

MACHINE LEARNING & DEEP LEARNING

MSC. RENZO CLAURE

1



El desbalance

MSC. RENZO CLAURE

2

¿Por qué es un problema?

Porque los modelos aprenden a **predecir la clase mayoritaria**, no a entender los patrones de la minoritaria. Esto genera:

- Alta precisión global (accuracy), pero baja capacidad para detectar la clase rara (recall bajo).
- Riesgo de sobreajuste a la clase dominante.
- Falsas métricas de éxito (confianza en modelos que en realidad no sirven).

Un modelo que predice siempre la clase 0 tiene 95% de accuracy... pero **no detecta ningún caso real** de la clase 1.

Clase	Frecuencia
0	950
1	50

MSC. RENZO CLAURE

3

Opciones para implementar balanceo

1. Sobremuestreo (Oversampling)
 - Random OverSampling: copias aleatorias de datos minoritarios.
 - SMOTE: crea nuevas instancias sintéticas basadas en interpolación de vecinos.
 - ADASYN: similar a SMOTE pero enfocado en puntos más difíciles.
2. Submuestreo (Undersampling)
 - Random UnderSampling: elimina ejemplos al azar.
 - Tomek Links: elimina ejemplos "conflictivos".
 - Cluster Centroids: agrupa y resume los datos de la clase mayoritaria.
3. Cambiar el algoritmo o su penalización
 - Algunos modelos permiten poner pesos a las clases para que aprendan a prestar más atención a las minoritarias.
4. Cambiar la métrica de evaluación
 - Precision, Recall, F1-score
 - Matriz de confusión
 - AUC-ROC (especialmente útil en desbalance)

MSC. RENZO CLAURE

4



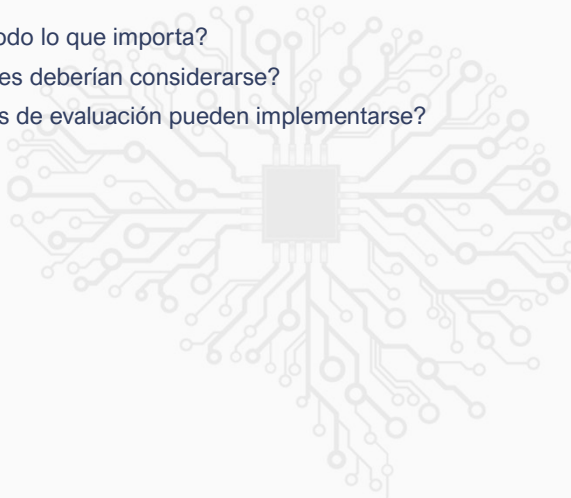
MÉTODOS DE EVALUACIÓN DE MODELOS

MSC. RENZO CLAURE

5

Métodos de evaluación de modelos

- ¿Es la exactitud todo lo que importa?
- ¿Qué otros factores deberían considerarse?
- ¿Qué otras formas de evaluación pueden implementarse?



MSC. RENZO CLAURE

6

Métodos de evaluación de modelos

- Otras medidas que pueden implementarse, que pueden depender del negocio:
 - Satisfacción del usuario
 - Incremento en ventas cruzadas
 - Monto de ingresos generados
 - Monto de costos ahorrados
 - Incremento en la permanencia de los usuarios
 - Diagnósticos exitosos de pacientes con cáncer
- Se deben elegir las formas de medición más apropiadas al inicio del análisis

MSC. RENZO CLAURE

7

Construcción de predictores tontos

¿qué pasa si nuestro modelos predice solo por suerte?

- ¿Qué es un predictor tonto?
- Para variables discretas:
 - Por la clase más frecuente
 - Aleatorio, según la distribución de los datos de entrenamiento
 - Aleatorio uniforme, sin importar los datos de entrenamiento predecir cualquier valor aleatoriamente
 - Clase única, predecir siempre la misma clase
- En el caso de variables continuas:
 - Media
 - Mediana
 - Quantil
 - Constante

MSC. RENZO CLAURE

8

Construcción de predictores tontos

¿qué pasa si nuestro modelos predice solo por suerte?

