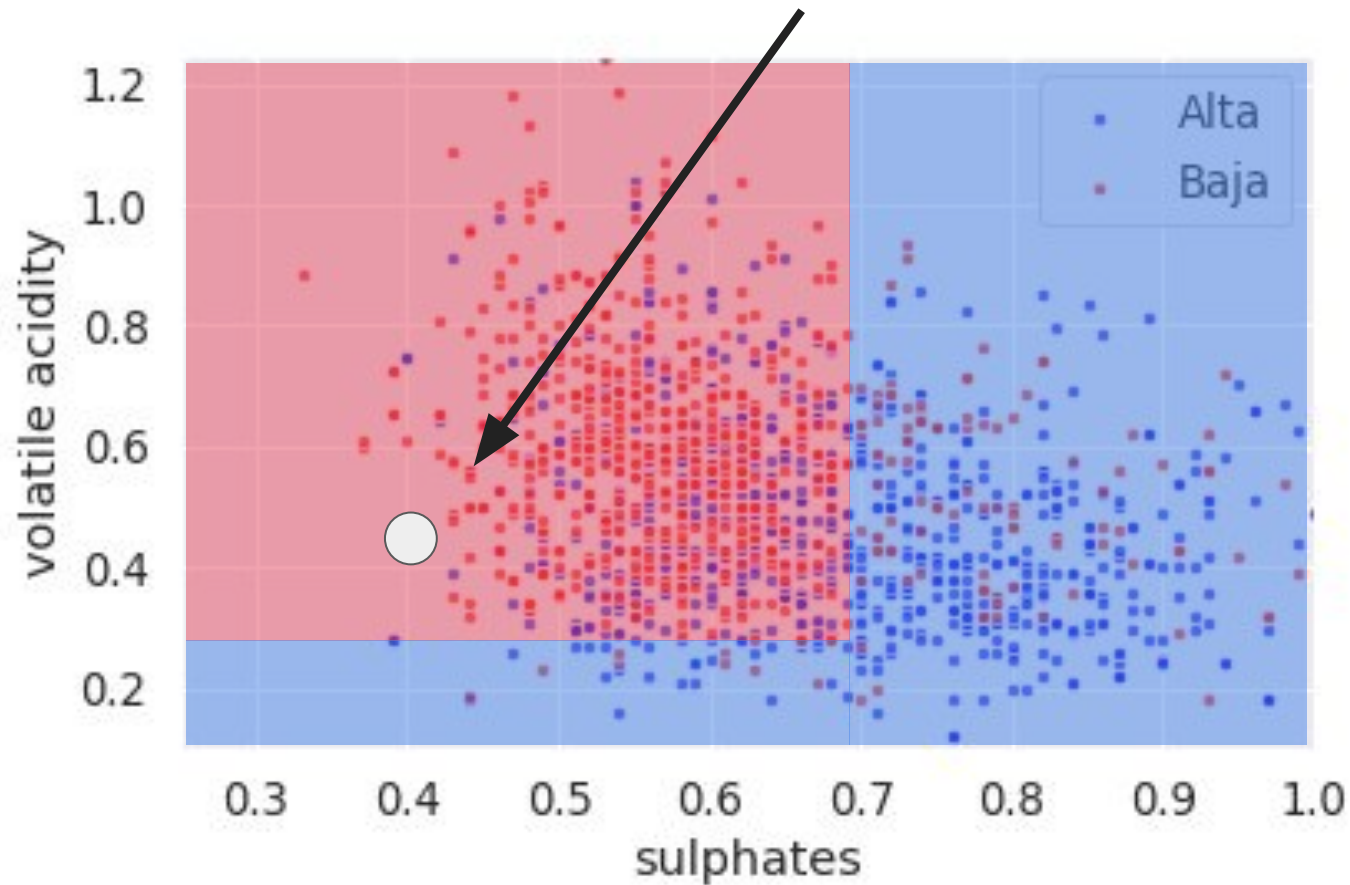


¿Qué diferencia hay entre los features de los vinos de calidad Alta y Baja ?



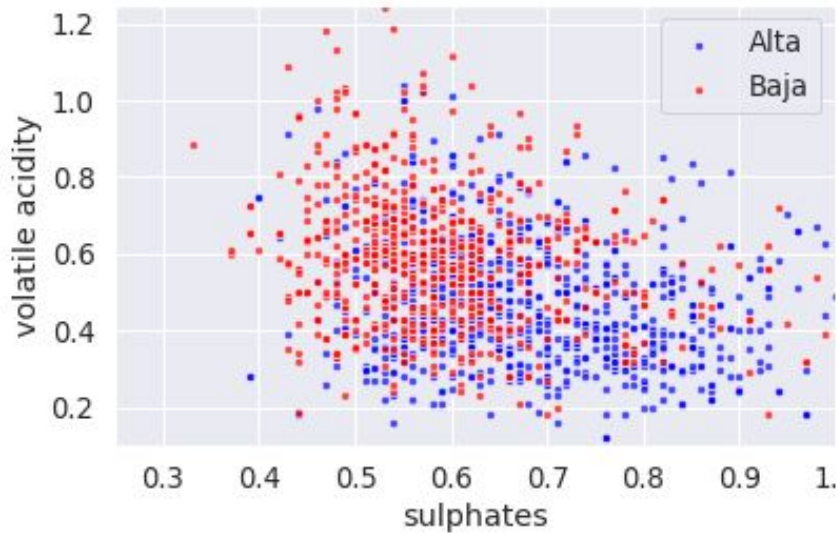
Vino nuevo

¿Calidad Alta o Baja?

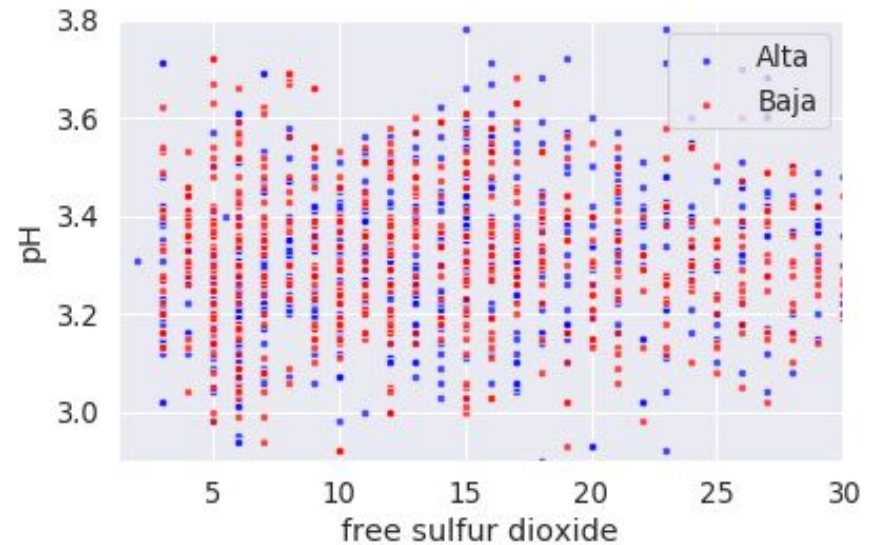


No todos los features son útiles

(sulphates , volatile acidity)



(free sulfur dioxide , pH)



Fin de Parte 1



Parte 2





Machine Learning



Machine Learning



Uso naive



**Phd en
Computación**



Machine Learning

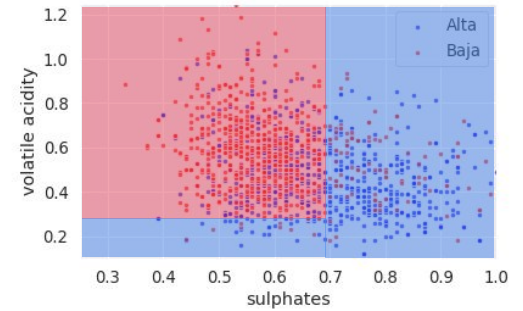
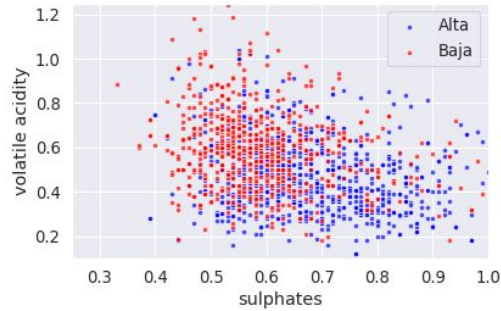


