

Curvas ROC y PR

Aprendizaje Automático

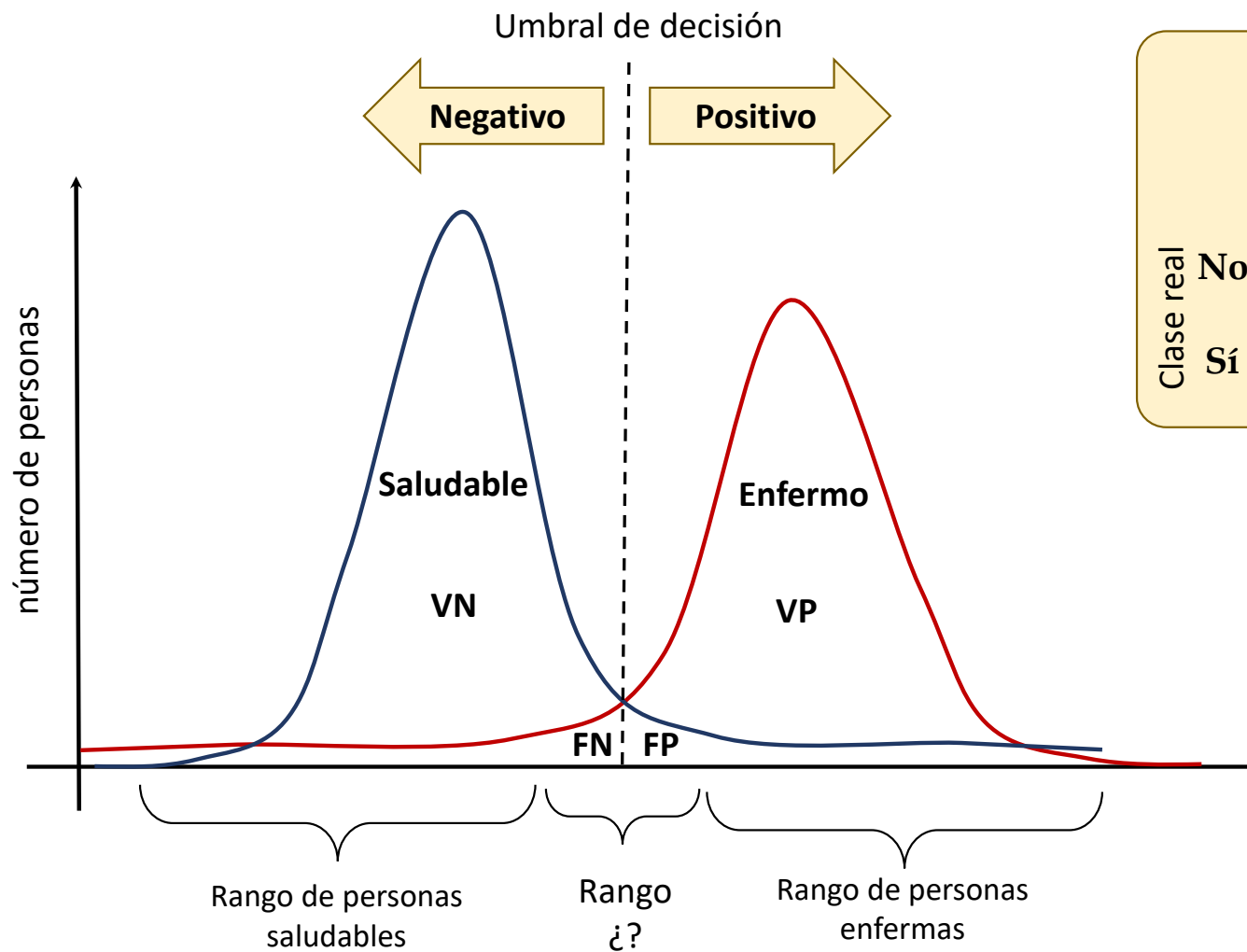


Tecnológico
de Monterrey

Dr. Luis Eduardo Falcón Morales

ITESM

Campus Guadalajara



		¿Enfermo?	
		Predicción	
Clase real	No	VN	FP
	Sí	FN	VP

Exactitud

De los resultados de la matriz de confusión, definimos ahora la exactitud de manera más formal:

Exactitud o tasa de éxito:

$$exactitud = \frac{VP + VN}{VP + VN + FP + FN}$$

Tasa de error, es decir, tasa de clases mal clasificadas:

$$tasa\ de\ error = 1 - exactitud$$

A la *exactitud* algunos autores también la llaman *precisión*, término que no usaremos en este curso para evitar confusión con el término definido más adelante.

falsas alarmas

		Predicción del modelo: ¿Pertenece a la clase positiva?	
		No	Sí
Sabemos que realmente: está en la clase positiva	No	Verdadero Negativo (VN)	Falso Positivo (FP)
	Sí	Falso Negativo (FN)	Verdadero Positivo (VP)

Sensibilidad (Recall) y 1 – Especificidad

Estas métricas ayudan a dar un buen seguimiento de los verdaderos positivos y verdaderos negativos, estén estos bien o mal clasificados por nuestro estimador.

