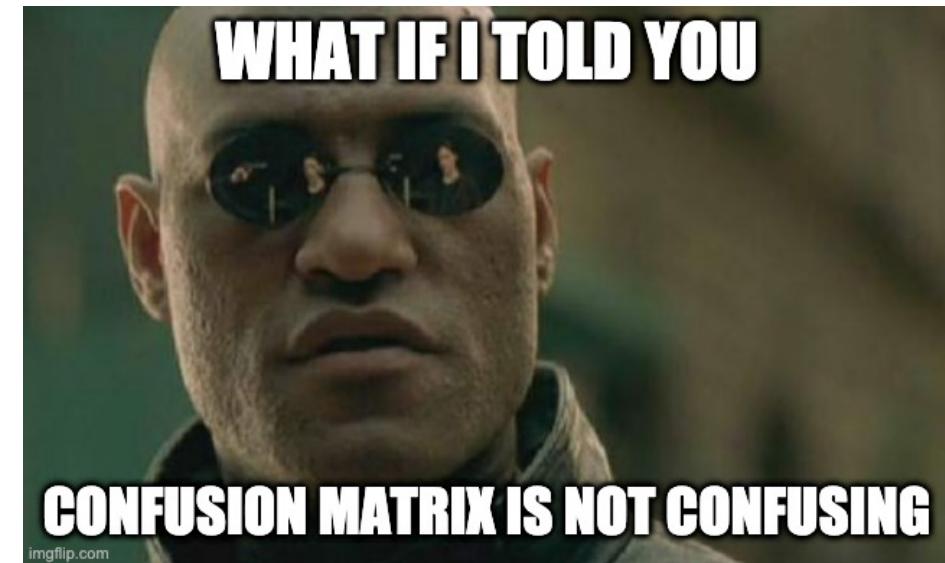


Matriz De Confusión

Pontificia Universidad Javeriana
Francisco Carlos Calderón Ph.D

Objetivos:

- Usar la matriz de confusión como la fuente de una métrica de evaluación de clasificadores.
- Calcular la matriz de confusión y sus derivados para un problema de clasificación binario.
- Reconocer las diferencias entre las métricas derivadas de la matriz de confusión.



<https://imgflip.com/i/4fv8km>

Matriz de confusión

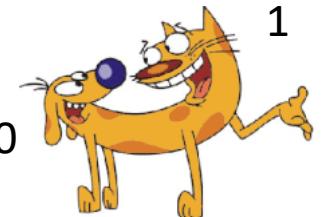
También conocida como matriz de errores,

- Es un diseño de tabla específico que permite la visualización del rendimiento de un algoritmo, de aprendizaje supervisado.
- A partir de esta pueden calcularse otras métricas puntuales que dan una idea del rendimiento del algoritmo.

Aprendiendo del ejemplo

Supongamos que tenemos 13 imágenes:

- 5 de perros 🐶 (Clase 0)
- 8 de gatos 🐱 (Clase 1).



Real = [1,1,1,1,1,1,1,1,0,0,0,0,0]

Predicho = [0,0,0,1,1,1,1,1,0,0,0,1,1]



Estas fotos están bajo licencia [CC BY-SA](#)

Aprendiendo del ejemplo

Real = [1,1,1,1,1,1,1,1,0,0,0,0,0]

Predicho = [0,0,0,1,1,1,1,1,0,0,0,1,1]

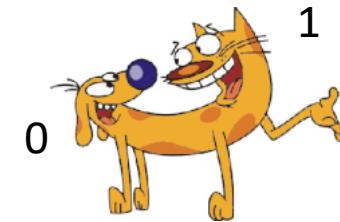
Comparando:

Hay 8 predicciones bien.

- 5 gatos
- 3 perros

Hay 5 predicciones mal.

- 3 veces eran gatos y se predijo perros,
- 2 veces eran perros y se predijo gatos



Aprendiendo del ejemplo

Ya podemos hacer la matriz de confusión:

		Clase real	
		Gato	Perro
Clase predicha	Gato	5	2
	Perro	3	3

Hay 8 predicciones bien.