Capacidades y limitaciones

Vamos a Mordor

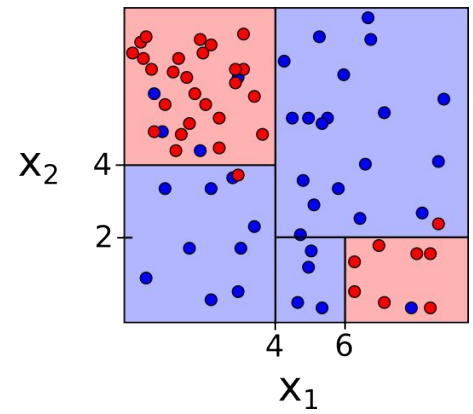
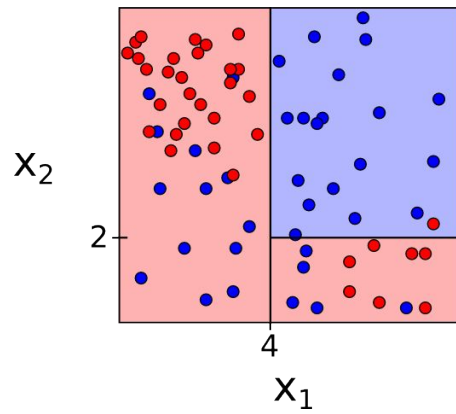
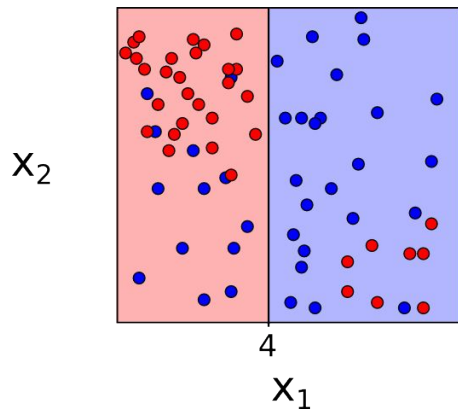


Algunos algoritmos de Machine Learning

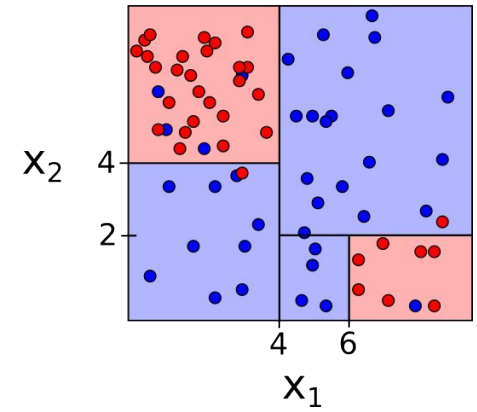
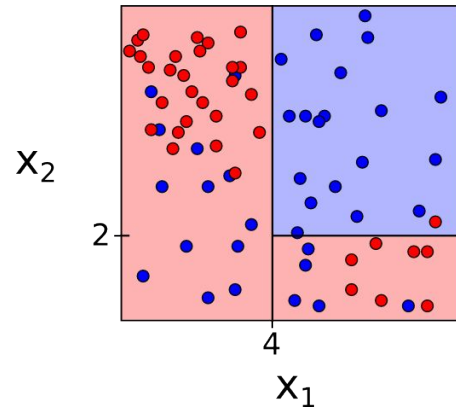
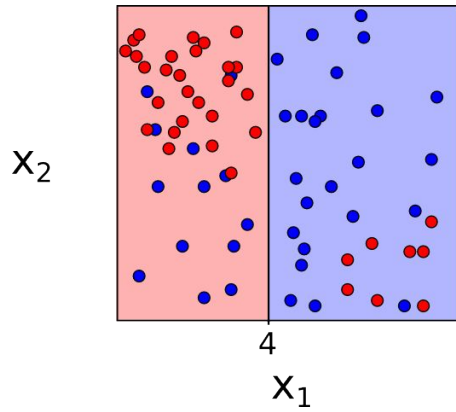


Machine Learning

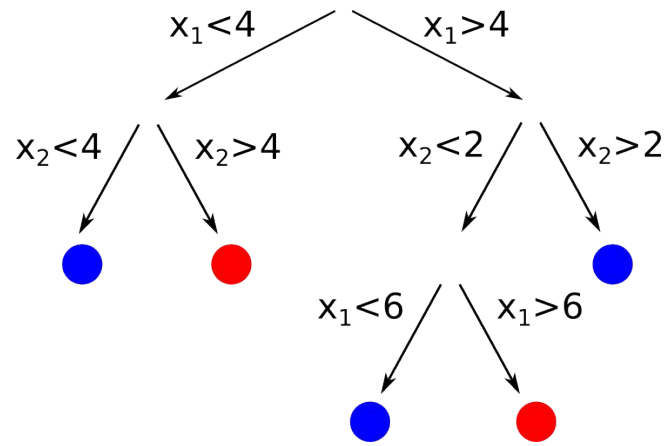
Árboles de decisión



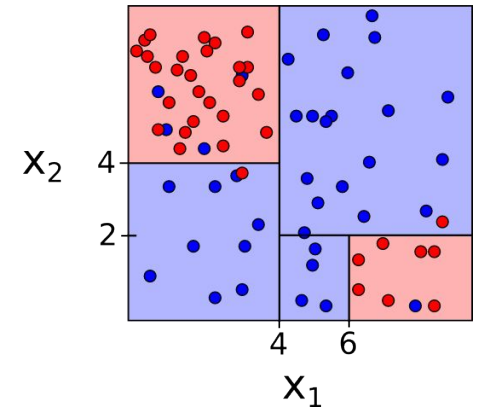
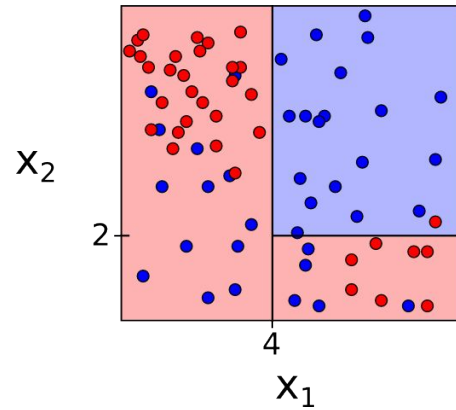
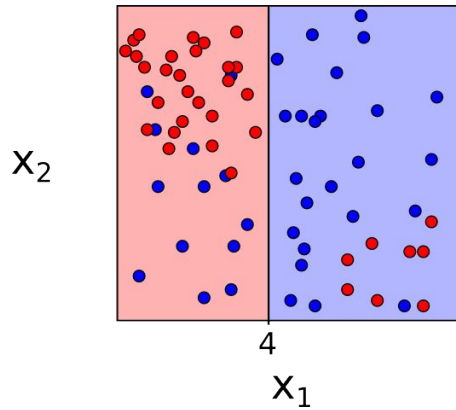
Árboles de decisión