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

**Sensibilidad
(recall)**

$$\text{sensibilidad}_{(\text{recall})} = \frac{VP}{VP + FN}$$

Tasa de verdaderos positivos, con respecto a todos los positivos reales

Nos da la proporción de los positivos reales que están bien clasificados.

exhaustividad = *sensibilidad* = recall

**1 – Especificidad
(specificity)**

$$1 - \text{especificidad} = \frac{FP}{VN + FP}$$

Tasa de falsos positivos, con respecto a todos los negativos reales.

tasa de verdaderos negativos: *especificidad* = $\frac{VN}{VN+FP}$

NOTA: Ambas métricas están entre 0 y 1. Se usarán para generar la curva ROC.

Precisión y Exhaustividad

Nos dan una medida sobre qué tan relevante son los resultados dados por un modelo.

**Precisión
(precision)**

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

Tasa de predicciones positivas.

Cuando un modelo predice una clase positiva, ¿qué porcentaje de acierto hay en ello?

Y definimos la **tasa de falsas alarmas**:
 $1 - precision$

**Sensibilidad
(recall)**

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

(recall)

Tasa de verdaderos positivos. Es decir, ¿qué porcentaje de los elementos de la clase positiva real son pronosticados correctamente?

Observa que:

$$exhaustividad = sensibilidad = recall$$

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

NOTA: Ambas métricas están entre 0 y 1. Se usarán para generar la curva PR.

```

falcon@falcon-dltest: ~/Simples_CNN/Simple_LeNet_MNIST
3062.8449041843414
[INFO] evaluating network...

```

	precision	recall	f1-score	support
0	0.99	0.99	0.99	1677
1	0.98	1.00	0.99	1935
2	0.99	0.98	0.99	1767
3	0.97	0.99	0.98	1766
4	1.00	0.99	0.99	1691
5	0.98	0.98	0.98	1653
6	1.00	0.98	0.99	1754
7	0.98	0.99	0.99	1846
8	0.98	0.97	0.98	1702
9	0.98	0.98	0.98	1709
micro avg	0.98	0.98	0.98	17500
macro avg	0.98	0.98	0.98	17500
weighted avg	0.98	0.98	0.98	17500

```

(micai) falcon@falcon-dltest:~/Simples_CNN/Simple_LeNet_MNIST

```

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

El reporte estándar que trae Python:

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

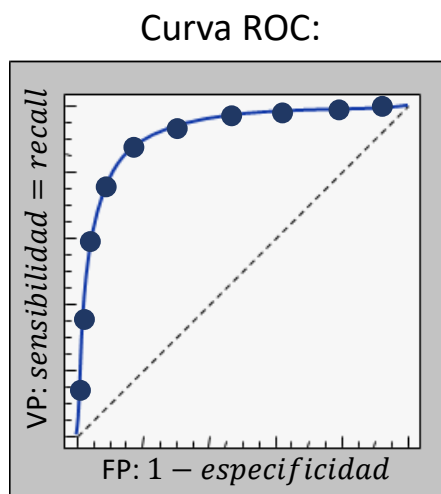
$$sensibilidad (recall) = \frac{VP}{VP + FN}$$

exhaustividad = *sensibilidad* = recall

$$\frac{1}{f1score} = \frac{1}{2} \left(\frac{1}{precision} + \frac{1}{recall} \right)$$

Curva ROC / AUC (Característica Operativa del Receptor / Área Bajo la Curva)

- La curva ROC nos proporciona una representación gráfica de la $(1 - \text{especificidad})$ contra la *sensibilidad*.
- También podemos decir que es la representación gráfica de la tasa de falsos positivos contra la tasa de verdaderos positivos.
- Se puede interpretar como una relación de Costo-Beneficio, ya que se busca encontrar un balance adecuado entre la tasa de falsos positivos (costo), frente a la tasa de verdaderos positivos (beneficio).