5 gatos

3 perros

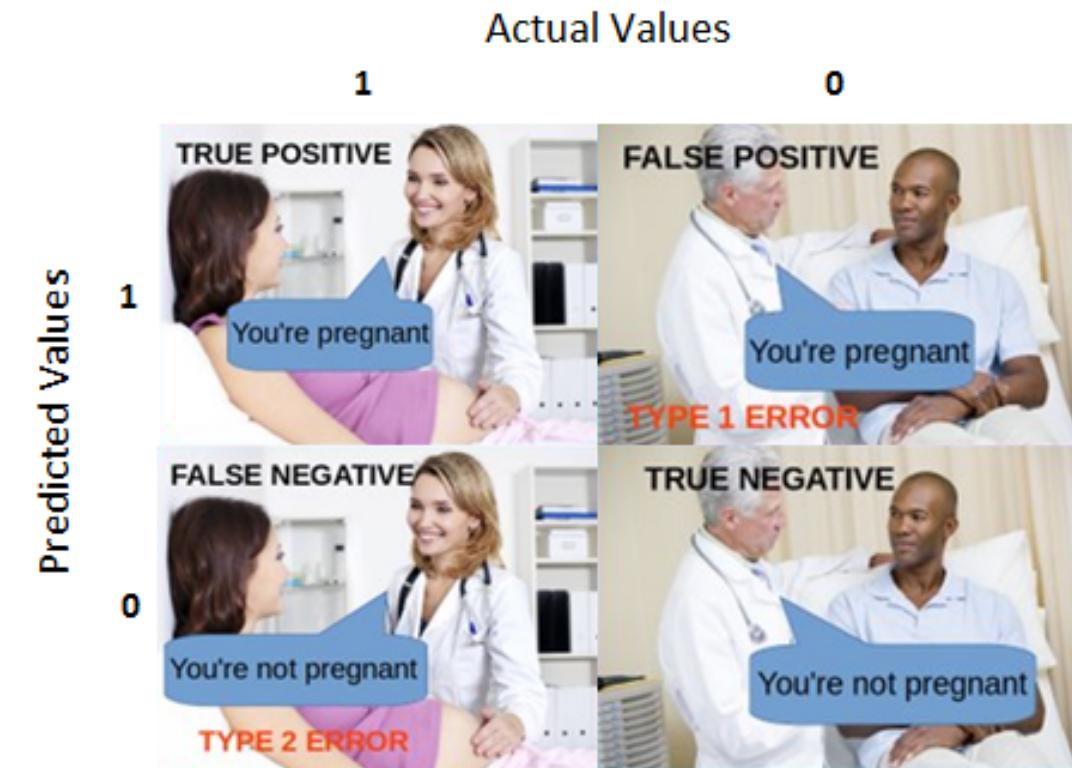
Hay 5 predicciones mal.

3 veces eran gatos y se predijo perros,

2 veces eran perros y se predijo gatos

Matriz de confusión binaria.

		Clase real	
		1	0
Clase predicha	1	Verdaderos Positivos	Falsos Positivos Error tipo 1
	0	Falsos Negativos Error Tipo 2	Verdaderos negativos



		Actual		
		P	N	
pre	P	TP	FP	
	N	FN	TN	

Derivaciones de la CM.

Verdadero positivo (TP)

Eqv. con acierto

Verdadero negativo (TN)

eqv. con correcto rechazo

Falso positivo (FP)

eqv. con falsa alarma, error tipo I

Falso negativo (FN)

eqv. con falla, error tipo II

Condición positiva (P)

El número de casos positivos reales en los datos. $P=TP+FN$

Condición negativa (N)

El número de casos negativos reales en los datos $N=TN+FP$

Sensibilidad, recuperación, tasa de aciertos o tasa de verdaderos positivos (TPR)

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

sensitivity, recall, hit rate, or true positive rate (TPR)

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR$$

Mide la proporción de positivos que se identifican correctamente. E.g. qué tan precisa es la prueba de detección para identificar la enfermedad en las personas que realmente la padecen

0 malo -> 1 bueno

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

Especificidad, selectividad o tasa de verdaderos negativos (TNR)

specificity, selectivity or true negative rate (TNR)

$$TNR = \frac{TN}{N} = \frac{TN}{TN + FP} = 1 - FPR$$

Mide la proporción de negativos que se identifican correctamente, Es la probabilidad de que los sujetos no enfermos sean clasificados como no enfermos por la prueba de detección.