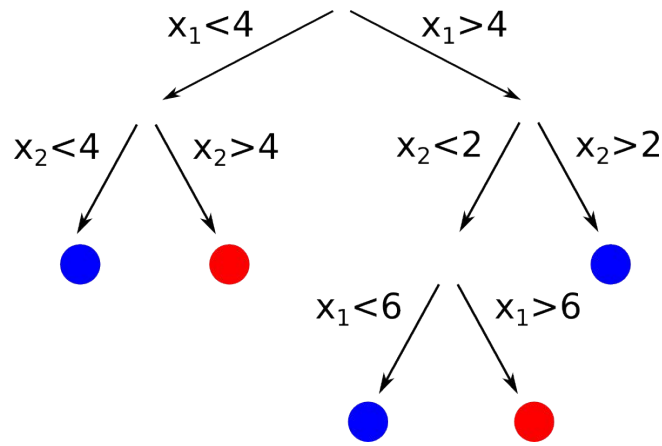
Árbol
entrenado



Árboles de decisión



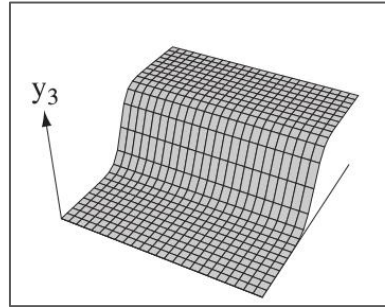
Árbol
entrenado



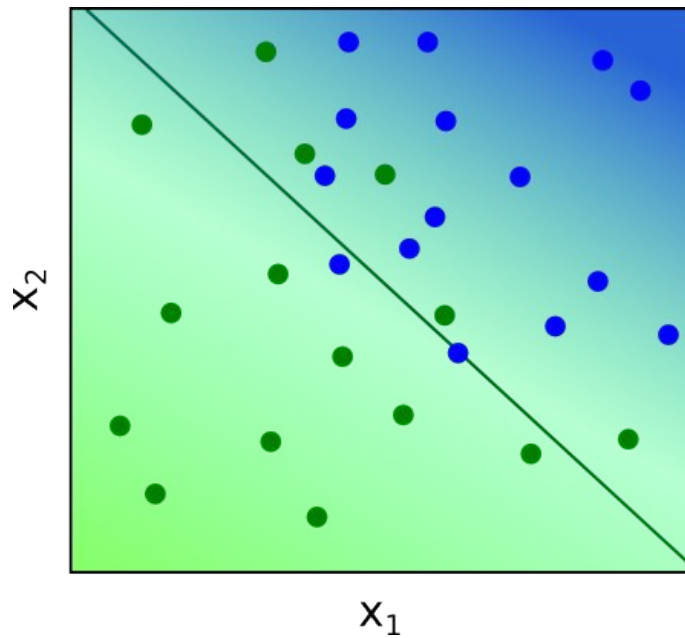
Predicción de
instancia nueva

$(5, 1) \longrightarrow ?$

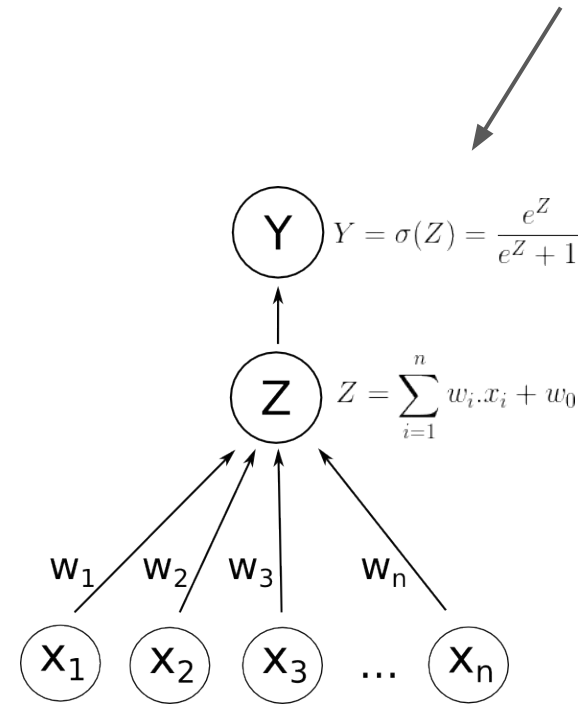
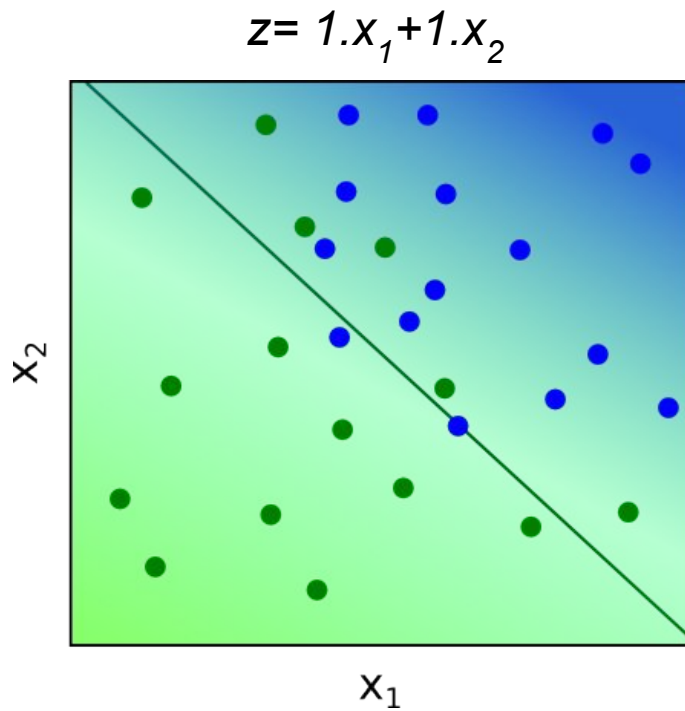
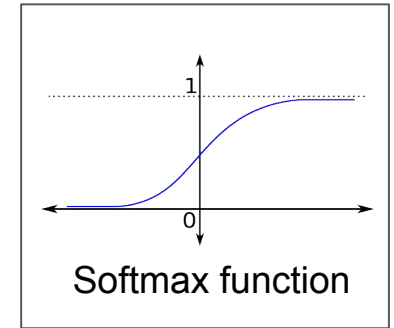
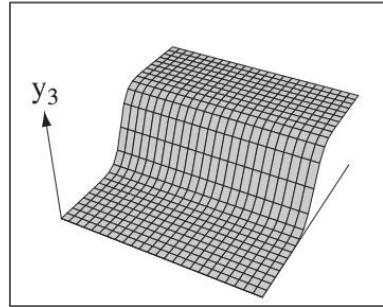
Redes Neuronales



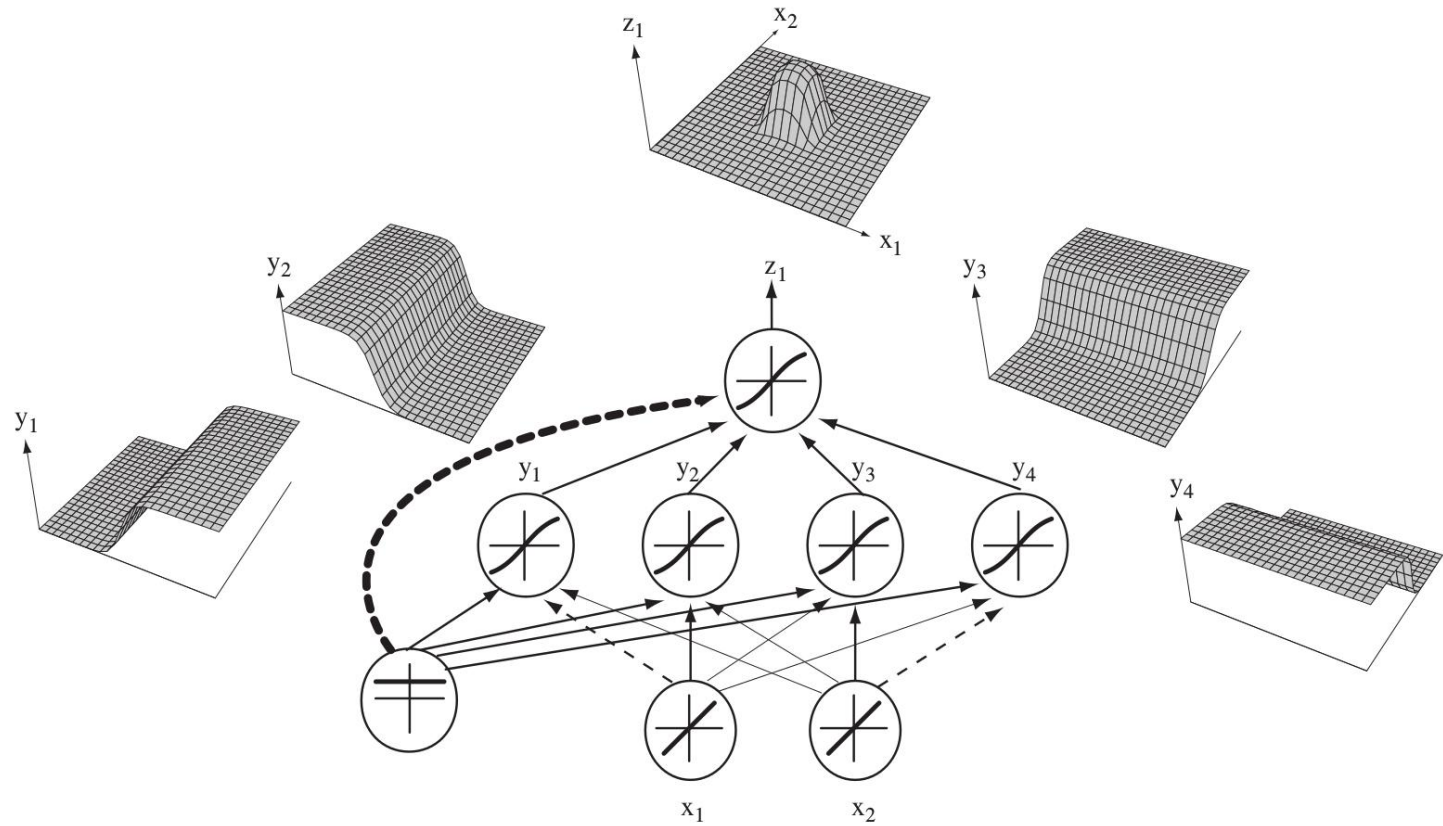
$$z = 1 \cdot x_1 + 1 \cdot x_2$$



Redes Neuronales



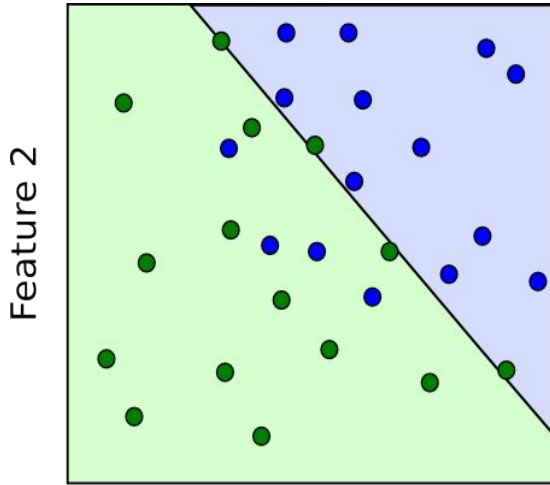
Redes Neuronales



Evaluación de los métodos

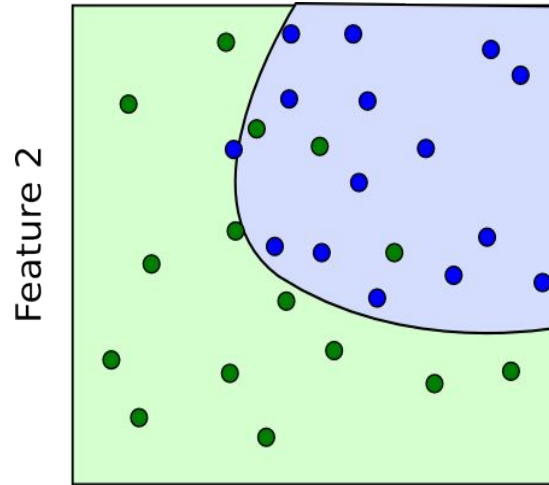
¿Qué clasificador uso?

