

# Aprendizaje automatizado

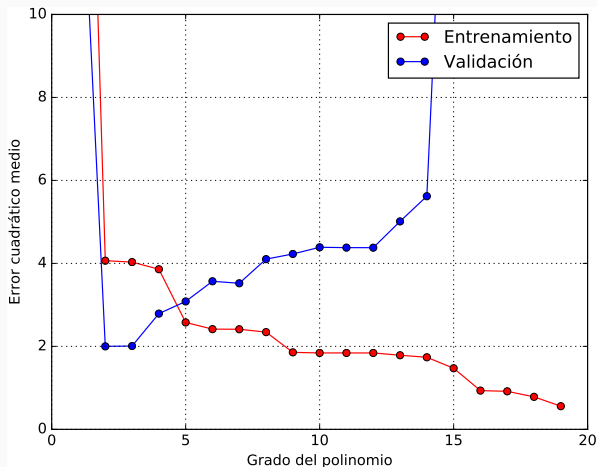
## EVALUACIÓN DE MODELOS

---

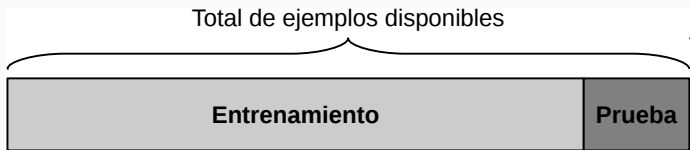
Gibran Fuentes-Pineda

Marzo 2023

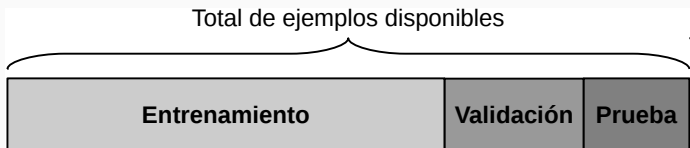
# El problema de la generalización revisitado



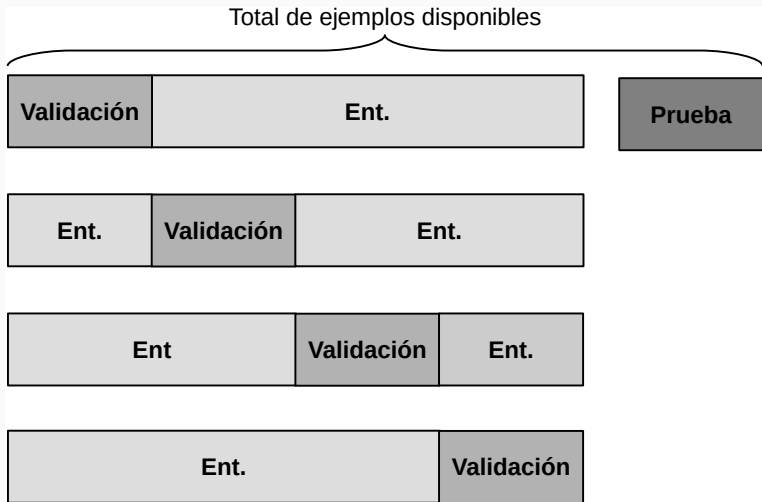
## Partición de los datos en entrenamiento y prueba



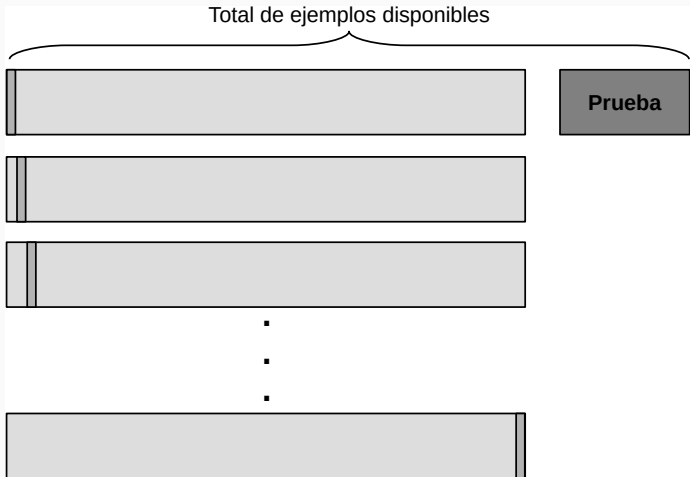
## Dividiendo los datos en entrenamiento, validación y prueba



