

Validación de Modelos

Analítica de Datos, Universidad de San Andrés

Matriz de Confusión Binaria

		Real	
Predicho	Real	VP	FP
	Predicho	FN	VN

		Predicho	
Real	Predicho	VP	FN
	Real	FP	VN

- VP: Verdaderos Positivos
- VN: Verdaderos Negativos
- FP: Falsos Positivos
- FN: Falsos Negativos

Matriz de Confusión Multiclase

		R		
		L	N	S
P	L	85	10	5
	N	8	82	10
	S	7	12	81

- L: Lluvia
- N: Nublado
- S: Soleado

Accuracy (Exactitud)

$$\text{Accuracy} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{VN} + \text{FP} + \text{FN}}$$

Es el porcentaje total de predicciones correctas que hace nuestro modelo.

Precision (Precisión)

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

Es qué tan confiables son nuestras predicciones positivas.

Specificity (Especificidad) / True Negative Rate (TNR)

$$\text{Specificity} = \frac{\text{VN}}{\text{VN} + \text{FP}}$$

Es qué tan bien identificamos los casos negativos correctamente.

Recall (Sensibilidad) / True Positive Rate (TPR)

$$\text{Recall (TPR)} = \frac{\text{VP}}{\text{VP} + \text{FN}}$$

Es qué tan bien detectamos todos los casos positivos que existen.

Fall-out / False Positive Rate (FPR)

$$\text{Fall-out (FPR)} = \frac{\text{FP}}{\text{FP} + \text{VN}}$$

Es qué tan seguido nos equivocamos prediciendo casos positivos.

Miss Rate / False Negative Rate (FNR)

$$\text{Miss Rate (FNR)} = \frac{\text{FN}}{\text{FN} + \text{VP}}$$

Es qué tan seguido nos perdemos casos positivos que deberíamos haber detectado.

$$\text{F1-Score} = 2 \times \frac{\textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

Es un balance entre la precisión y el recall, útil cuando las clases están desbalanceadas.

False Discovery Rate (FDR)

$$\text{FDR} = \frac{\text{FP}}{\text{FP} + \text{VP}}$$

Es qué tan seguido nos equivocamos cuando hacemos predicciones positivas.

False Omission Rate (FOR)

$$\text{FOR} = \frac{\text{FN}}{\text{FN} + \text{VN}}$$

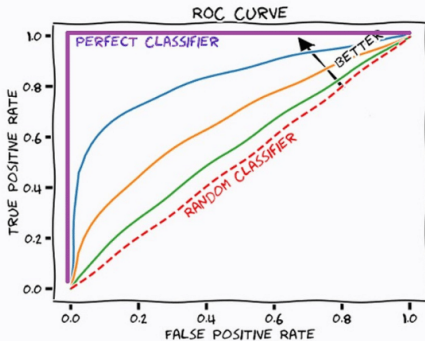
Es qué tan seguido nos equivocamos cuando hacemos predicciones negativas.

Negative Predictive Value (NPV)

$$\text{NPV} = \frac{\text{VN}}{\text{VN} + \text{FN}}$$

Es qué tan confiables son nuestras predicciones negativas.

Curva ROC



La curva ROC muestra la relación entre TPR y FPR para diferentes umbrales. El área bajo la curva (AUC) indica el rendimiento del modelo.

¿Cómo entender la curva ROC?

Para cada punto de umbral existe una tasa de falsos positivos y una tasa de verdaderos positivos. Las coordenadas que trazamos son (FPR, TPR) en el eje X e Y respectivamente. Recordamos que:

$$\text{Fall-out (FPR)} = \frac{\text{FP}}{\text{FP} + \text{VN}}$$

$$\text{Recall (TPR)} = \frac{\text{VP}}{\text{VP} + \text{FN}}$$

Umbral 0 - Clasificar todo como positivo (1/2)

En el caso de un umbral 0, todo esta siendo calificado como positivo porque esta a la derecha del umbral, entonces:

		R	
P		α (VP)	β (FP)
		0 (FN)	0 (VN)

Umbral 0 - Clasificar todo como positivo (2/2)

$$\text{Fall-out (FPR)} = \frac{\text{FP}}{\text{FP} + \text{VN}} = \frac{\beta}{\beta + 0} = 1$$

$$\text{Recall (TPR)} = \frac{\text{VP}}{\text{VP} + \text{FN}} = \frac{\alpha}{\alpha + 0} = 0$$

Umbral 1 - Clasificar todo como negativo (1/2)

En el caso de un umbral 1, todo esta siendo calificado como negativo porque esta a la izquierda del umbral, entonces:

		R	
P		0 (VP)	0 (FP)
		α (FN)	β (VN)

Umbral 1 - Clasificar todo como negativo (2/2)

$$\text{Fall-out (FPR)} = \frac{\text{FP}}{\text{FP} + \text{VN}} = \frac{0}{0 + \beta} = 0$$

$$\text{Recall (TPR)} = \frac{\text{VP}}{\text{VP} + \text{FN}} = \frac{0}{0 + \alpha} = 0$$

Curva PR

La curva PR (Precision-Recall) es similar a la ROC pero con métricas diferentes:

- Eje X: Precisión
- Eje Y: Recall

Para cada umbral U:

- Calculamos precisión = $\frac{VP}{VP + FP}$
- Calculamos recall = $\frac{VP}{VP + FN}$

¿Cuándo usar la curva PR?

La curva PR es especialmente útil cuando:

- Tenemos clases desbalanceadas
- Nos interesa más la precisión que los falsos positivos
- El costo de los falsos negativos es alto

Salud

- Priorizar recall sobre precisión
- Usar curvas PR
- No usar accuracy como métrica principal

Justicia

- Usar curvas ROC
- Considerar impacto social
- Evaluar transparencia

Ejemplo Práctico

	R	
P	80	20
	10	90

- Accuracy = $\frac{80+90}{200} = 0.85$
- Precision = $\frac{80}{90} = 0.89$
- Specificity (TNR) = $\frac{90}{100} = 0.90$
- Recall (TPR) = $\frac{80}{100} = 0.80$
- Fall-out (FPR) = $\frac{10}{100} = 0.10$

Ejemplo Práctico

	R	
P	80	20
	10	90

- Miss Rate (FNR) = $\frac{20}{100} = 0.20$
- $F1 = 2 \times \frac{0,89 \times 0,80}{0,89 + 0,80} = 0.84$
- $FDR = \frac{10}{90} = 0.11$
- $FOR = \frac{20}{110} = 0.18$
- $NPV = \frac{90}{110} = 0.82$

Implementación en Python

```
1 from sklearn.metrics import accuracy_score, precision_score
2 from sklearn.metrics import recall_score, f1_score
3 from sklearn.metrics import confusion_matrix
4 from sklearn.metrics import roc_curve, auc
5 from sklearn.metrics import precision_recall_curve
6
7 # Métricas básicas
8 accuracy = accuracy_score(y_true, y_pred)
9 precision = precision_score(y_true, y_pred)
10 recall = recall_score(y_true, y_pred)
11 f1 = f1_score(y_true, y_pred)
12
13 # Matriz de confusión
14 conf_matrix = confusion_matrix(y_true, y_pred)
15
16 # Curva ROC
17 fpr, tpr, _ = roc_curve(y_true, y_pred_proba)
18 roc_auc = auc(fpr, tpr)
19
20 # Curva PR
21 precision, recall, _ = precision_recall_curve(y_true, y_pred_proba)
```