Complejidad del modelo



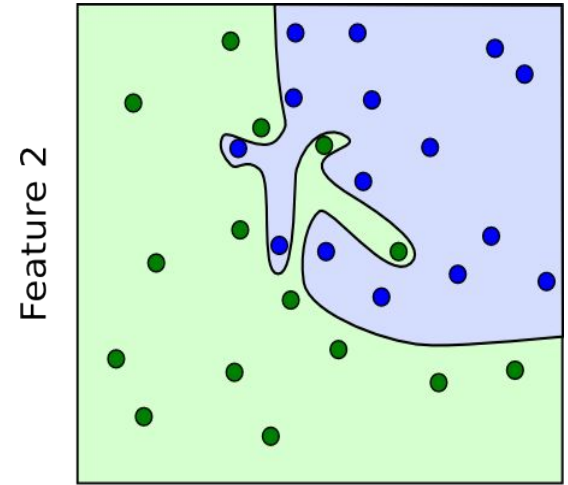
Feature 1

Accuracy = 80%
(24/30)



Feature 1

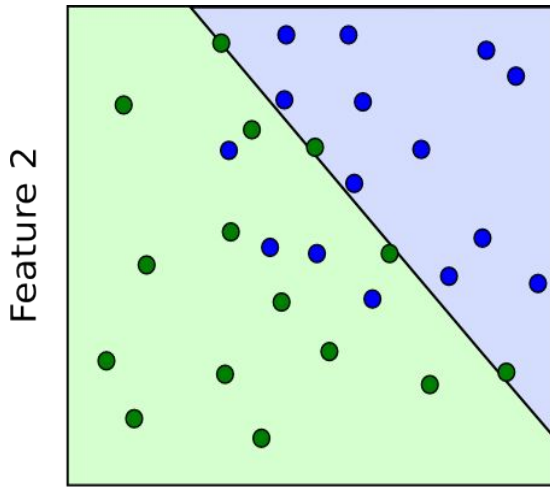
Accuracy = 87%
(26/30)



Feature 1

Accuracy = 100%
(30/30)

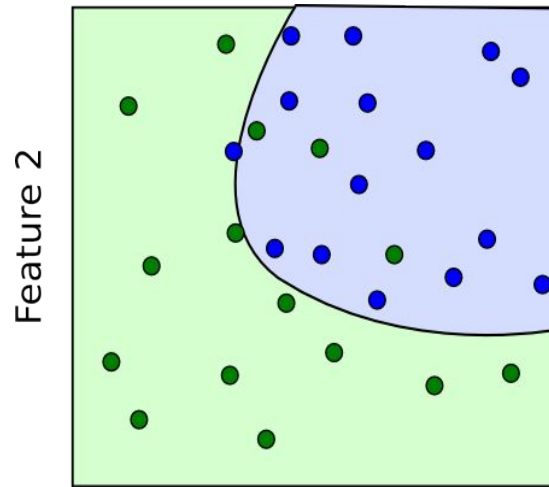
Complejidad del modelo



Feature 1

Accuracy = 80%
(24/30)

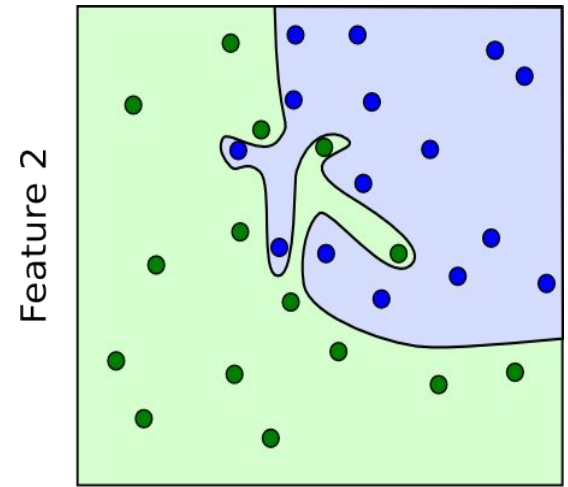
Under-fitting



Feature 1

Accuracy = 87%
(26/30)

Appropriate fitting



Feature 1

Accuracy = 100%
(30/30)

Over-fitting

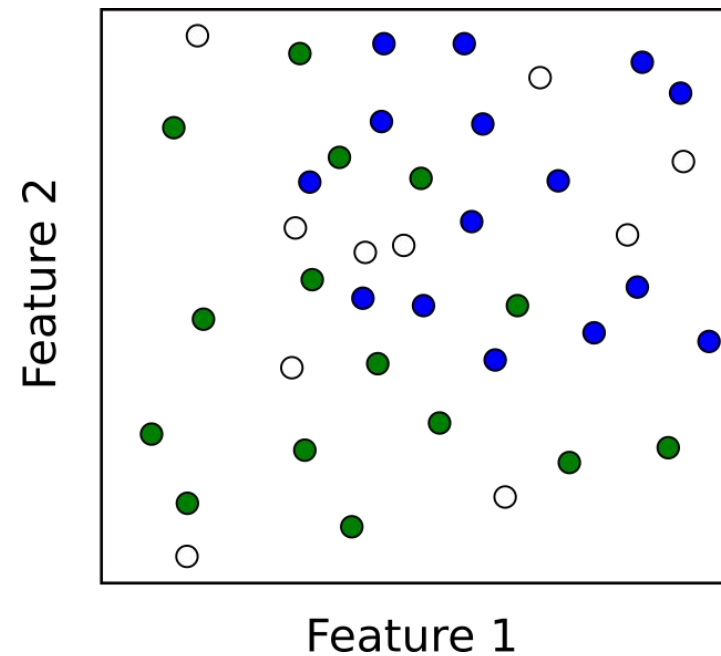
Estimar el performance en el dataset de entrenamiento es trampa!!!

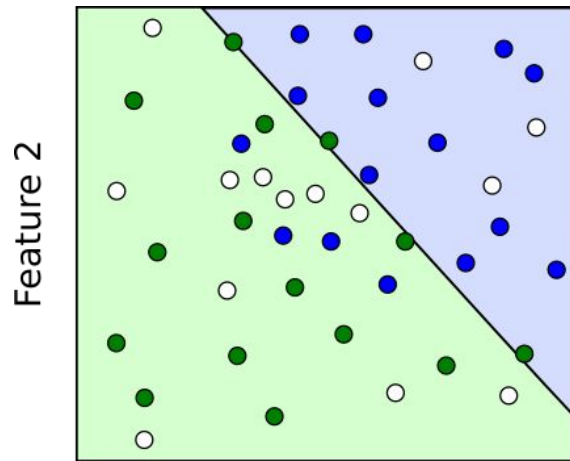
Training set
80%

pH	densidad	alcohol	calidad
3.4	0.990	102	Alta
3.1	0.989	11.5	Alta
3.2	1.003	12.2	Baja
3.7	0.998	10.3	Alta
3.2	1.004	11.9	Baja
...

Test set
20%

pH	densidad	alcohol	calidad
3.2	0.996	9.2	Alta
3.2	0.995	10.8	Alta
3.1	1.003	11.3	Baja
...

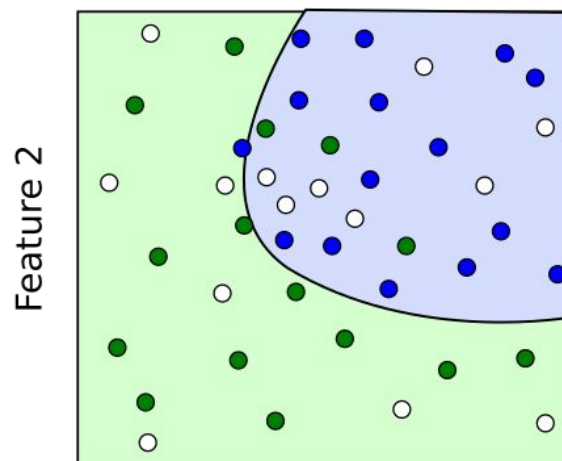




Feature 1

Accuracy = 60%
(12/20)

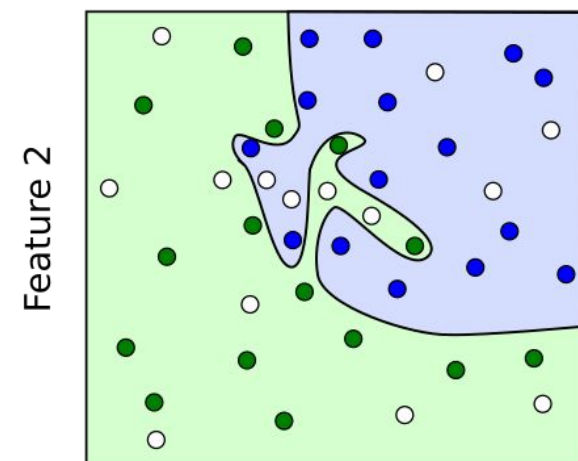
Under-fitting



Feature 1

Accuracy = 85%
(17/20)