0 malo -> 1 bueno

Sensibilidad y especificidad

La sensibilidad es la medida en que los verdaderos positivos no se pasan por alto,

- Por lo que los falsos negativos son “pocos” si $TPR \rightarrow 1$

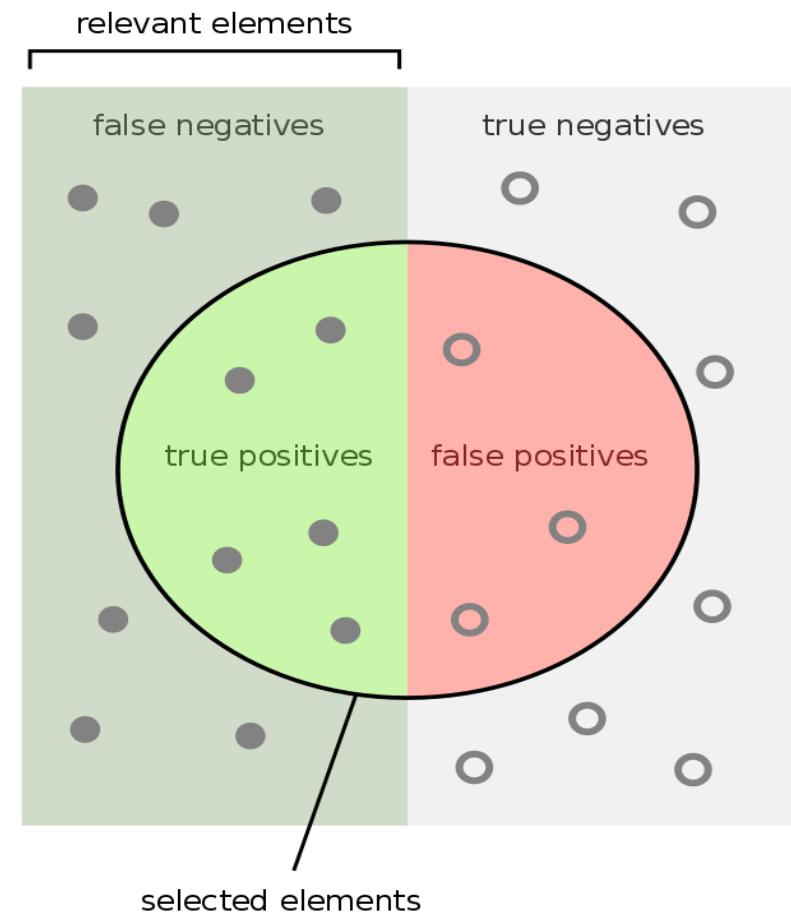
La especificidad es la medida en que los verdaderos negativos se clasifican como tales,

- Por lo que los falsos positivos son “pocos” si $TNR \rightarrow 1$

El promedio del TNR y TPR se le conoce como la Exactitud Balanceada o *Balanced Accuracy*

Matriz de confusión

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN



How many relevant items are selected?
e.g. How many sick people are correctly identified as having the condition.