- Para qué sirven
 - Baseline
 - Comparativo como el peor escenario
- Por qué un modelo está cerca de la misma efectividad que un predictor tonto?
 - Desbalanceo
 - Características o variables independientes triviales o inefectivas
 - Mala elección de Kernel, o parámetros gamma, lambda o C

MSC. RENZO CLAURE

9

Matriz de confusión

	Predicho Negativo	Predicho Postivo	
Real Negativo	Verdadero Negativo, VN, TN	Falso Positivo, FP (Tipo I)	
Real Positivo	Falso negativo, FN (Tipo II)	Verdadero Positivo, VP, TP	

MSC. RENZO CLAURE

10

Matriz de confusión

	Predicho Negativo (0)	Predicho Positivo (1)	
Real Negativo	90	0	
Real Positivo	10	0	
	100	0	$(90+0)/100=0,9$

MSC. RENZO CLAURE

11

Matriz de confusión

	Predicho Negativo (0)	Predicho Positivo (1)	
Real Negativo	82	2	
Real Positivo	8	8	
	90	10	$(82+8)/100=0,90$

MSC. RENZO CLAURE

12

Matriz de confusión

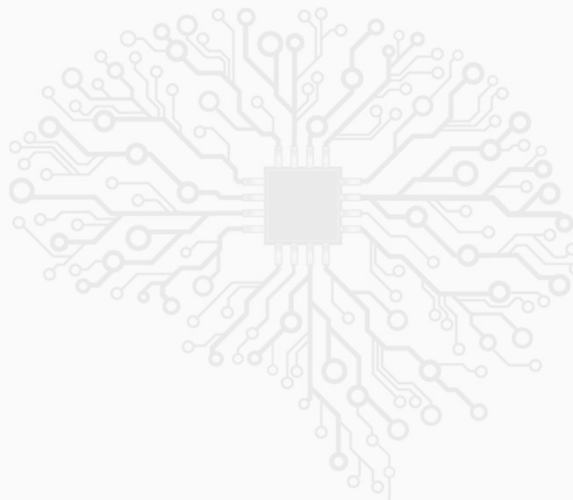
	Predicho Negativo (0)	Predicho Positivo	
Real Negativo	89	1	
Real Positivo	1	9	
	90	10	$(89+9)/100=0,98$

MSC. RENZO CLAURE

13

Ejemplo: predicción de cáncer

por que la precisión no los es todo, que pasa con el desbalanceo



MSC. RENZO CLAURE

14

Matriz de confusión

MACHINE LEARNING

CASOS TOTALES = P + N		PREDICCIÓN	
		Negativo (PN)	Positivo (PP)
REAL	Negativo (N)	Verdadero Negativo, VN, TN	Falso Positivo, FP (Tipo I)
	Positivo (P)	Falso negativo, FN (Tipo II)	Verdadero Positivo, VP, TP

ACIERTO GLOBAL $ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$

SENSIBILIDAD
RECALL, TPR $TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR$

PRECISIÓN $PPV = \frac{TP}{TP + FP} = 1 - FDR$

ESPECIFICIDAD $TNR = \frac{TN}{N} = \frac{TN}{TN + FP} = 1 - FPR$

FNR $FNR = \frac{FN}{P} = \frac{FN}{FN + TP} = 1 - TPR$

F1 SCORE $F_1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$ MSC. RENZO CLAURE

CARACTERÍSTICAS

- Contrasta modelos vs realidad
- Ayuda a identificar las fuentes de desviación
- Tipos de error
- Permite extraer varios indicadores de precisión
 - Sensibilidad: Acierto de positivos
 - Especificidad: Acierto de negativos

15

Matriz de confusión

Recall = 0,9

Prec = 0,73

VALOR REAL		VALOR PREDICHO		
		N	P	Total
N	Cantidad	21	8	29
	% total			0,48
P	Cantidad	3	28	31
	% total			0,52
Total	Cantidad	24	36	60
	% total	0,4	0,6	0,81

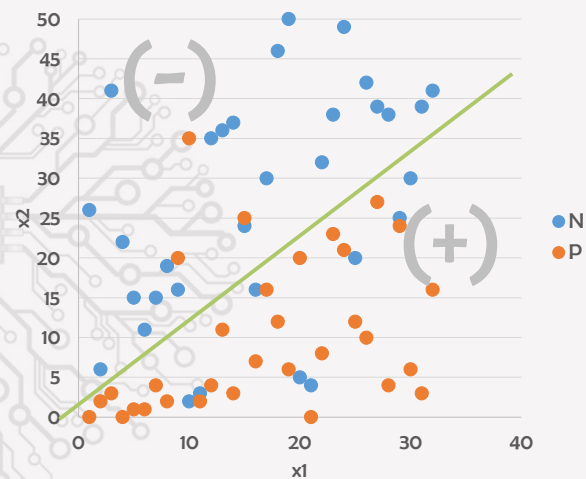
ACIERTO GLOBAL $ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$

SENSIBILIDAD
RECALL, TPR $TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR$

PRECISIÓN $PPV = \frac{TP}{TP + FP} = 1 - FDR$

CASOS TOTALES = P + N		PREDICCIÓN	
		Negativo (PN)	Positivo (PP)
REAL	Negativo (N)	Verdadero Negativo, VN, TN	Falso Positivo, FP
	Positivo (P)	Falso negativo, FN	Verdadero Positivo, VP, TP

