

Maestría en Inteligencia Artificial Aplicada

Curva ROC



Tecnológico
de Monterrey

Dr. Luis Eduardo Falcón Morales

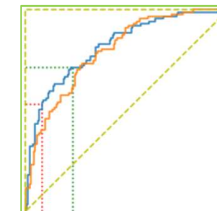
ITESM

Campus Guadalajara

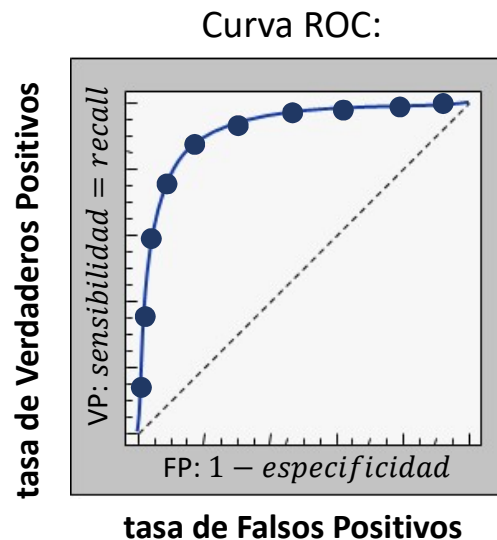
Curva ROC / AUC (Característica Operativa del Receptor / Área Bajo la Curva)

- Las curvas ROC (Receiver Operating Characteristic) han sido utilizadas en gran diversidad de áreas de la ciencia y se ha extendido su uso en el área de aprendizaje automático.
- Es una representación visual del desempeño de un clasificador binario.
- En general son utilizadas para comparar el desempeño de clasificadores binarios, y se pueden extender a varias clases con el criterio *uno contra todos*. Así, si tienes k clases, se tendrían k curvas ROC.
- Las curvas ROC se empezaron a utilizar durante la segunda guerra mundial para la detección de aviones, submarinos o en general de objetos enemigos mediante el radar.

Verdadero Negativo (VN)	Falso Positivo (FP)
Falso Negativo (FN)	Verdadero Positivo (VP)



- 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).



		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN

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

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

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

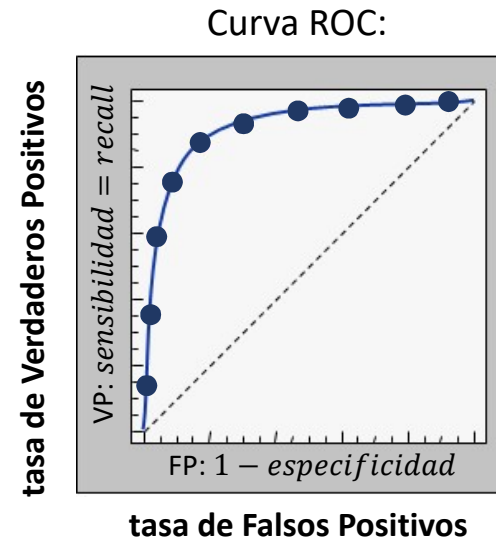
tasa de verdaderos positivos: $\text{sensibilidad} = \frac{VP}{VP + FN}$

tasa de falsos positivos: $1 - \text{especificidad} = \frac{FP}{FN + FP}$

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

Observa que la curva ROC grafica información con respecto a la Clase Real, es decir, “lo mal clasificado del conjunto de las clases negativas reales”, contra “lo bien clasificado del conjunto de las clases positivas reales”.