How many negative selected elements are truly negative?
e.g. How many healthy people are identified as not having the condition.

$$\text{Sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

$$\text{Specificity} = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$$

Tasa de errores o tasa de falsos negativos (FNR) y Tasa de caída o falso positivo (FPR)

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

miss rate or false negative rate (FNR)

$$FNR = \frac{FN}{P} = \frac{FN}{FN + TP} = 1 - TPR$$

1 malo -> 0 bueno

fall-out or false positive rate (FPR)

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} = 1 - TNR$$

1 malo -> 0 bueno

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

Precisión o valor predictivo positivo (PPV)

precision or positive predictive value (PPV)

$$PPV = \frac{TP}{TP + FP} = 1 - FDR$$

Es la probabilidad de que los sujetos con una prueba de detección positiva realmente tengan la enfermedad.

0 malo -> 1 bueno

		Actual		
		P	N	
pre	P	TP	FP	
	N	FN	TN	

Valor predictivo negativo (VPN)

negative predictive value (NPV)

$$NPV = \frac{TN}{TN + FN} = 1 - FOR$$

Es la probabilidad de que los sujetos con una prueba de detección negativa realmente no tengan la enfermedad.

0 malo -> 1 bueno

Tasa de descubrimiento falso (FDR) y Tasa de falsas omisiones (FOR)

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

False discovery rate (FDR)

$$FDR = \frac{FP}{FP + TP} = 1 - PPV$$

1 malo -> 0 bueno

False omission rate (FOR)

$$FOR = \frac{FN}{FN + TN} = 1 - NVP$$

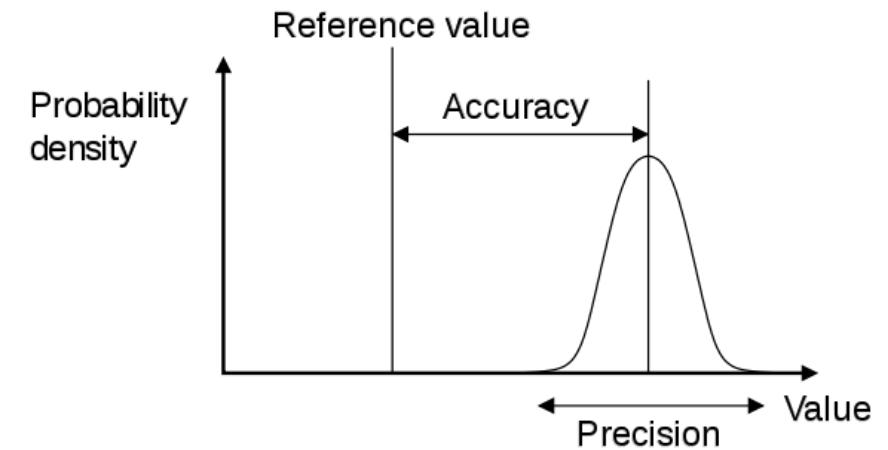
1 malo -> 0 bueno

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

Exactitud (Accuracy)

La Exactitud es la cercanía de las mediciones a un valor específico, mientras que la precisión es la cercanía de las mediciones entre sí.

$$ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$$



0 malo -> 1 bueno

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

F1 Score

Es una medida de la exactitud, puede calcularse a partir de la precisión y la sensibilidad usando la media armónica de estas.

$$F1 = 2 * \frac{PPV \cdot TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$$

también se conoce como coeficiente de Sørensen-Dice o coeficiente de similitud de Dice (DSC)

0 malo -> 1 bueno

Coeficiente de correlación de Matthews

Tiene en cuenta los positivos y negativos verdaderos y falsos.
En general, se considera una medida equilibrada que se puede utilizar incluso si las clases son de tamaños muy diferentes.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

$$MCC = \sqrt{PPV \cdot TPR \cdot TNR \cdot NPV} - \sqrt{FDR \cdot FNR \cdot FPR \cdot FOR}$$

-1 Muy malo en desacuerdo total -> 0 Malo no mejor que una moneda -> 1 Muy bueno

https://www.researchgate.net/publication/338351315_The_advantages_of_the_Matthews_correlation_coefficient_MCC_over_F1_score_and_accuracy_in_binary_classification_evaluation

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5721660/>

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

		Actual	
		P	N
pre	P	TP	FP
	N	FN	TN

Resumiendo:

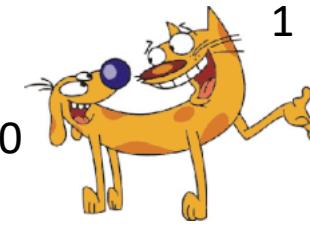
	True condition				
Total population	Condition positive	Condition negative	Prevalence = $\frac{\sum \text{Condition positive}}{\sum \text{Total population}}$	Accuracy (ACC) = $\frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total population}}$	
Predicted condition	True positive	False positive, Type I error	Positive predictive value (PPV), Precision = $\frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}}$	False discovery rate (FDR) = $\frac{\sum \text{False positive}}{\sum \text{Predicted condition positive}}$	
Predicted condition negative	False negative, Type II error	True negative	False omission rate (FOR) = $\frac{\sum \text{False negative}}{\sum \text{Predicted condition negative}}$	Negative predictive value (NPV) = $\frac{\sum \text{True negative}}{\sum \text{Predicted condition negative}}$	
	True positive rate (TPR), Recall, Sensitivity, probability of detection, Power $= \frac{\sum \text{True positive}}{\sum \text{Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm $= \frac{\sum \text{False positive}}{\sum \text{Condition negative}}$	Positive likelihood ratio (LR+) $= \frac{\text{TPR}}{\text{FPR}}$	Diagnostic odds ratio (DOR) $= \frac{\text{LR+}}{\text{LR-}}$	F_1 score = $2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$
	False negative rate (FNR), Miss rate = $\frac{\sum \text{False negative}}{\sum \text{Condition positive}}$	Specificity (SPC), Selectivity, True negative rate (TNR) $= \frac{\sum \text{True negative}}{\sum \text{Condition negative}}$	Negative likelihood ratio (LR-) $= \frac{\text{FNR}}{\text{TNR}}$		

Tomado de Wikimedia commons

hiperparámetros

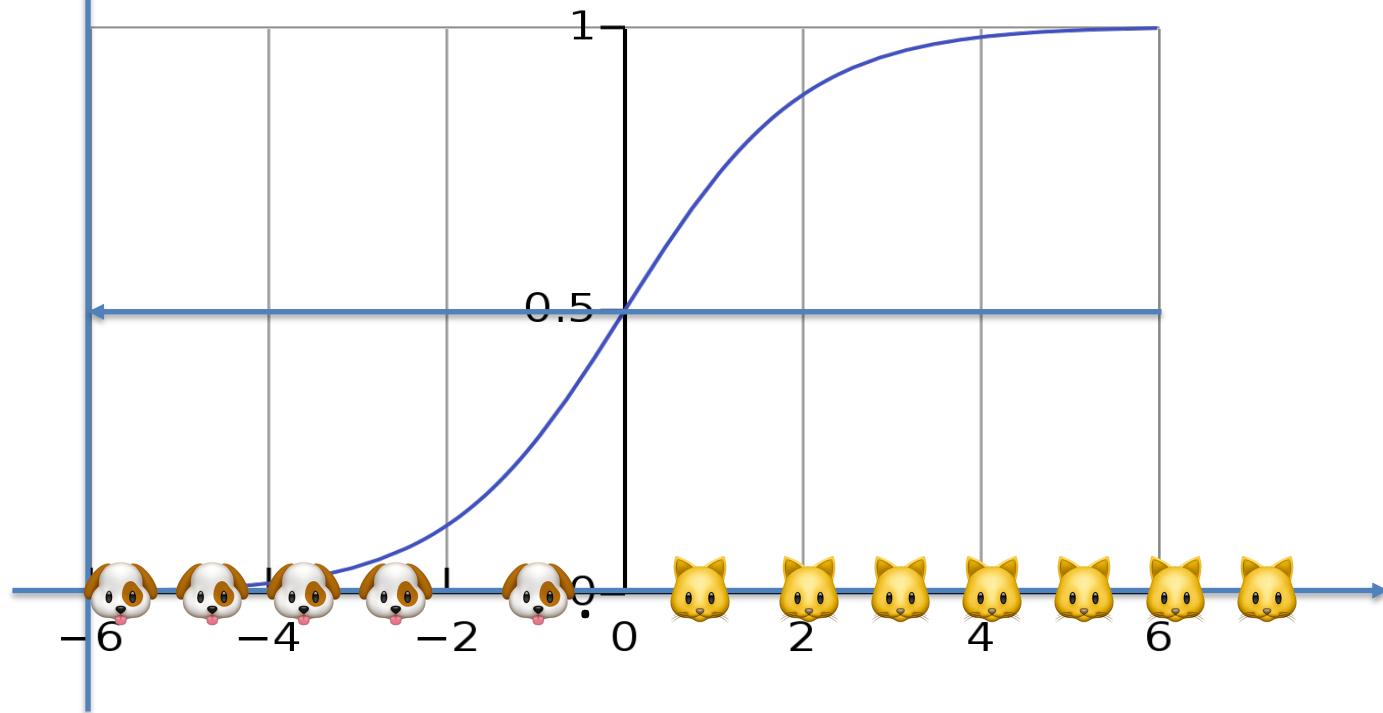
Son propiedades del modelo estadístico del algoritmo, que pueden influir en:

- Su complejidad,
- Su velocidad de aprendizaje,
- Los resultados de su aplicación.



ROC y AUC-ROC

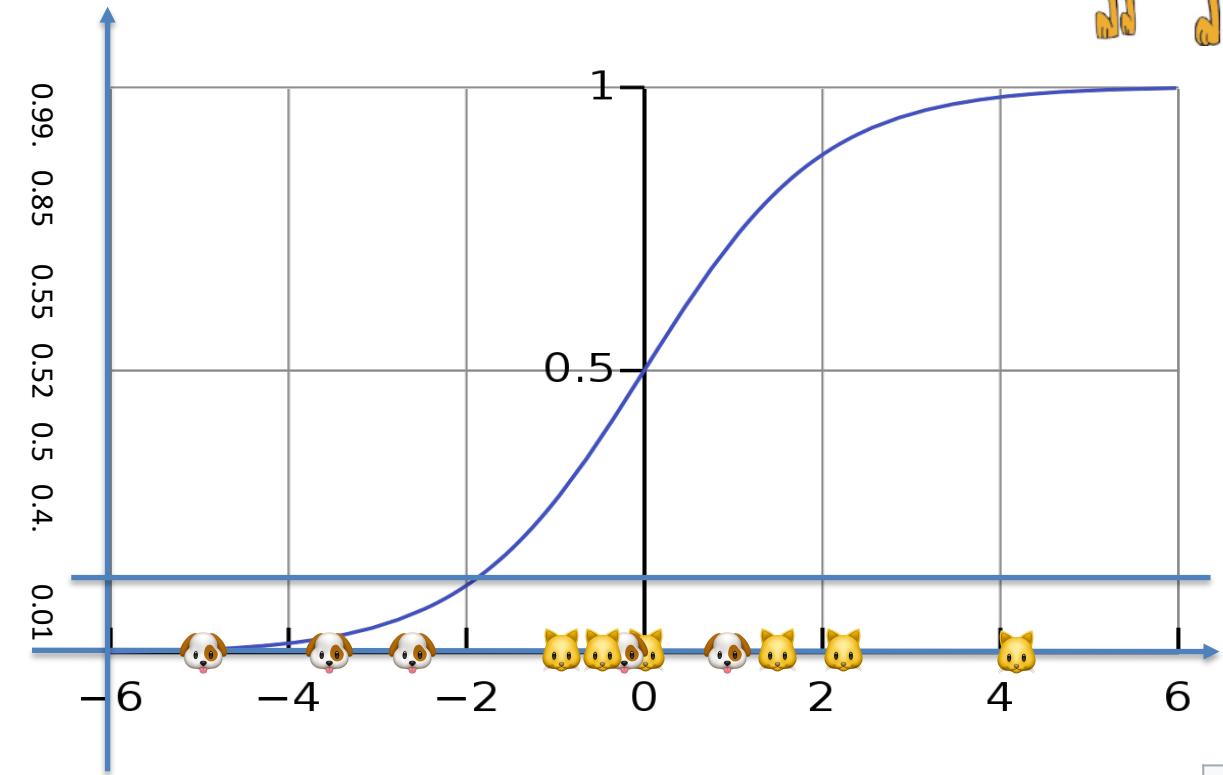
La curva ROC se crea trazando la tasa de verdaderos positivos (TPR) contra la tasa de falsos positivos (FPR) para varios valores de umbral. “Diferentes Hiperparámetros”



