

Análisis de Desempeño de los Modelos

CURSO GRUPO BANCOLOMBIA

Universidad Nacional de Colombia



Análisis de Desempeño para Clasificadores

- Evaluar el modelo de aprendizaje de máquina es una parte fundamental.
- Hay diferentes medidas disponibles.



Exactitud (Accuracy)

- La exactitud es la relación entre el número de las predicciones correctas y el número de predicciones totales. O sea, es la proporción de las predicciones correctas.

$$Accuracy = \frac{\# \text{ predicciones correctas}}{\# \text{ predicciones totales}}$$

- El mejor valor posible es 1.

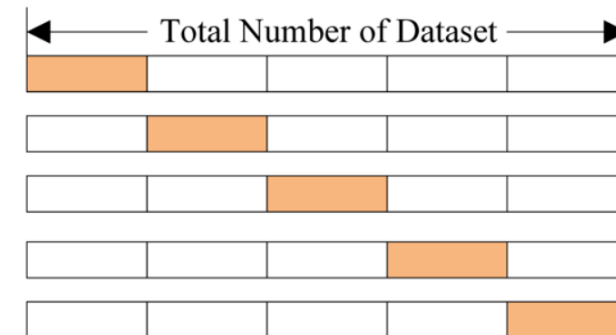
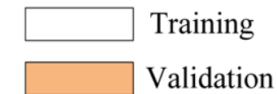
Validación cruzada (Cross-Validation)

La validación cruzada es una implementación excelente para medir el desempeño del modelo.

- Lo que se hace es dividir el conjunto de datos, por ejemplo en cinco.
- Para cada porción de los datos se genera un clon del clasificador.
- Se entrena el modelo con la porción de los datos
- Se evalúa el modelo con la medida de exactitud.
- Y finalmente se comparan los cinco resultados.

Nota: cuando se tiene una base de datos desbalanceada, es decir, algunas de las clases tiene más ejemplo que otras, la exactitud no es la medida de rendimiento más adecuada.

Tomado de: <https://www.kaggle.com/kashnitsky/topic-3-decision-trees-and-knn>





Matriz de confusión

La matriz de confusión es uno de los mejores métodos para evaluar un clasificador.

- Básicamente es contar el número de veces que una clase A es clasificada como B.
- Cada columna representa la clase real y cada fila representa la clase predicha.

En el ejemplo:

- Se tiene una base de datos con 100 ejemplos ($n=100$).
- El número 40 son los Verdaderos Negativos (VN), los casos en los que se dijo NO y la clase a la que pertenece es efectivamente NO.
- El número 10 son los Falsos Positivos (FP), los casos en los que se dijo SÍ y en realidad eran NO.
- El número 5 son los Falsos Negativos (FN), los casos en los que se dijo NO y en realidad eran SÍ.
- El número 45 son los Verdaderos Positivos (VP), los casos en los que se dijo SÍ y el resultado real era SÍ.
- Un clasificador perfecto sólo tendría VP y VN (la diagonal principal)

N=100	Predicción SI	Predicción NO
	Clase: SI	Clase: NO
	45	5
	10	40

Matriz de confusión

La matriz de confusión proporciona mucha información del clasificador.

- La exactitud: predicciones correctas en proporción a las predicciones totales.
- Precisión: predicciones correctas en proporción a las predicciones positivas.
- *Recall o Sensibilidad o tasa de verdaderos positivos (TPR)*: es la proporción de instancias positivas que se clasificaron correctamente.
- F1 Score: es la medida armónica de la Precisión y el Recall. Sólo tendrá un valor alto si ambos tiene un valor alto.

$$F1_{score} = 2 \frac{Precision * Recall}{Precision + Recall} = \frac{VP}{VP + \frac{FN + FP}{2}}$$

$$Accuracy = \frac{VP + VN}{N}$$

$$Precision = \frac{VP}{VP + FP}$$

$$Recall = \frac{VP}{VP + FN}$$

N	Predicción SI	Predicción NO
	Verdaderos Positivos (VP)	Falsos Negativos (FN)
Clase: SI		
Clase: NO	Falsos Positivos (FP)	Verdaderos Negativos (VN)



Matriz de confusión

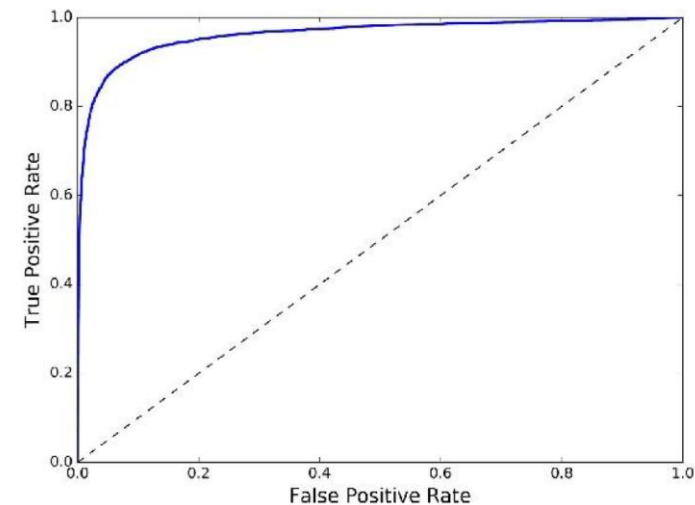
La medida preferida dependerá del proyecto, por ejemplo:

- Un clasificador de videos para niños, es preferible que el clasificador rechace muchos videos buenos (**Recall** baja), pero sí rechace todos los videos negativos (alta **Precision**).
- Un clasificador para detectar transacciones fraudulentas, probablemente está bien un clasificador con un 30% de **Precision**. Pero debe tener un 99% **Recall**. Está claro que se recibirán falsas alarmas, pero casi todas las transacciones fraudulentas serán revisadas.

Curva ROC

Característica Operativa del Receptor (receiver operating characteristic).

- Es la gráfica de la *tasa positiva verdadera (Recall)* vs la *tasa de falsos positivos (FPR)*.
- El FPR es la proporción de instancias negativas que se clasifican incorrectamente como positivas.
- En la gráfica, La línea punteada representa la curva ROC de un clasificador puramente aleatorio.
- Un buen clasificador se mantiene lo más alejado posible de esa línea (hacia la esquina superior izquierda).





¡A codificar!