tasa de verdaderos positivos: $sensibilidad = \frac{VP}{VP+FN}$

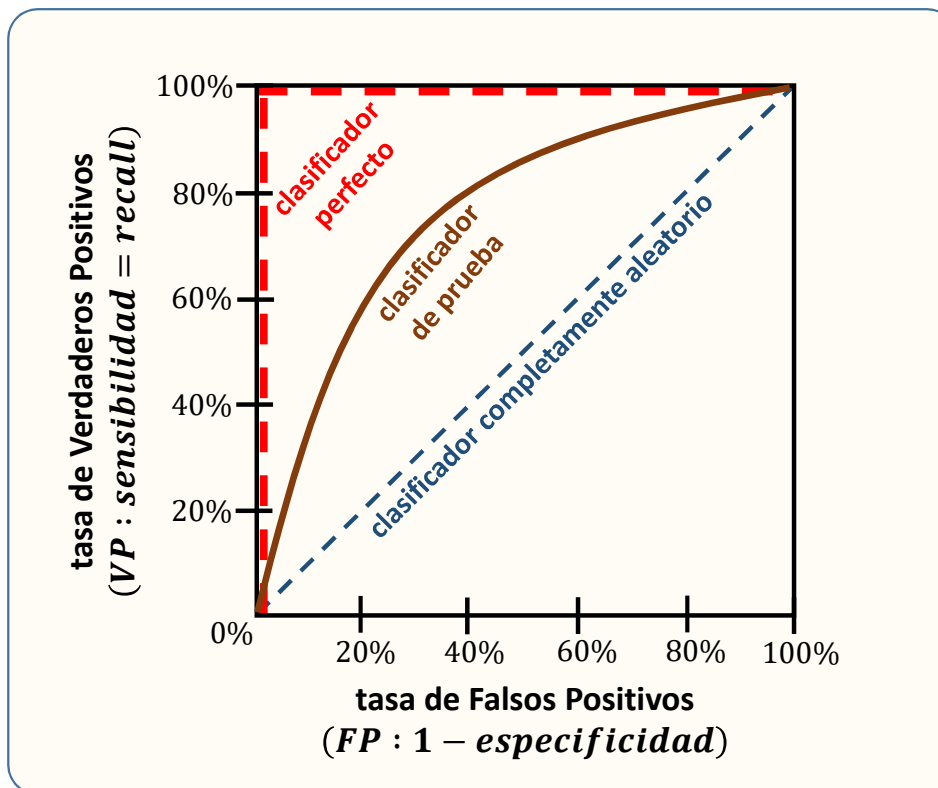


		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN

tasa de falsos positivos: $1 - especificidad = \frac{FP}{VN+FP}$

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

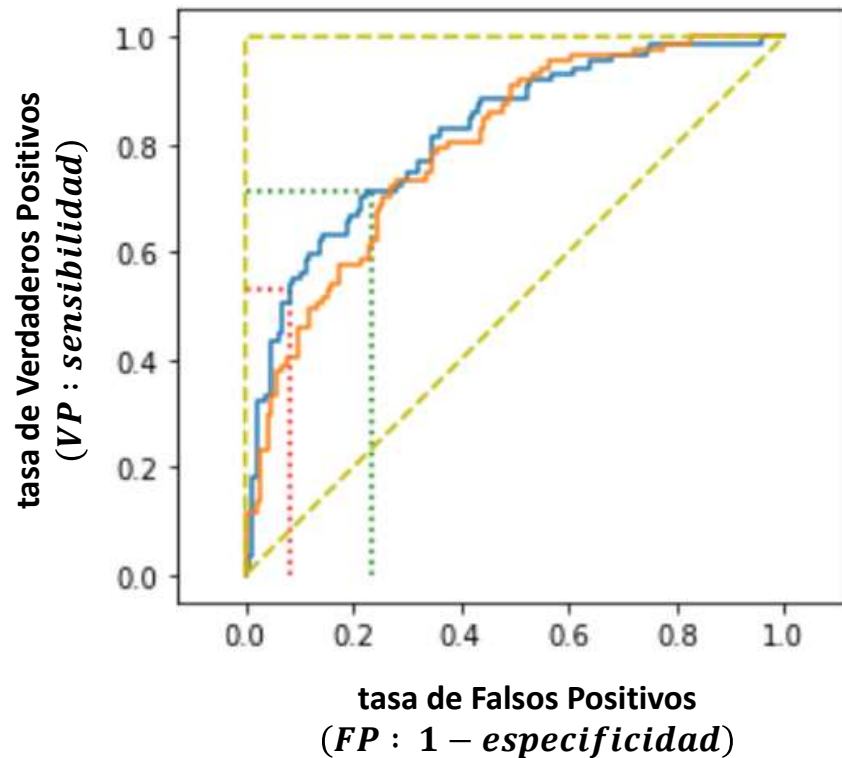
De manera general las gráficas de las curvas ROC tienen la siguiente forma:



		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN

- Mientras más arriba y distante esté del clasificador de no determinación (clasificador completamente aleatorio), mejor desempeño del modelo.
- Valores abajo del clasificador de no determinación indican los peores desempeños, inclusive peor que una decisión tomada de manera aleatoria.