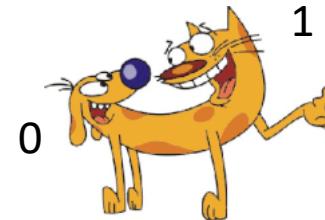
La curva característica de funcionamiento, o curva ROC, es un diagrama gráfico que ilustra la capacidad de diagnóstico de un sistema clasificador binario a medida que varía su umbral de discriminación.

El método fue desarrollado para operadores de receptores de radar militares, razón por la cual se llama así. Receiver Operating Characteristic

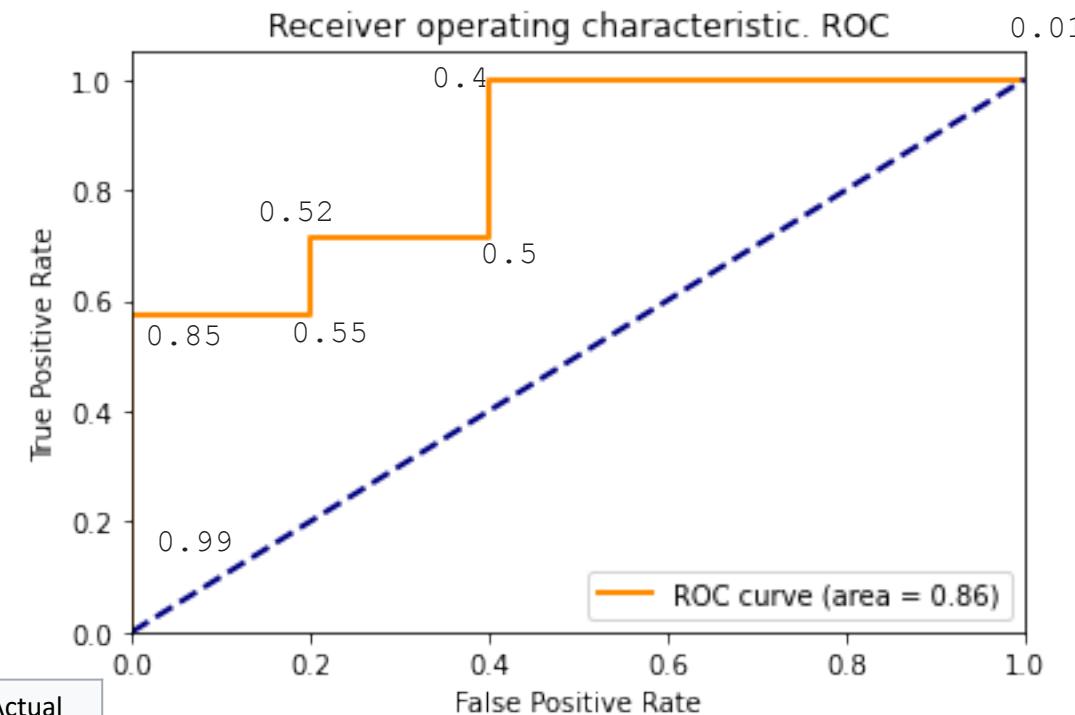
ROC y AUC-ROC



Clases=[1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0]
Scores=[0.99, 0.94, 0.92, 0.85, 0.52, 0.45, 0.40, 0.55, 0.5, 0.22, 0.1, 0.01]