		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

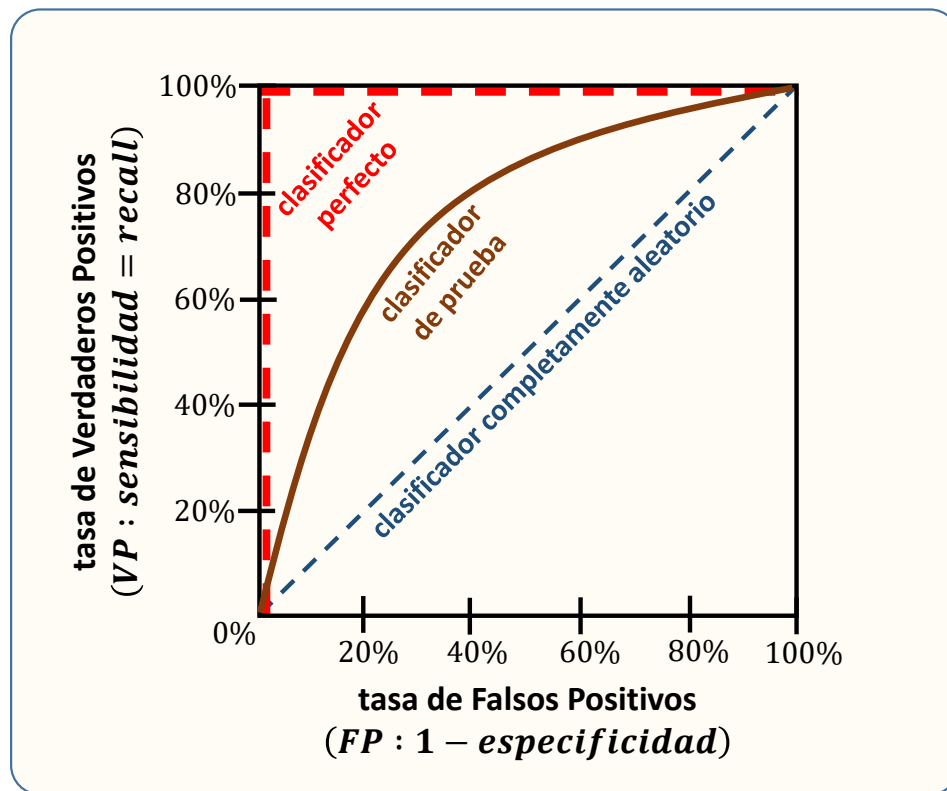
$$\text{especificidad} = \frac{VN}{FP + VN}$$

$$1 - \text{especificidad} = \frac{FP}{FP + VN}$$

$$\text{sensibilidad} = \frac{VP}{VP + FN}$$

$$\text{exhaustividad} = \text{sensibilidad} = \text{recall}$$

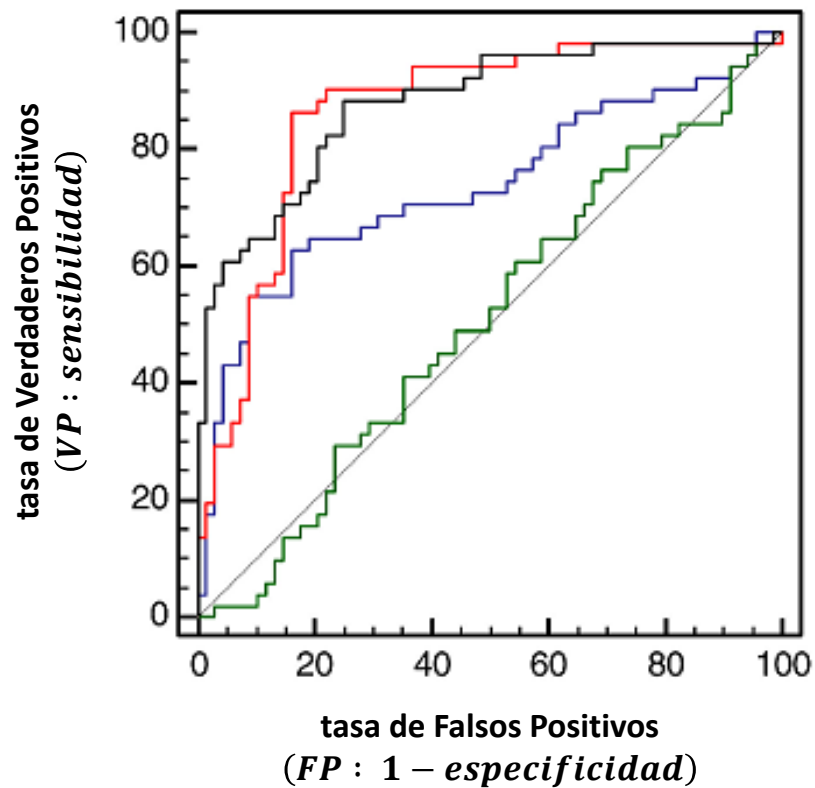
- Las curvas ROC tratan de incrementar la tasa de verdaderos positivos, disminuyendo la tasa de falsos positivos.