Appropriate fitting

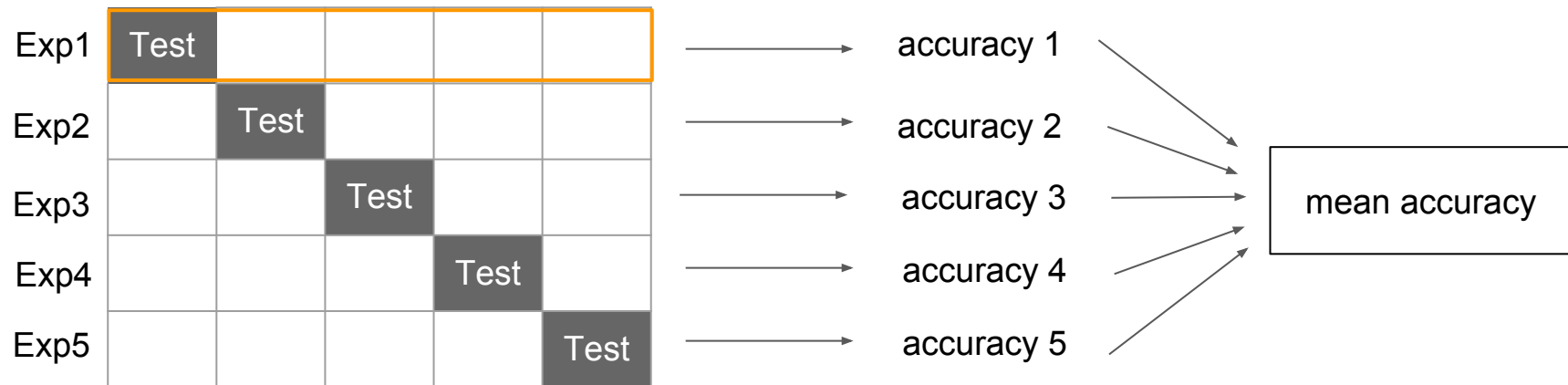


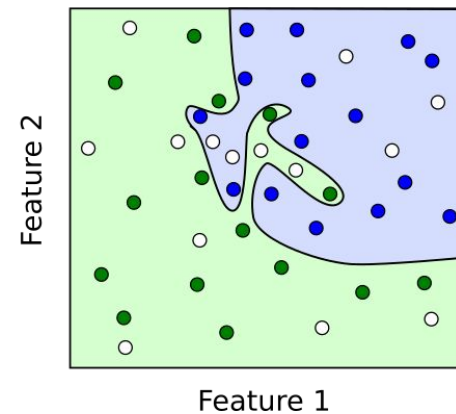
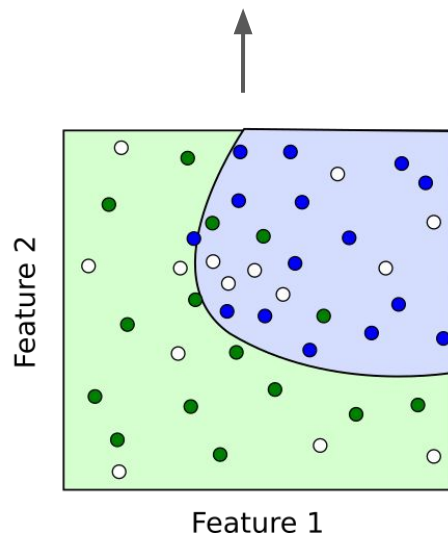
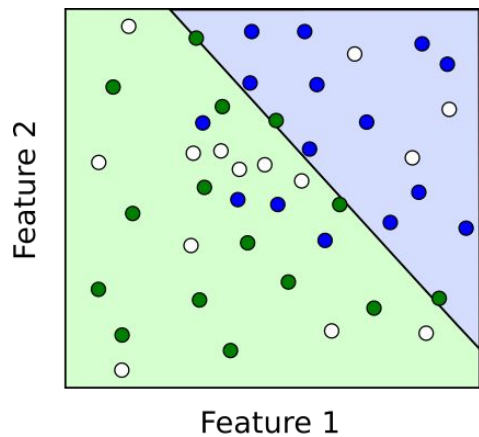
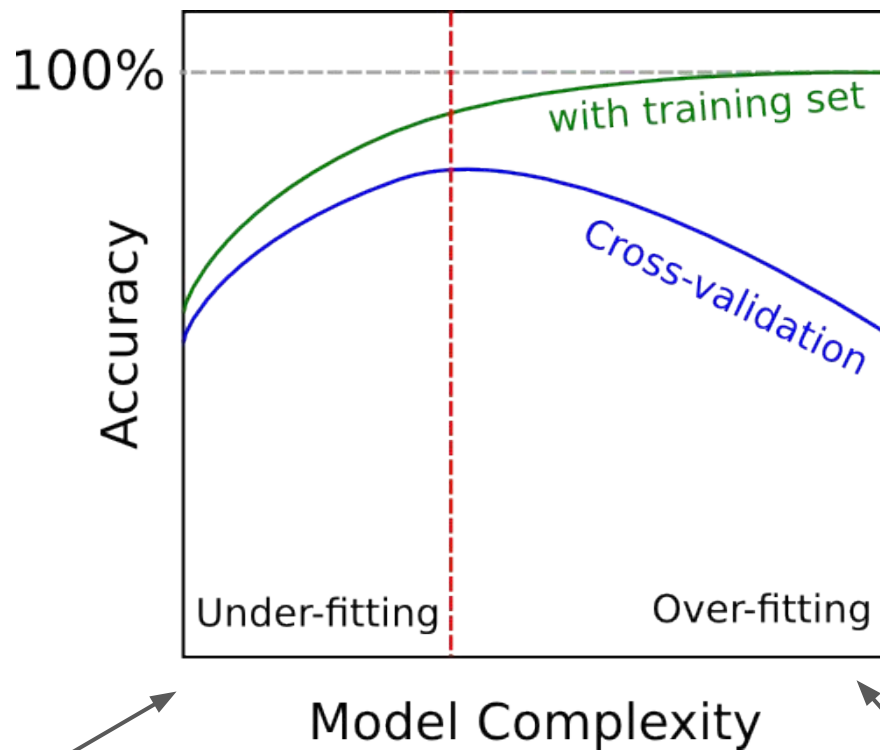
Feature 1

Accuracy = 75%
(15/20)

Over-fitting

5-Fold Cross-Validation





Pipeline general

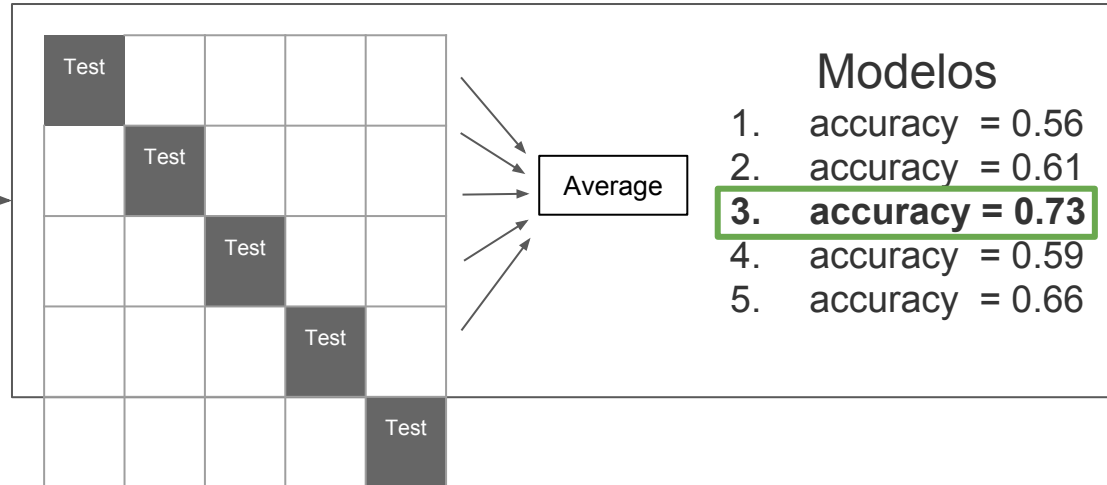
Training set
80%

pH	densidad	alcohol	calidad
3.4	0.990	102	Alta
3.1	0.989	11.5	Alta
3.2	1.003	12.2	Baja
3.7	0.998	10.3	Alta
3.2	1.004	11.9	Baja
...

Test set
20%

pH	densidad	alcohol	calidad
3.2	0.996	9.2	Alta
3.2	0.995	10.8	Alta
3.1	1.003	11.3	Baja
...

