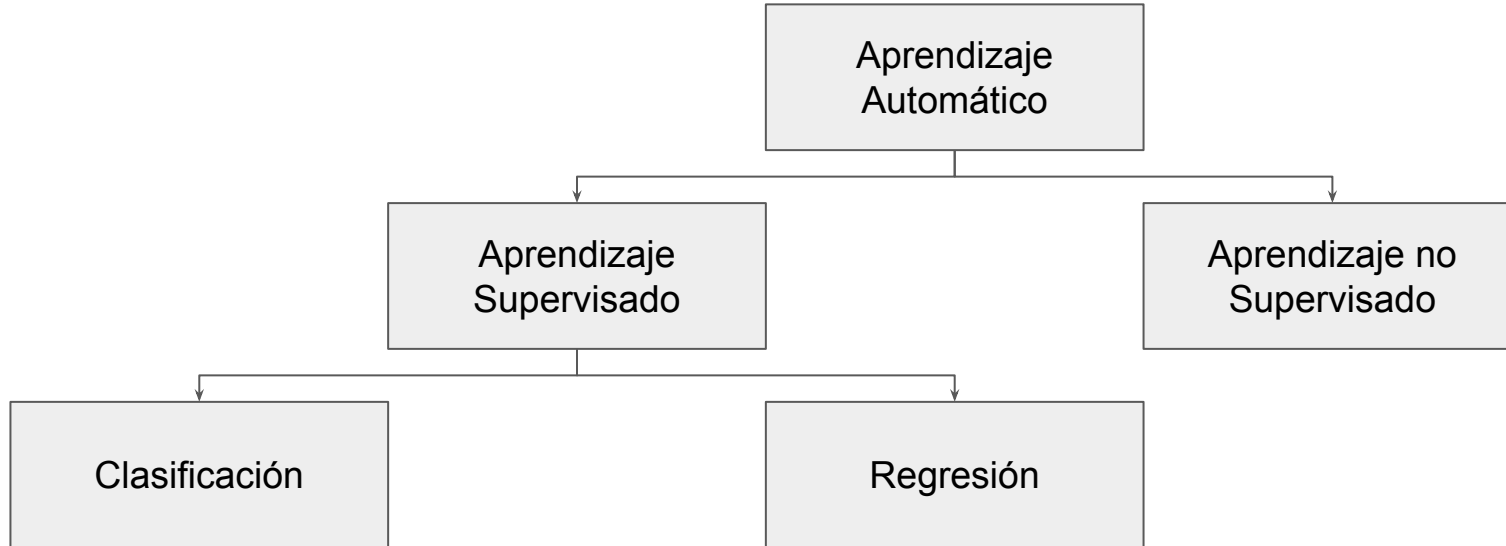


Aprendizaje supervisado

Fabián Villena

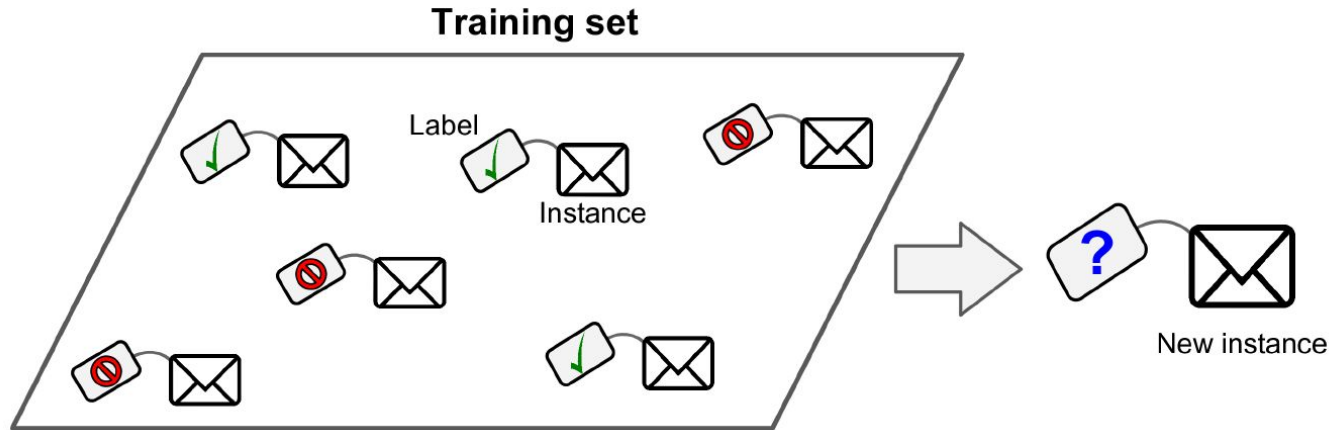
Aprendizaje supervisado

En el aprendizaje supervisado, los datos de entrenamiento que se ingresan al algoritmo incluyen las soluciones.