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.
 - Observa que un modelo puede tener mejores resultados en cierto rango, pero ser superado por otros modelos en otros rangos.
 - Por este traslape entre diferentes curvas ROC, la gráfica puede ser engañosa, por lo que conviene incluir siempre el valor AUC.
- **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).

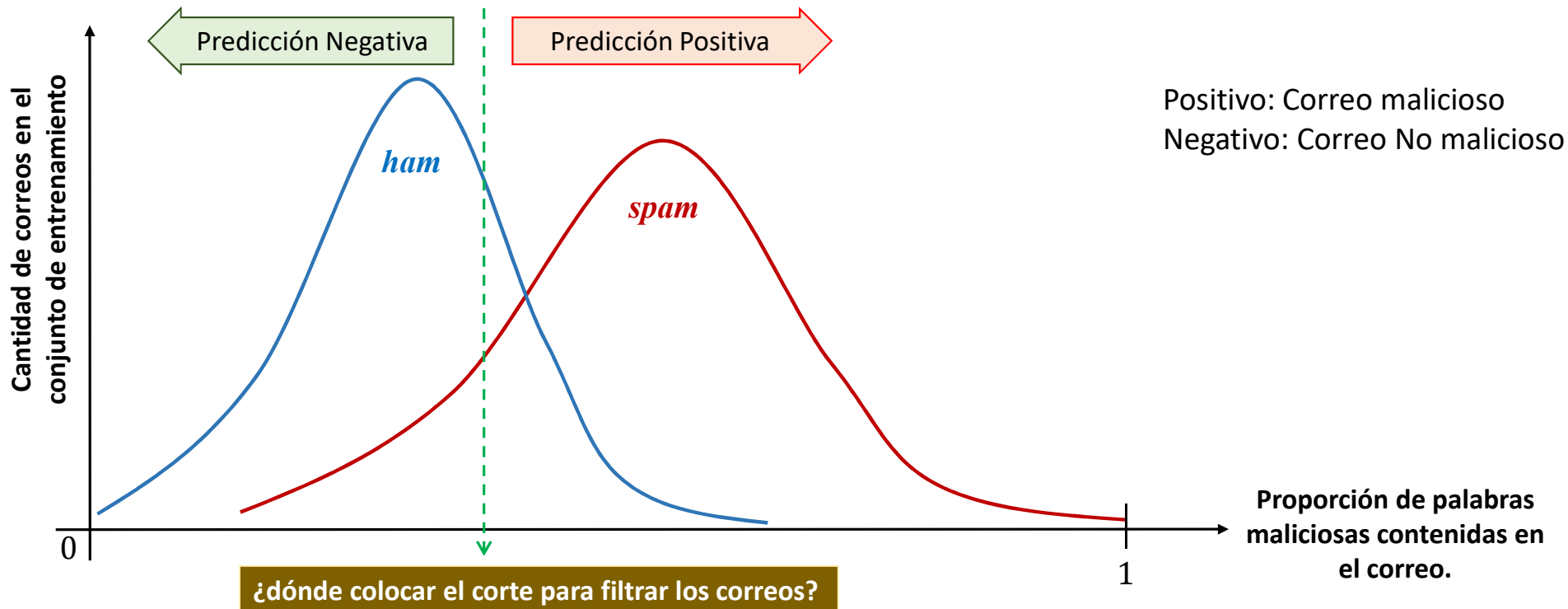
Las curvas ROC se utilizan frecuentemente para comparar dos o más algoritmos, ya que nos proporciona medidas globales, independientemente del punto de corte.