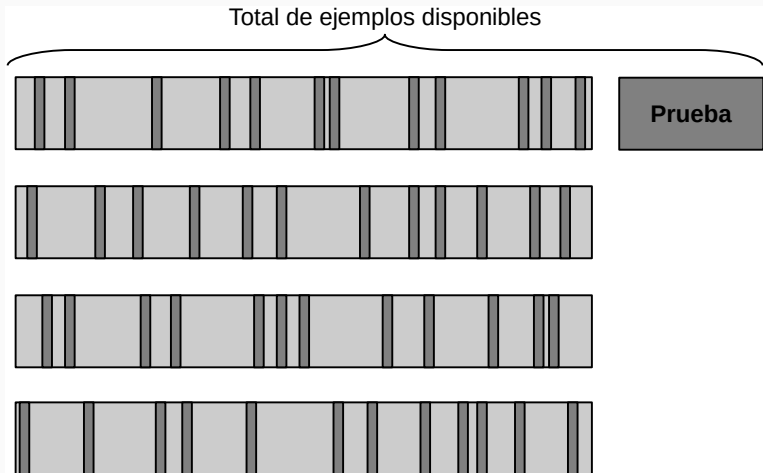
# Validación cruzada con $K$ particiones



## Validación cruzada dejando uno fuera (LOOCV)



# Validación cruzada aleatoria



# Cálculo del error en validación cruzada

- Promedio de los errores en cada partición

$$E = \frac{1}{K} \sum_{i=1}^K E_i$$

- En el caso de LOOCV

$$E = \frac{1}{n} \sum_{i=1}^n E_i$$



# Medidas de rendimiento para regresión

- Error cuadrático medio (ECM)

$$ECM(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n} \sum_{i=1}^n \left( \hat{y}^{(i)} - y^{(i)} \right)^2$$

- Raíz del error cuadrático medio (RECM)

$$RECM(\mathbf{y}, \hat{\mathbf{y}}) = \sqrt{ECM(\mathbf{y}, \hat{\mathbf{y}})}$$

- Erro absoluto medio (EAM)

$$EAM(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n} \sum_{i=1}^n \left| \hat{y}^{(i)} - y^{(i)} \right|$$

# Medidas de rendimiento para regresión

- Coeficiente de determinación ( $R^2$ )

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

donde

$$SS_{tot} = \sum_{i=1}^n \left( y^{(i)} - \mu \right)^2$$

$$SS_{res} = \sum_{i=1}^n \left( y^{(i)} - \hat{y}^{(i)} \right)^2$$

$$\mu = \frac{1}{n} \cdot \sum_{i=1}^n y^{(i)}$$

# Medidas de rendimiento de clasificadores binarios

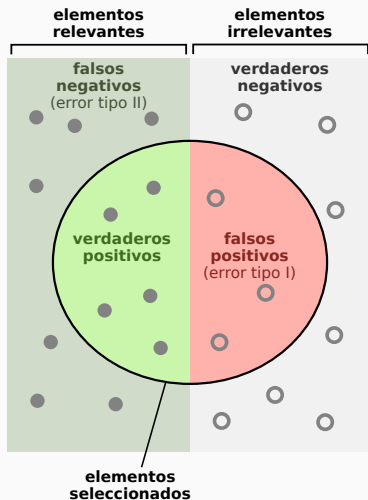


Figura traducida de Wikipedia (entrada de *Precision and Recall*)

# Medidas de rendimiento de clasificadores binarios

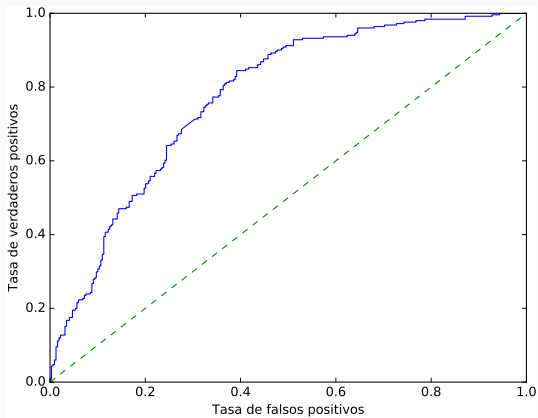
$$\text{precisión} = \frac{|\text{verdaderos positivos}|}{|\text{elementos seleccionados}|}$$

$$\text{exhaustividad} = \frac{|\text{verdaderos positivos}|}{|\text{elementos relevantes}|}$$