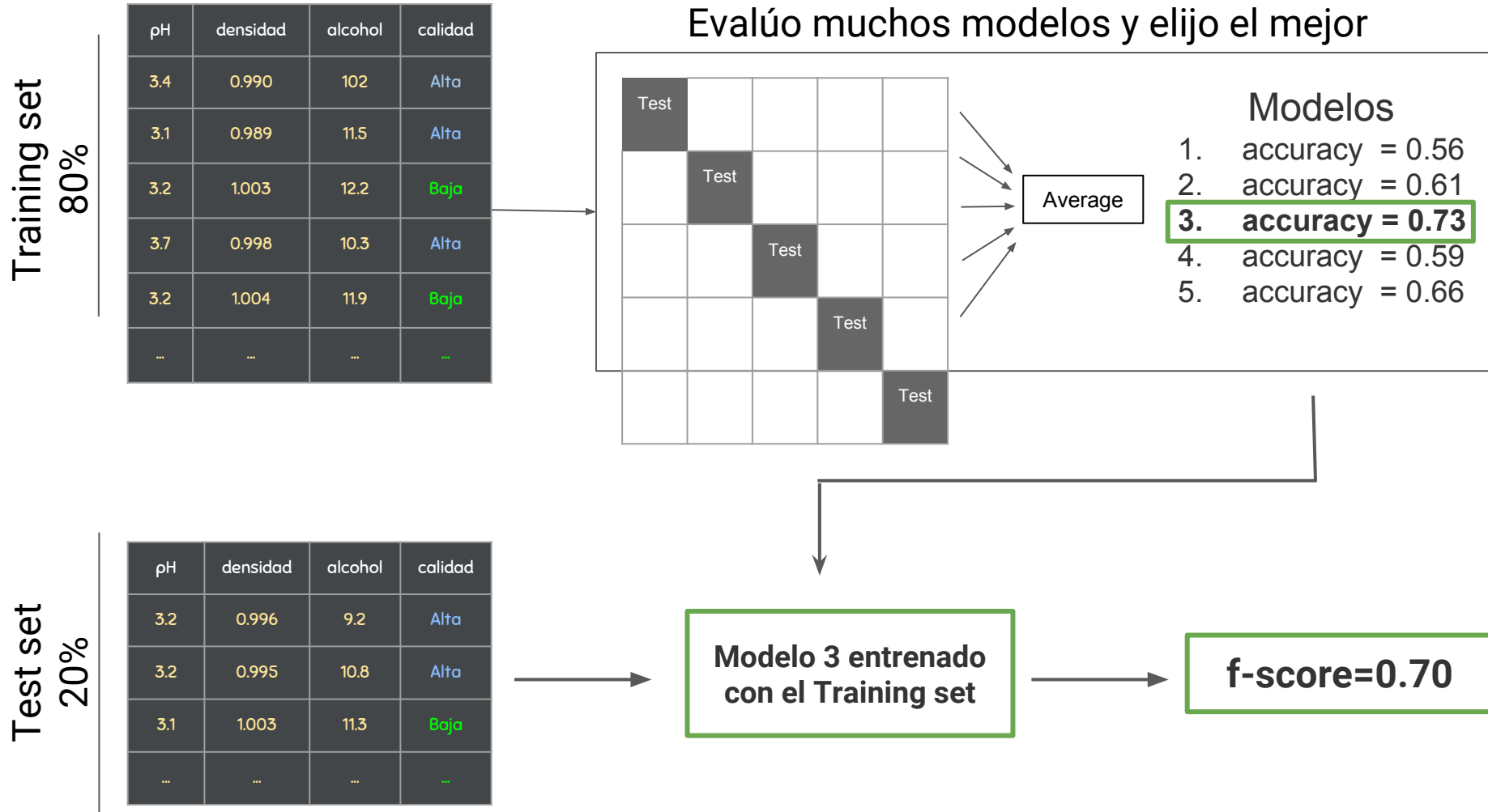
Evalúo muchos modelos y elijo el mejor



Modelos

1. accuracy = 0.56
2. accuracy = 0.61
3. **accuracy = 0.73**
4. accuracy = 0.59
5. accuracy = 0.66

Pipeline general



Métricas de evaluación

Identificación de vinos :
Sobre el test set

	era Alta	era Baja
Identificados como Alta	8	8
Identificados como Baja	2	92

Matriz de confusión

Métricas de evaluación

Identificación de vinos :
Sobre el test set

	era Alta	era Baja
Identificados como Alta	8	8
Identificados como Baja	2	92

Matriz de confusión

$$Accuracy = \frac{8+92}{8+8+2+92} = 0.91$$

Caso extremo: Siempre digo **Baja**

	era Alta	era Baja
Identificados como Alta	0	0
Identificados como Baja	10	100

Matriz de confusión

$$Accuracy = \frac{100}{10+100} = 0.91$$

Evalúo la capacidad de identificar los vinos de calidad Alta

	era Alta	era Baja
Identificados como Alta	8 (tp)	8 (fp)
Identificados como Baja	2 (fn)	92 (tn)

Matriz de confusión

Precision y Recall

Precision:

fracción de los identificados como **Alta** que fueron correctamente clasificados

Recall:

fracción de los que eran **Alta**, que efectivamente fueron identificados como **Alta**

	era Alta	era Baja
identificados como Alta	8 (tp)	8 (fp)
identificados como Baja	2 (fn)	92 (tn)

$$Precision = \frac{tp}{tp+fp} = \frac{8}{16} \approx 0.5$$

$$Recall = \frac{tp}{tp+fn} = \frac{8}{10} \approx 0.8$$

Precision y Recall

Casos límites:

Clasifico siempre como **Alta**

	eran Alta	eran Baja
Identificados como Alta	10 (tp)	100 (fp)
Identificados como Baja	0 (fn)	0 (tn)

$$Precision = \frac{tp}{tp+fp} = \frac{10}{110} \approx 0.09$$

$$Recall = \frac{tp}{tp+fn} = \frac{10}{10} \approx 1$$

Precision y Recall

Casos límites:

Clasifico como **Alta** solo si estoy muuuy seguro

	eran Alta	eran Baja
Identificados como Alta	3 (tp)	1 (fp)
Identificados como Baja	7 (fn)	99 (tn)

$$Precision = \frac{tp}{tp+fp} = \frac{3}{4} \approx 0.75$$

$$Recall = \frac{tp}{tp+fn} = \frac{3}{10} \approx 0.3$$

F-measure

El F-measure (F1-score) es un trade off entre el Precision y el Recall y se calcula como el promedio armónico entre ambos

$$F = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2PR}{P+R}$$

$$\begin{array}{lcl} \textit{Precision} & = & \frac{tp}{tp+fp} = \frac{8}{16} \approx 0.5 \\ \textit{Recall} & = & \frac{tp}{tp+fn} = \frac{8}{10} \approx 0.8 \end{array} \longrightarrow F = 0.616$$