		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

Mientras más distante esté del clasificador de no determinación (clasificador completamente aleatorio), se tiene un mejor desempeño del modelo.

Curva ROC / Análisis ROC / Espacio ROC / Área AUC



- El análisis ROC a través de las gráficas en el llamado espacio ROC nos permite comparar el desempeño de diferentes modelos.
- **AUC: área bajo la curva.** Medida del desempeño del modelo dada por el valor del área entre el clasificador completamente aleatorio y la curva de prueba.
- El AUC toma valores entre 0.5 (no discrimina) y 1 (discrimina de manera excelente).

Clases Desbalanceadas - Igualmente Importantes FN y FP
<métrica G-mean>

Sensibilidad (recall)

$$\text{sensibilidad} = \frac{VP}{VP + FN}$$

Tasa de verdaderos positivos

Especificidad (specificity)

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

Tasa de verdaderos negativos

$$G\text{-mean} = \sqrt{\text{Recall} \times \text{Specificity}}$$

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

Media geométrica de
sensibilidad y especificidad:
un valor bajo en cualquier
de los dos, implica un valor
bajo para G-mean.
 $0 \leq G\text{-mean} \leq 1$

Si las clases son no balanceadas, al menos un 80%-20% aproximadamente, conviene usar la métrica *G-mean* en lugar del *accuracy*. Esta métrica es independiente de la distribución de los casos entre clases.

Métricas de Desempeño

Exactitud o tasa de éxito:

$$exactitud = \frac{VP + VN}{VP + VN + FP + FN}$$

Precisión (precision)

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

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

Métrica G-mean

$$G-mean = \sqrt{recall \times especificidad}$$

Especificidad (specificity)

$$especificidad = \frac{VN}{VN + FP}$$

Valor_F

$$f1\text{-score} = \frac{2VP}{2VP + FP + FN}$$

Exhaustividad / Sensibilidad
(recall / sensitivity)

$$exhaustividad\ (recall) = \frac{VP}{VP + FN}$$

Curva PR (precision – recall)

- En los casos de clases desbalanceadas, la curva ROC no es lo más adecuado ya que generalmente proporciona resultados demasiado optimistas.
- Mientras mayor desbalanceadas estén las clases, mayor el sesgo que pueda tener la curva ROC debido a la poca cantidad de datos de la clase positiva.
- La alternativa para clases muy desbalanceadas es la gráfica de la curva PR.

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

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

Tasa de predicciones positivas

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

Tasa de verdaderos positivos

Sensibilidad=Exhaustividad=Recall

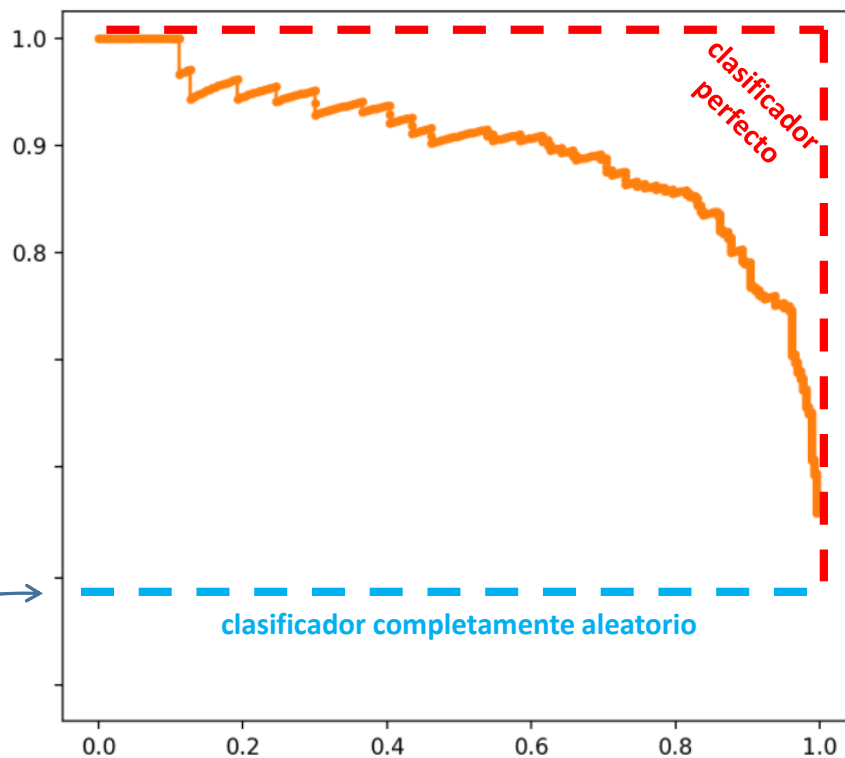
Curva PR : (precision – recall)