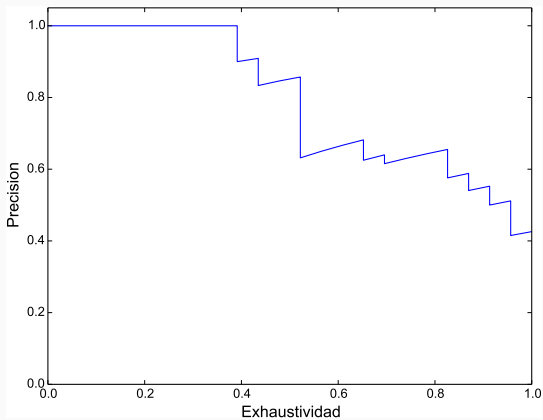
$$\text{tasa de verdaderos positivos} = \text{exhaustividad}$$

$$\text{tasa de falsos positivos} = \frac{|\text{falsos positivos}|}{|\text{elementos irrelevantes}|}$$

# Curva ROC



# Curva de precisión-exhaustividad



# Matriz de confusión

		Clase Verdadera	
		Cáncer	No Cáncer
Clase Predicha	Cáncer	5 VP	3 FP
	No Cáncer	10 FN	6 VN

- **Compacidad:** Mide qué tan cerca están los elementos del mismo clústeres
- **Separación:** Mide qué tan separados están los elementos de diferentes clústeres



- **Pureza:** Mide la proporción de la clase con mayor número de elementos en el clúster con respecto al tamaño del mismo

- **Jaccard:**

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{VP}{VP + FP + FN}$$

# Métricas con clases desbalanceadas

- Considera la tarea de clasificación de correo no deseado.

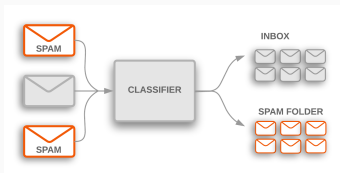


Figura reproducida de <https://developers.google.com/machine-learning/guides/text-classification>

# Métricas con clases desbalanceadas

- Considera la tarea de clasificación de correo no deseado.

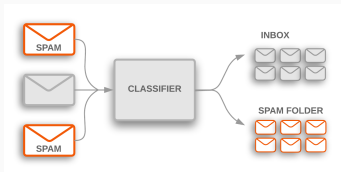


Figura reproducida de <https://developers.google.com/machine-learning/guides/text-classification>

- Nuestro conjunto de datos disponible contiene 96 % de correo normal y tan sólo 4 % de correo no deseado.

# Métricas con clases desbalanceadas

- Considera la tarea de clasificación de correo no deseado.

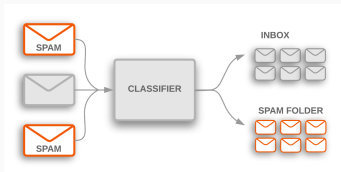


Figura reproducida de <https://developers.google.com/machine-learning/guides/text-classification>

- Nuestro conjunto de datos disponible contiene 96 % de correo normal y tan sólo 4 % de correo no deseado.
- Entrenamos un clasificador con un subconjunto de estos datos y evaluamos su exactitud con el restante, obteniendo un 96 % de exactitud. ¿Es este un buen modelo?

# Impacto del desbalance en el aprendizaje

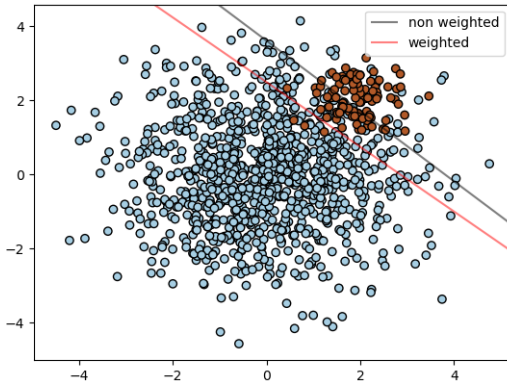


Figura reproducida de [https://scikit-learn.org/stable/auto\\_examples/svm/plot\\_separating\\_hyperplane\\_unbalanced.html](https://scikit-learn.org/stable/auto_examples/svm/plot_separating_hyperplane_unbalanced.html)

# Estrategias de aprendizaje para clases desbalanceadas

- Generar ejemplos artificiales de clase más escasa (*oversampling*)
- Elegir un subconjunto más pequeño de las clases más comunes (*undersampling*)
- Usar función de pérdida pesada