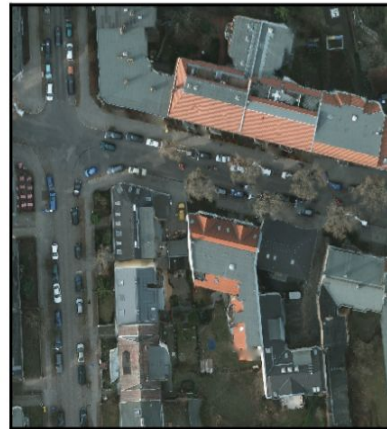
Parte 3



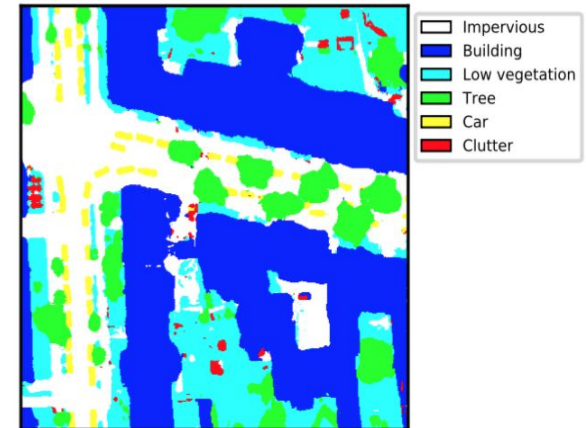
Detección de Imágenes



RGB



Prediction

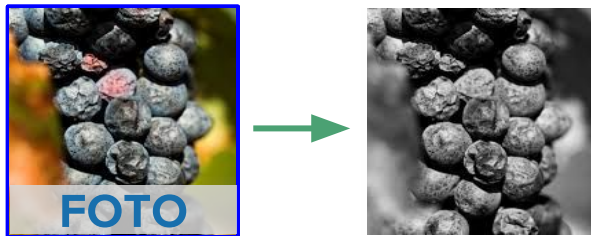


Lewis Fishgold and Rob Emanuele, 2017

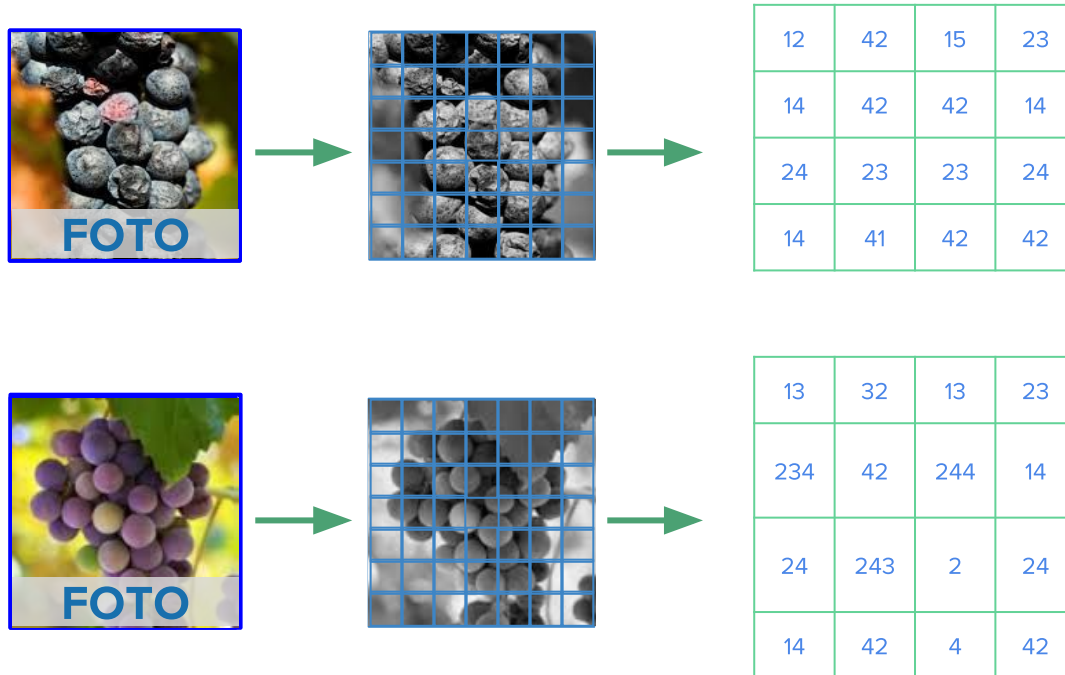
Extracción de Features



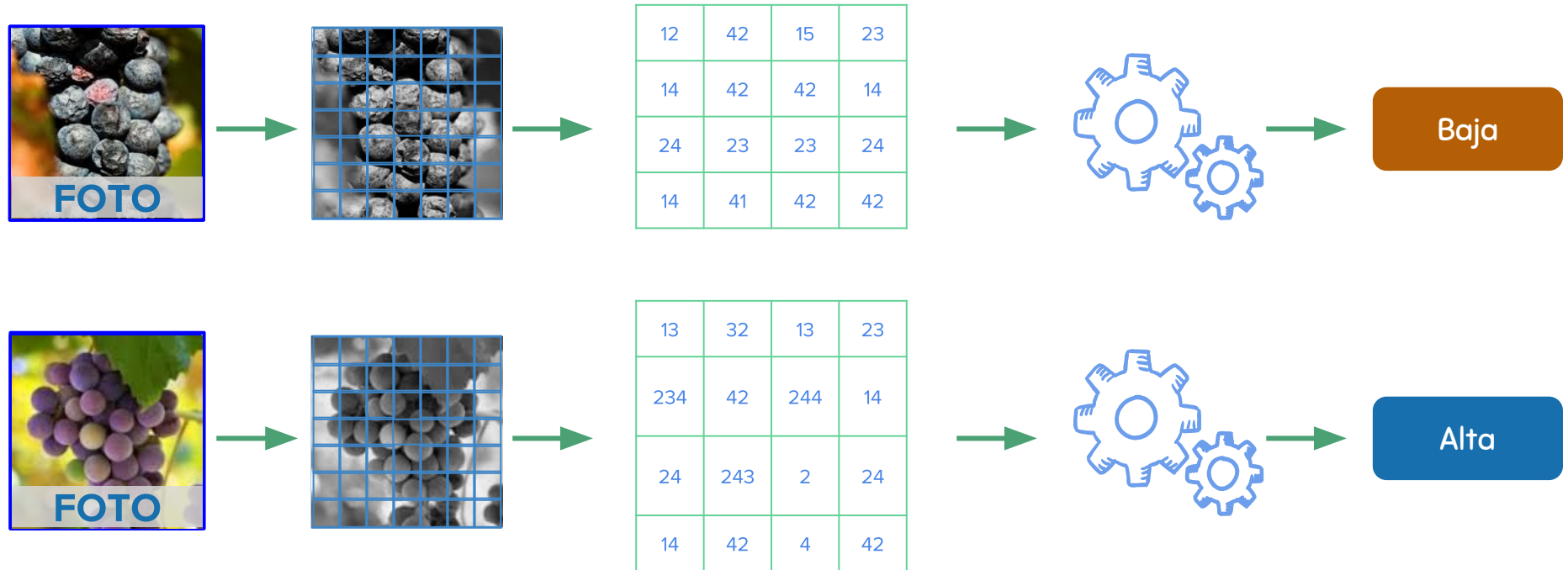
Extracción de Features



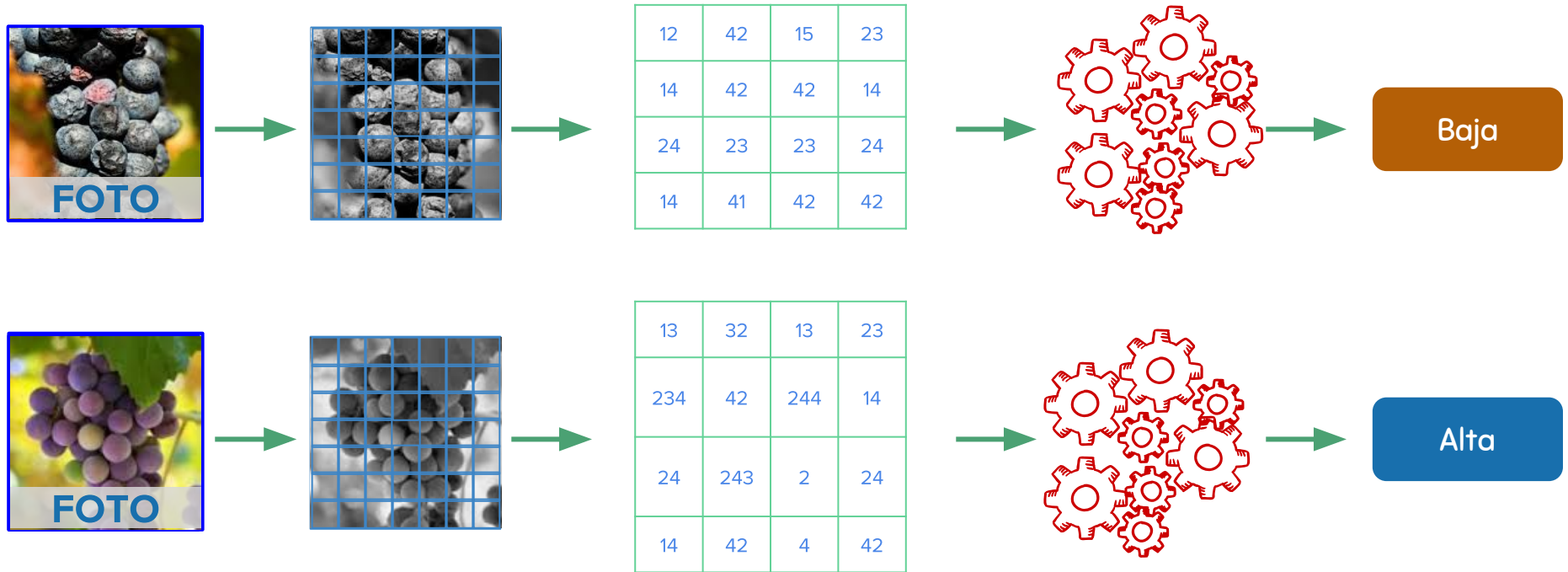
Extracción de Features



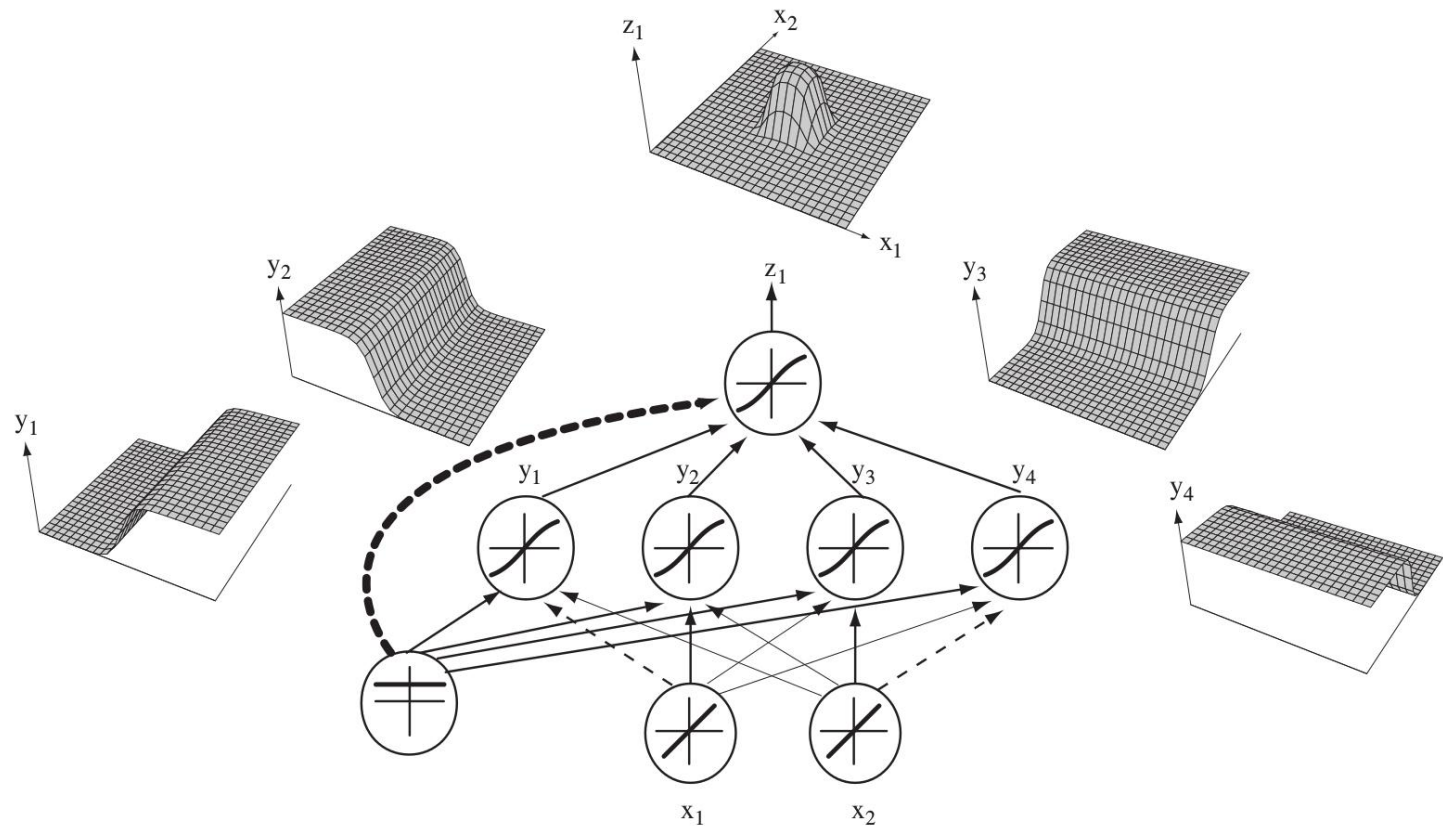
Extracción de Features



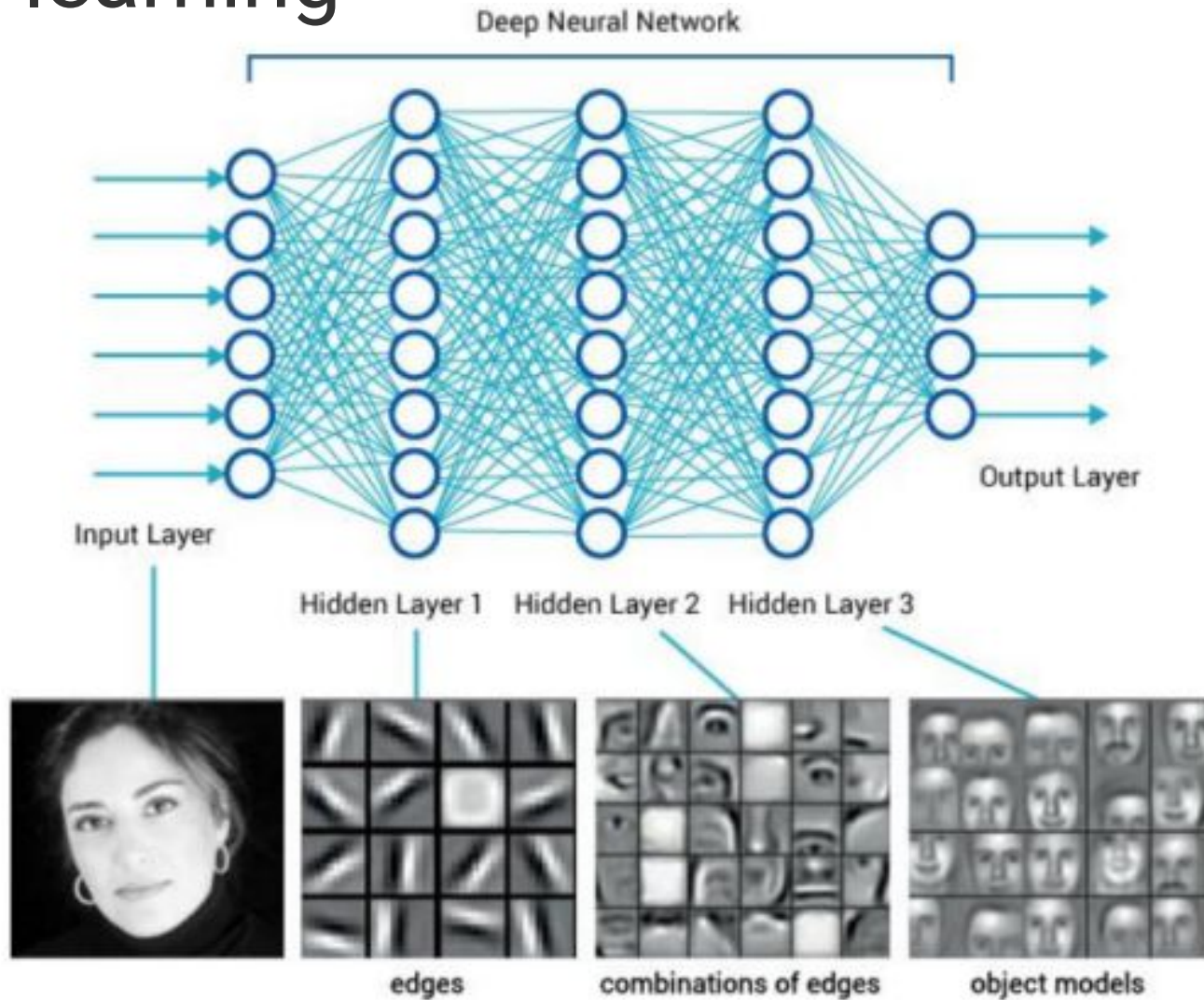
Extracción de Features



Redes Neuronales

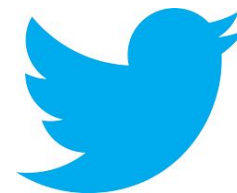


Deep learning



Procesamiento de Lenguaje Natural

Google



Bag of Words