Existen varios criterios para la interpretación del valor AUC, pero en general se pueden considerar los siguientes criterios para evaluar alguna prueba:

- $AUC \in (0.97, 1]$ prueba excelente
- $AUC \in (0.9, 0.97]$ muy buena prueba
- $AUC \in (0.7, 0.9]$ prueba aceptable
- $AUC \in (0.6, 0.7]$ prueba regular
- $AUC \in [0.5, 0.6]$ la prueba no discrimina

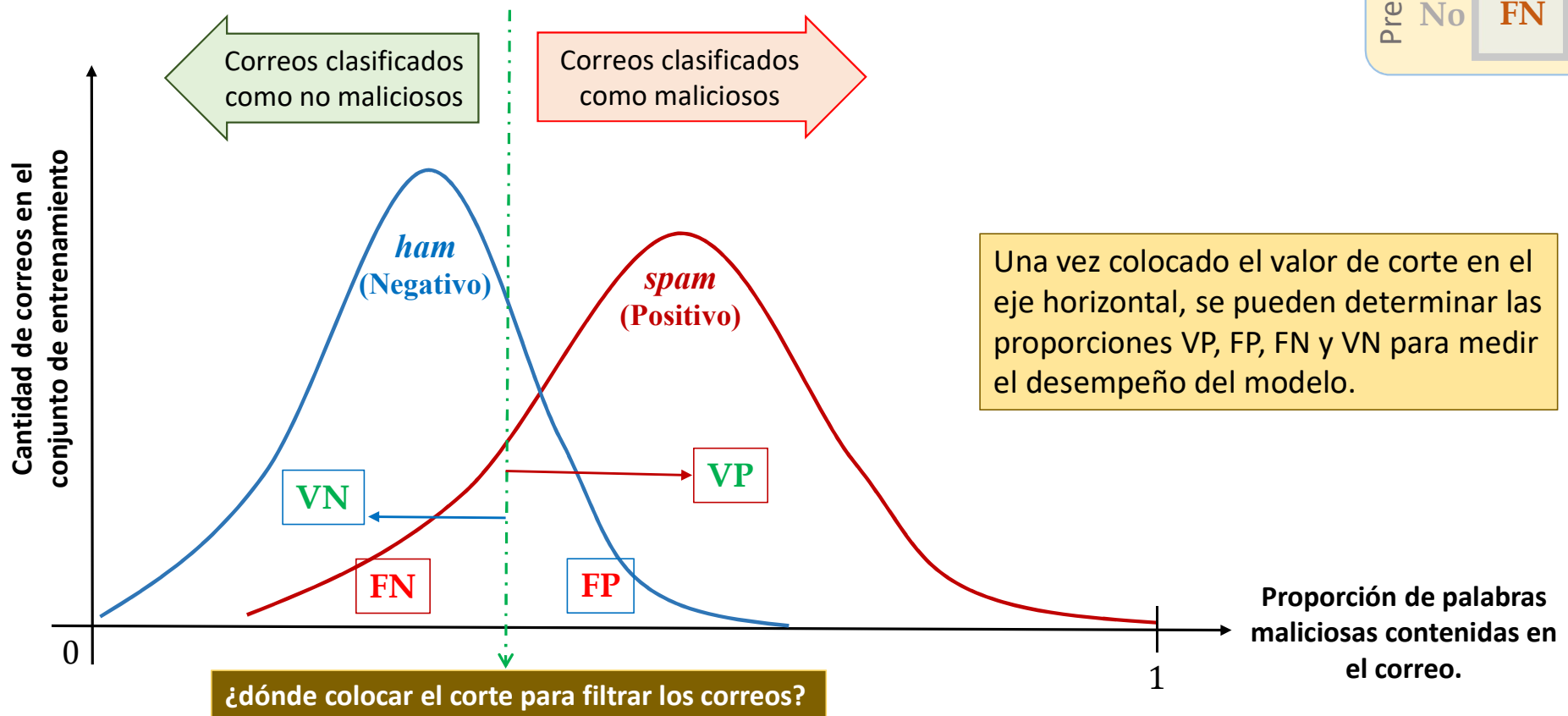
Ejemplo: Interpretación gráfica de la curva ROC:

Supongamos que tenemos un clasificador que determina si un correo es o no malicioso (*spam*) con base a la detección de cierta cantidad de palabras consideradas maliciosas. Se cuenta con un banco de palabras maliciosas. Al llegar un nuevo correo se buscarán y contarán cuántas de estas palabras maliciosas están contenidas en el correo y con base a dicho valor se sacará la proporción de dichas palabras. Esta proporción será entonces un valor entre 0 y 1, y lo ubicaremos en el eje horizontal de nuestro plano cartesiano. En el eje vertical se indicará el total de correos en el conjunto de entrenamiento con dicha proporción de palabras maliciosas. Se desea determinar el valor de dicha proporción a partir del cual filtrar los correos de la manera *más efectiva*.