Clasificación

Una de las tareas que se resuelven con aprendizaje supervisado es la clasificación, en donde la etiqueta asociada a cada ejemplo es una categoría o *clase*.



Métricas de rendimiento

Para poder entender el rendimiento de un modelo y poder valorar si está funcionando de manera correcta, debemos calcular un valor numérico. Cada métrica de rendimiento va a tener un significado específico y debemos interpretarlas de manera correcta.

	precision	recall	f1-score	support
0	0.77	0.86	0.81	37584
1	0.84	0.75	0.79	37577
accuracy			0.80	75161
macro avg	0.81	0.80	0.80	75161
weighted avg	0.81	0.80	0.80	75161

Matriz de confusión

Una manera de evaluar el rendimiento de un clasificador es observar la matriz de confusión. La idea general es contar la cantidad de veces que los ejemplos de la clase A son clasificados como clase B. Este cálculo normalmente se debe realizar en el subconjunto de prueba.

		Predicho	
		Negativo	Positivo
Real	Negativo	Verdaderos negativos (FN)	Falsos positivos (FN)
	Positivo	Falsos negativos (FN)	Verdaderos positivos (FN)

Accuracy

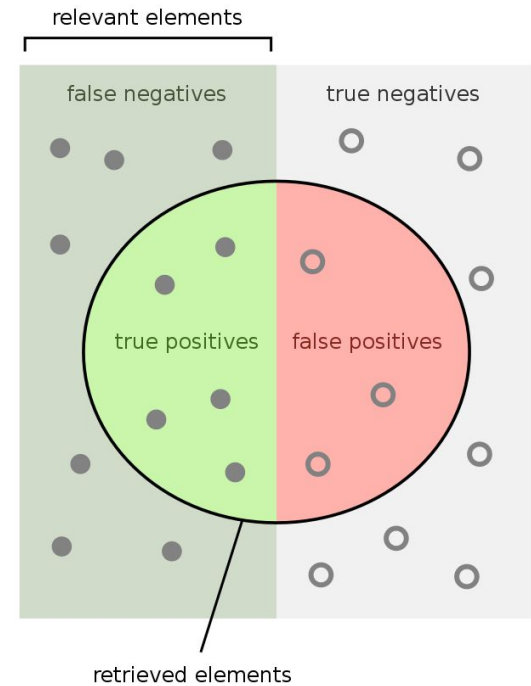
La *accuracy* es una de las métricas que podemos utilizar para evaluar el rendimiento de un modelo. La *accuracy* es la fracción de predicciones que nuestro modelo realizó correctamente. El problema es que en problemas desbalanceados, esta métrica puede llevarnos a falsas conclusiones.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision y recall

- La *precision* busca identificar la proporción de predicciones positivas que realmente eran positivas.
- El *recall* busca identificar la proporción de elementos realmente positivos que se predijeron como positivos.

Para evaluar el rendimiento de nuestro modelo debemos evaluar ambas métricas. Típicamente existe un compromiso entre ambas métricas que debemos manejar con cuidado.



How many retrieved items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are retrieved?