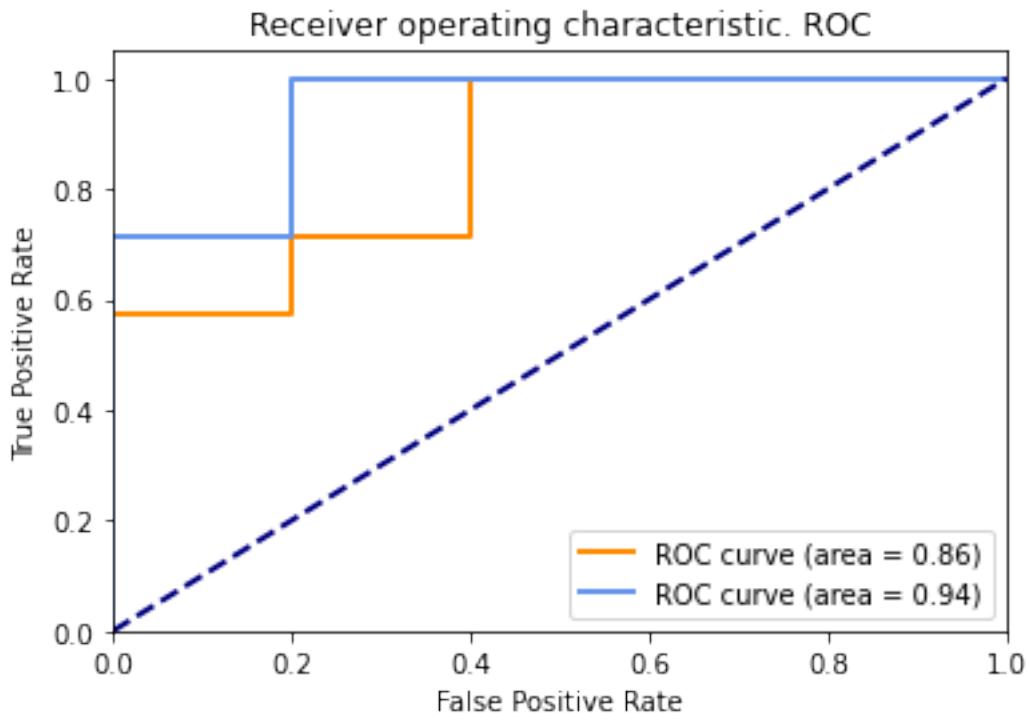
La curva ROC se crea trazando la tasa de verdaderos positivos (TPR) contra la tasa de falsos positivos (FPR) para varios valores de umbral. "Diferentes Hiperparámetros"



$$FPR = \frac{FP}{N} = \frac{FP}{FP+TN} = 1 - TNR$$

ROC y AUC-ROC

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$



$$FPR = \frac{FP}{N} = \frac{FP}{FP+TN} = 1 - TNR$$

```
Clases=[1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0]
Scores=[0.99, 0.94, 0.92, 0.85, 0.52, 0.45, 0.40, 0.55, 0.5, 0.22, 0.1, 0.01]
Scores2=[0.99, 0.94, 0.92, 0.85, 0.7, 0.55, 0.46, 0.58, 0.3, 0.22, 0.1, 0.01]
```

Multiclas

La Matriz de confusión
también puede usarse en
problemas de múltiples
clases.

Pero eso lo veremos la
semana siguiente!



Ejercicio en clase

Realizar un ejemplo de la ROC obtenida usando KNN cambiando los hiperparámetros del método para el problema de clasificación binaria suministrado.

Tarea:

Usando uno de los dos datasets:

<https://www.kaggle.com/c/titanic>

<https://www.kaggle.com/renanmav/which-team-will-win-this-dota-2-match>

Implemente el **mejor** clasificador que usted considere conveniente de los vistos en clase.

Bibliografía

<https://people.inf.elte.hu/kiss/11dwhdm/roc.pdf>

Matriz de confusión

Facultad de ingeniería
Deptº de electrónica