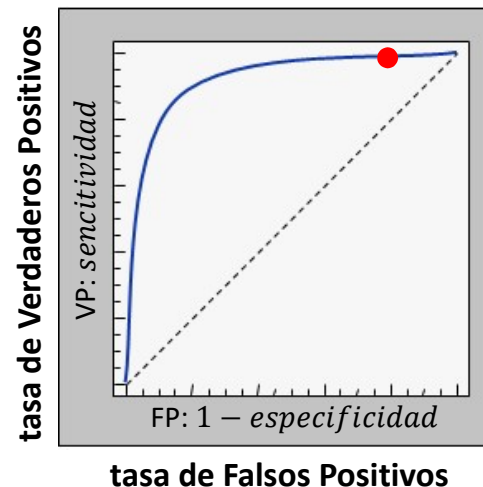
Resultado Positivo: Correo malicioso
Resultado Negativo: Correo No malicioso

		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN

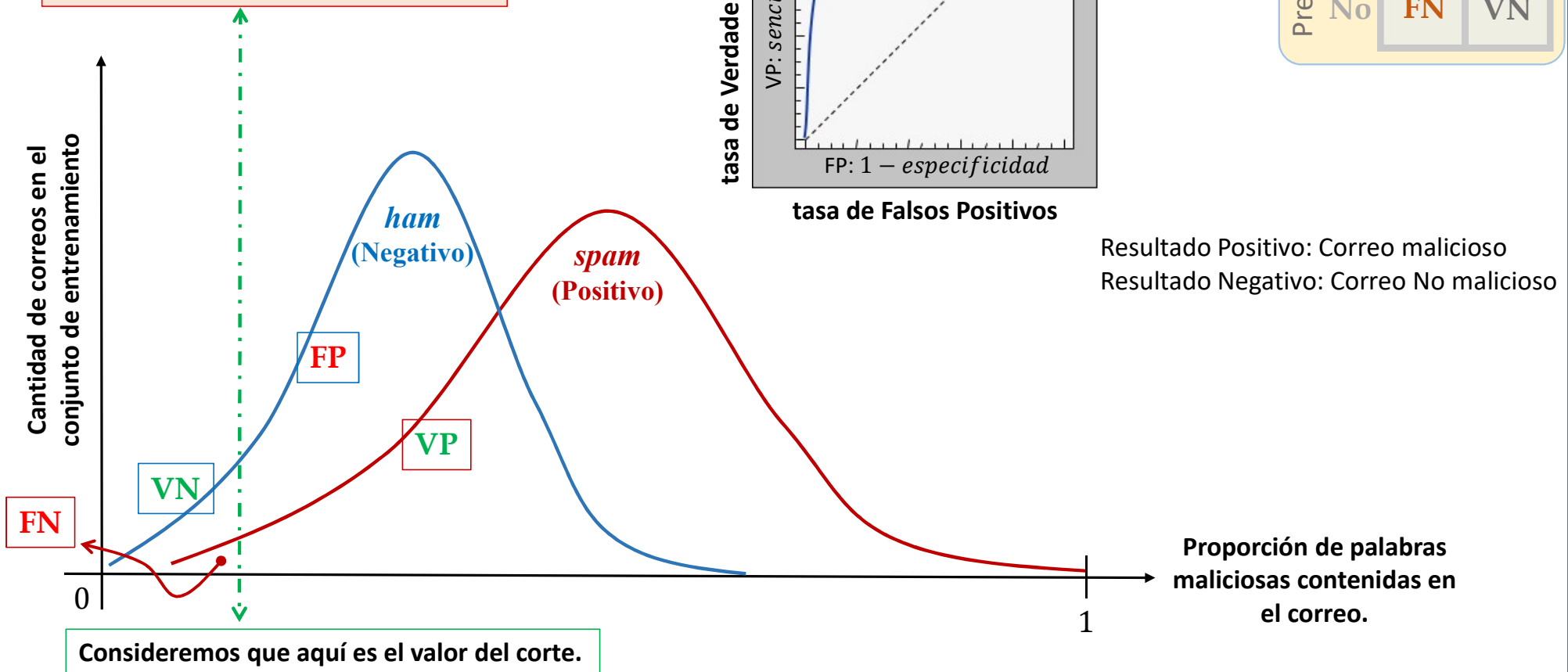


Si el corte fuera aquí, no estaría dejando pasar casi ningún correo: detecta la mayoría de los *spams*, a costa de filtrar también casi todos los correos *ham*.