MSC. RENZO CLAURE



16

Matriz de confusión

Recall = 0,97

Prec = 0,7

VALOR REAL		VALOR PREDICHO		
		N	P	Total
N	Cantidad	8	21	29
	% total			
P	Cantidad	1	30	31
	% total			
Total	Cantidad	9	51	60
	% total			0,63

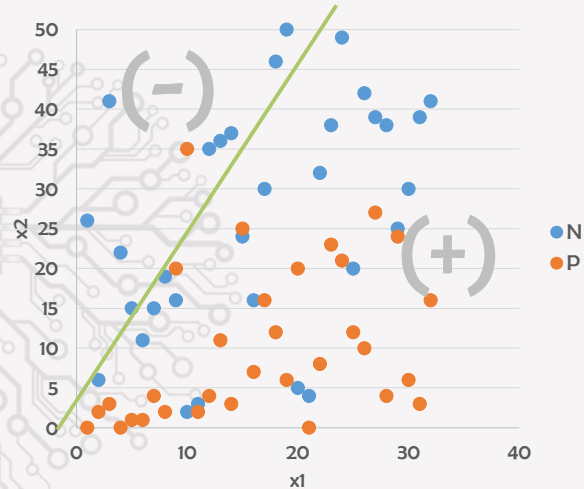
ACIERTO GLOBAL $ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$

SENSIBILIDAD
RECALL, TPR $TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR$

PRECISIÓN $PPV = \frac{TP}{TP + FP} = 1 - FDR$

CASOS TOTALES = P + N		PREDICCIÓN	
		Negativo (PN)	Positivo (PP)
REAL	Negativo (N)	Verdadero Negativo, VN, TN	Falso Positivo, FP
	Positivo (P)	Falso negativo, FN	Verdadero Positivo, VP, TP

MSC. RENZO CLAURE



17

Modelos según importancia de medición

recall vs precisión

- Modelos que se enfocarían en un mayor **Recall**
 - Detección de Cancer
 - Retener clientes de alto valor
 - ...
- Modelos orientados en una mayor **Precisión**
 - Predictor de culpabilidad
 - ...
 - ...

MSC. RENZO CLAURE

18

F1 score

media armónica

- Se basa en la media armónica de los resultados:
- Compara el Recall Vs la Precisión
 - $F1 = 2 \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}}$
- Es un medida global de precisión mas equilibrada

MSC. RENZO CLAURE

19

Revisión de la matriz de confusión

- NB_10

MSC. RENZO CLAURE

20

COMO AFECTA LA PROBABILIDAD A LA PRECISIÓN

PROBABILIDAD DE LA CLASIFICACIÓN

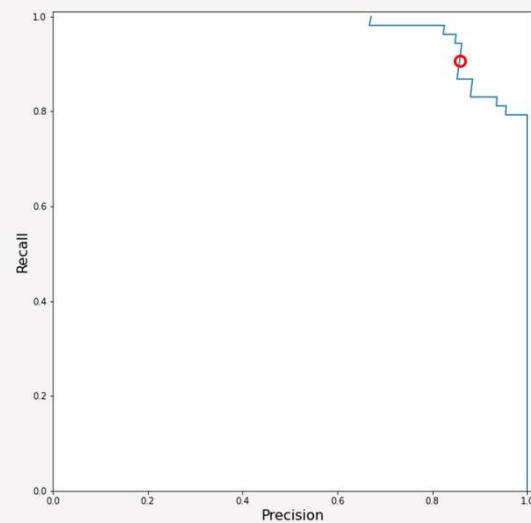
- Los modelos, por defecto, típicamente asignan el valor de 1 o positivo si la probabilidad es mayor al 50%
- Subir o bajar este umbral produce afectaciones en los resultados e indicadores, ¿De qué modo?
- Al subir el umbral, somos más exigentes y no queremos arriesgar la mala clasificación de negativos. Es decir, por ejemplo, si subimos el umbral de probabilidad al 70% exigimos que solo aquellos con una probabilidad superior al 70% sean considerados como 1, + o éxito. Ejemplos: Como afecta a la: Detección de Cancer, Venta cruzada
- Los modelos no siempre cuentan con este indicador

MSC. RENZO CLAURE

21

Precisión vs recall (sensibilidad)

- El punto óptimo Precisión=1, Recall=1
- Se modulan los umbrales para mejorar al máximo ambos scores