	dinero	les	perro	vino	mercado	dólar	...
Doc 1	1	0	3	2	0	0	...
Doc 2	0	2	0	0	0	0	...
Doc 3	0	1	2	0	0	0	...
Doc 4	1	0	0	0	0	2	...
Doc 5	0	0	0	0	3	0	...
...

Podemos cuantificar el significado
de las palabras?

¿25?

0



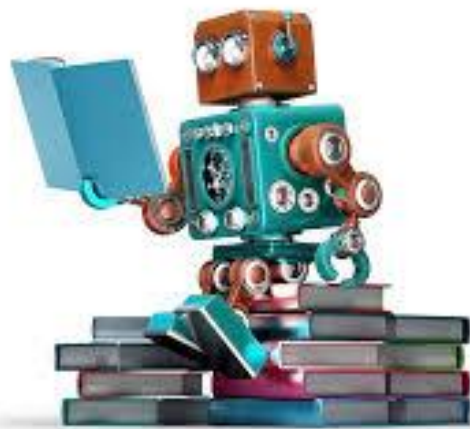
100

¿felicidad?

amor



comida



Corpus de textos

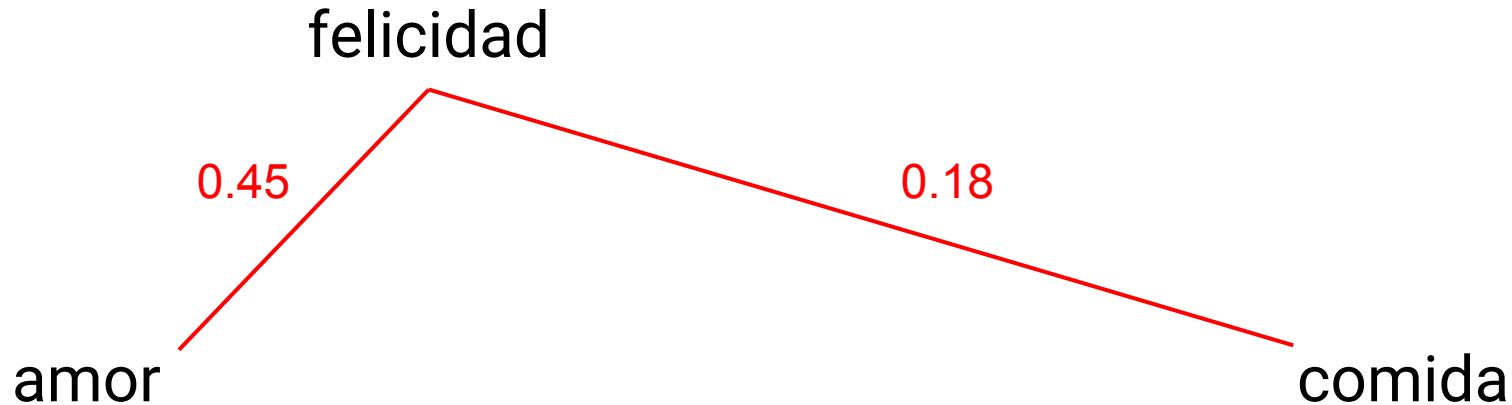
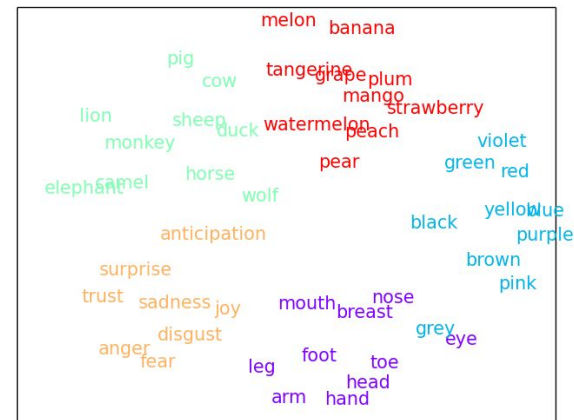


Word-embeddings



Cálculo de **cercanía** semántica entre palabras

Word-embeddings



Extracción de Features



Word-embeddings

Tinto	→ (0.4,0.9,...0.3)
de	→ (0.5,0.8,...0.5)
...	→ (...)
nariz	→ (0.1,0.9,...0.1)

50 dimensiones



promedio

(0.3,0.85,...,0.42)

50 dimensiones

FIN