Valor en la Curva ROC:

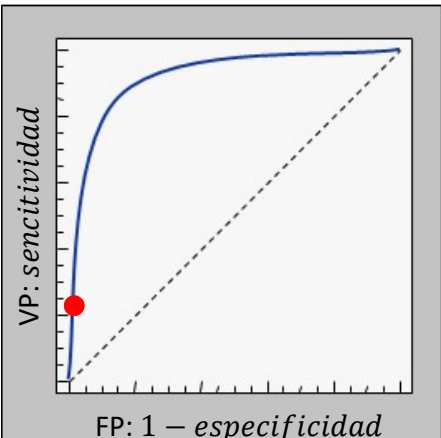


		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN



Valor en la Curva ROC:

tasa de Verdaderos Positivos



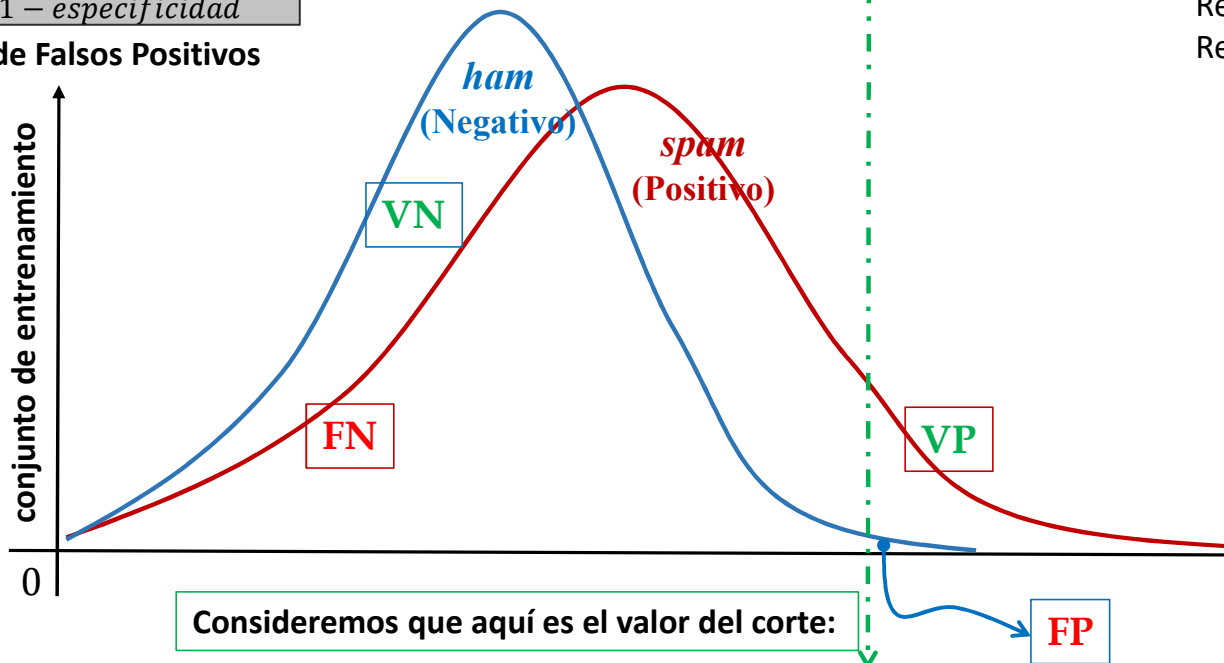
tasa de Falsos Positivos

Si el corte fuera aquí, estaría dejando pasar casi todo: detecta muy bien los *hams*, a costa de dejar pasar también casi todos los correos *ham*.