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

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

tasa de Predicciones Positivas
($FP : 1 - especificidad$)

Porcentaje de datos de la clase positiva real (clase minoritaria).
Para el caso balanceado esta línea estaría en 0.5.



tasa de Verdaderos Positivos
($VP : sensibilidad = recall$)

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

curva ROC vs curva PR

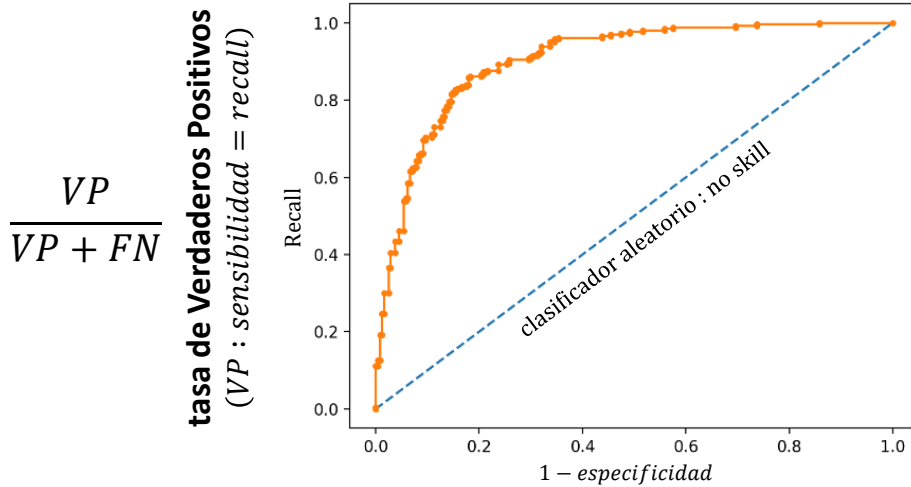
ROC: Receiver Operating Characteristic
PR : Precision - Recall

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

		Predicción	
		No(0)	Sí(1)
Clase real	No(0)	VN	FP
	Sí(1)	FN	VP

Se centra en la clase minoritaria.

Curva ROC:

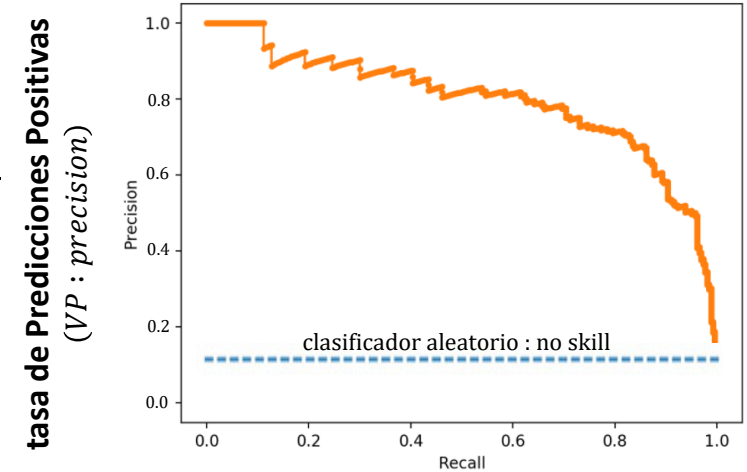


Se centra en ambas clases.

tasa de Falsos Positivos
(FP : 1 - especificidad)

$$\frac{FP}{FP + VN}$$

Curva PR:



tasa de Verdaderos Positivos
(VP : sensibilidad = recall)

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