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

F-score

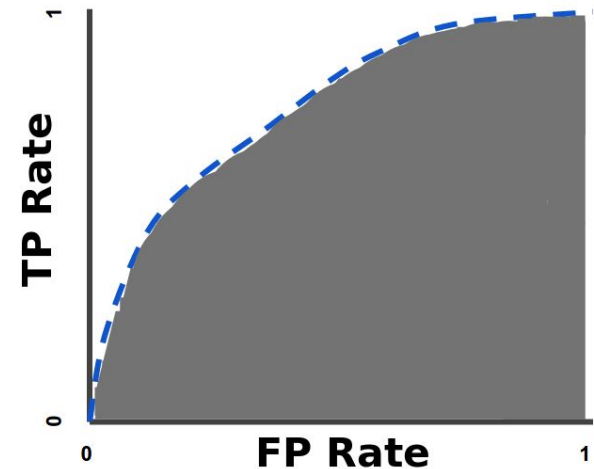
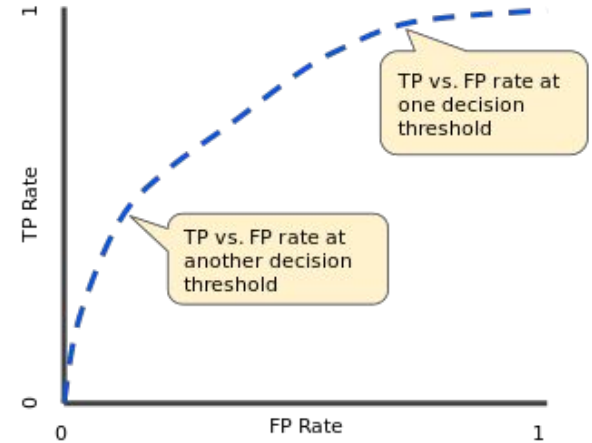
El *F*-score es una métrica resumen que nos comunica el resultado de la *precision* y el *recall* en un sólo valor. La versión más utilizada es el F_1 -score, que pondera con el mismo peso la *precision* y el *recall*.

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}$$

$$F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

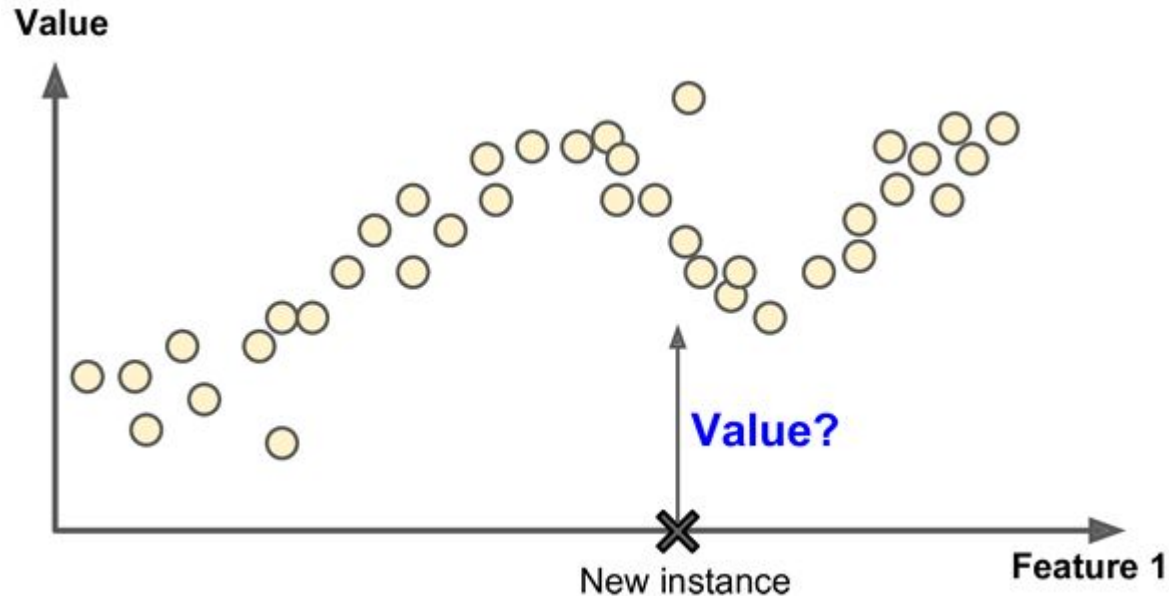
Área bajo la curva ROC

La curva ROC es una representación del cambio en el False Positive Rate (FPR; la proporción de elementos negativos que fueron incorrectamente clasificados como positivos) y el True Positive Rate (TPR; Recall) de un clasificador binario al modular su umbral de decisión. La métrica derivada de esta curva es el área bajo la misma.



Regresión

Una de las tareas que se resuelven con aprendizaje supervisado es la regresión, en donde la etiqueta asociada a cada ejemplo es un valor continuo.



Métricas de rendimiento

Las métricas más utilizadas para evaluar el rendimiento de un regresor son el *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE) y el *Root Mean Squared Error* (RMSE).

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

$$\text{MSE} = \frac{\sum_{i=1}^n |(y_i - \hat{y}_i)^2|}{n}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n |(y_i - \hat{y}_i)^2|}{n}}$$