		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN

Resultado Positivo: Correo malicioso
Resultado Negativo: Correo No malicioso

Cantidad de correos en el conjunto de entrenamiento

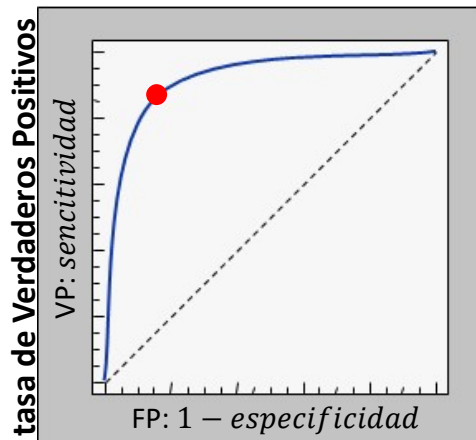


Consideremos que aquí es el valor del corte:

FP

Proporción de palabras maliciosas contenidas en el correo.

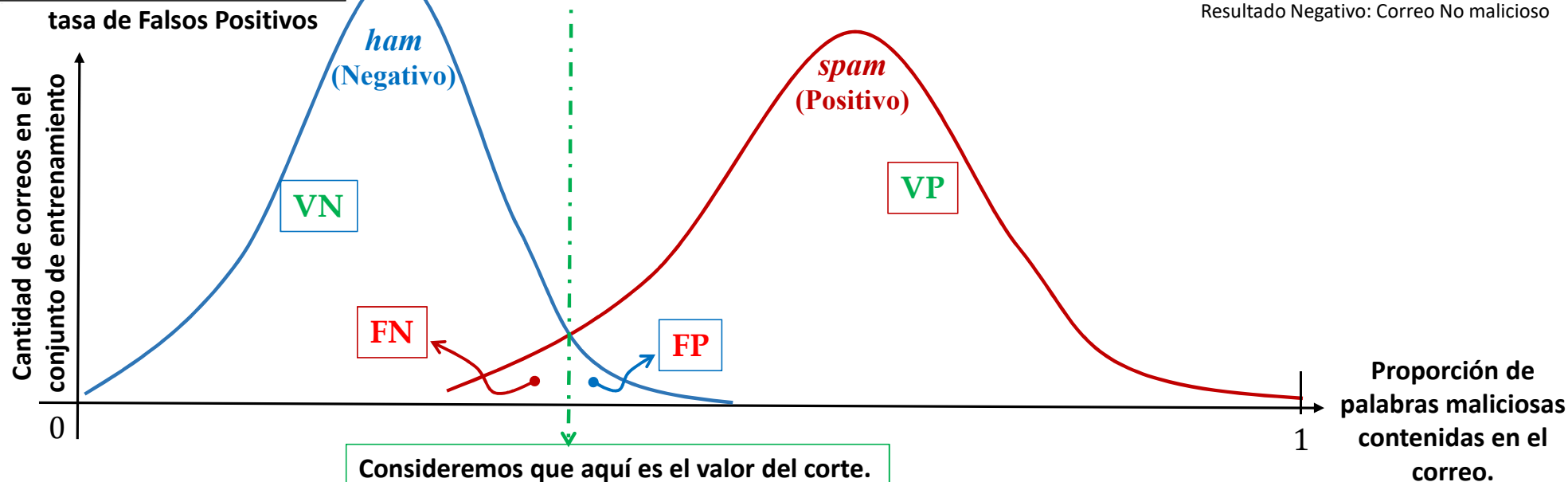
Valor en la Curva ROC:



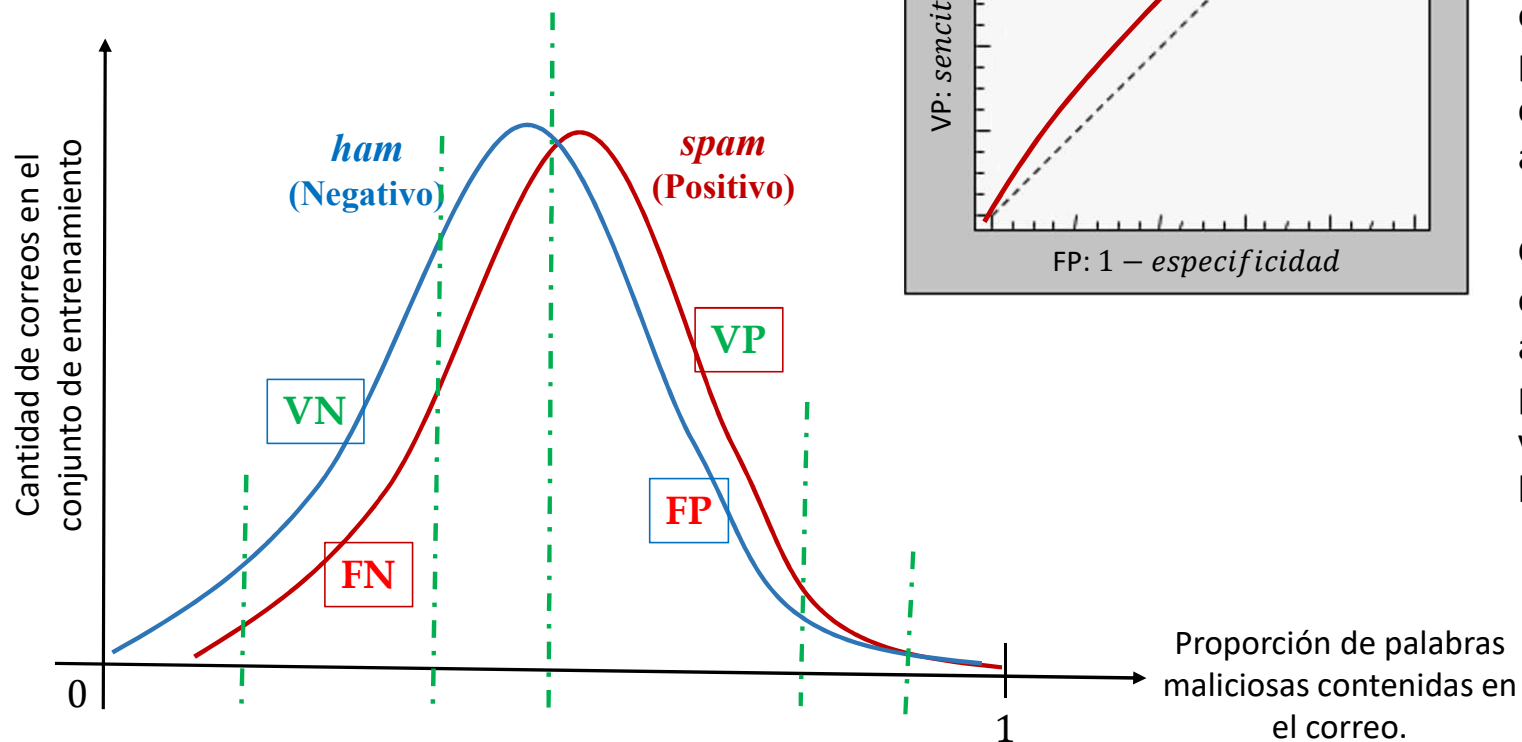
Esto es lo que se busca en un buen clasificador: que detecte casi todos los *spams* (VP), a la vez que es mínima la cantidad de correos *ham* que los clasificó mal (FP) .

		Clase real	
		Sí	No
Predicción	Sí	VP	FP
	No	FN	VN

Resultado Positivo: Correo malicioso
Resultado Negativo: Correo No malicioso



Ejemplo de una curva ROC de un algoritmo de detección que tiene un muy mal desempeño:



En un mal clasificador casi no hay diferencia entre la conclusión de la prueba y una decisión completamente aleatoria.

Gráficamente observamos que las áreas VP y FP son muy parecidas para cualquier valor del corte en el eje horizontal.

Ejemplo gráfico: Predicciones con los datos de Prueba: *pima indians diabetes* de la UCI:
<https://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes>