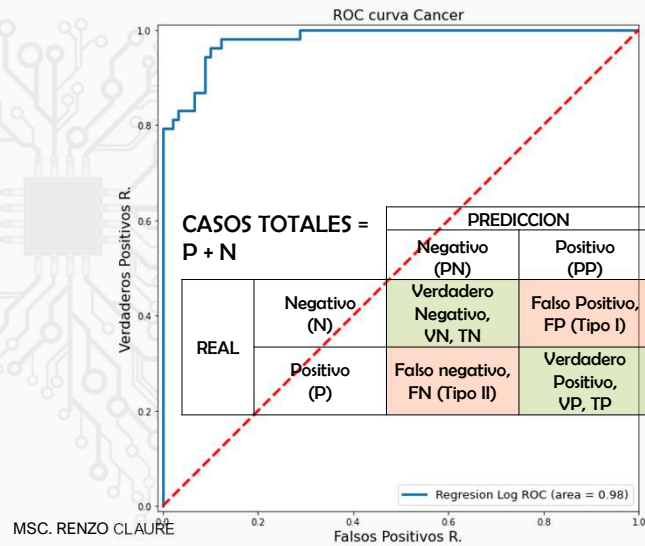
MSC. RENZO CLAURE

22

Curvas roc

(receiver operating characteristic)

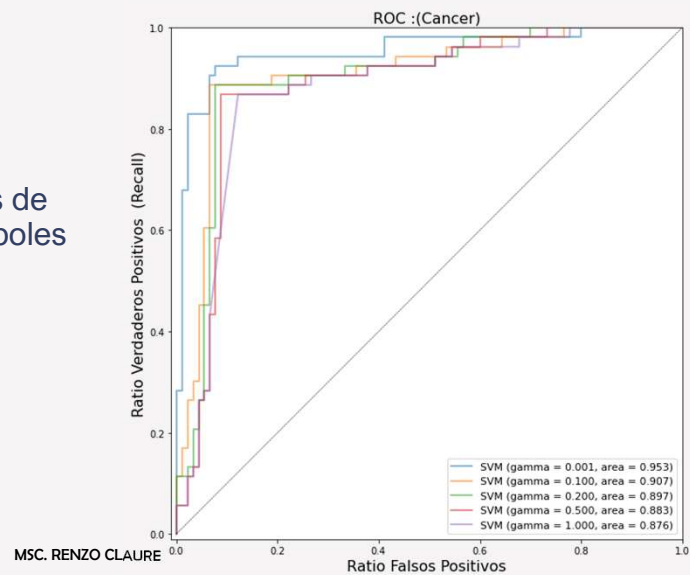
- Contrasta los ratios de Falso Positivo y Verdadero Positivo
- El punto óptimo se da cuando el ratio VP=1 y el FP=0
- La curva que más se aleja de la diagonal principal representa al mejor modelo
- ¿Cuál es el peor modelo?
- ¿Qué representa el área debajo de la curva?
- Sirve para comparar más de 1 modelo



23

Curvas roc comparativas

- ¿Cuál es el mejor modelo?
- ¿Pueden compararse modelos de clase distinta? Por ejemplo Árboles y Regresión Logística



24

Evaluación de multclasificadores

micro y macro recall

REAL	PRONOSTICO	ACIERTO	MACRO RECALL		MICRO RECALL	
P1	P1	1	2	67%	15	75%
P1	P2	0				
P1	P1	1				
P2	P2	1	7	78%		
P2	P2	1				
P2	P2	1				
P2	P2	1				
P2	P2	1				
P2	P2	1				
P2	P3	0				
P2	P2	1				
P2	P2	1				
P2	P1	0				
P3	P2	0	3	60%		
P3	P2	0				
P3	P3	1				
P3	P3	1				
P3	P3	1				
P4	P4	1	3	100%		
P4	P4	1				
P4	P4	1				
				76%		

MSC. RENZO CLAURE

- **Macro RECALL:**
 - Es el promedio de los diferentes niveles
 - Cada nivel o clase tiene el mismo peso
- **Micro RECALL:**
 - Cada caso tiene el mismo peso
 - Las clases más frecuentes tienen mas influencia
- Cuando las clases están igualmente distribuidas son similares
- Si la micro RECALL es más baja que la macro, la mala clasificación del nivel dominante puede estar afectando
- Si es lo contrario al anterior es necesario analizar la clasificación de los niveles menos frecuentes

25

Evaluación de la regresión

- El coeficiente de determinación o r^2 , r cuadrado, mide el grado de ajuste de una curva con respecto a datos reales
- Su principal limitación es que no nos dice si sobre estimamos o subestimamos un valor
- Varía entre 0 y 1
- ¿En que se diferencia del coeficiente de correlación lineal?
- Puede también medirse los errores, para comparar modelos:
 - Error absoluto promedio
 - Error cuadrático medio
 - Error absoluto mediano

MSC. RENZO CLAURE

26

Evaluación de la regresión

- Existen varias opciones para construir los baseline o regresores tontos:
 - Promedio
 - Media
 - Mediana
 - Cuantil



MSC. RENZO CLAURE

27



Validación cruzada

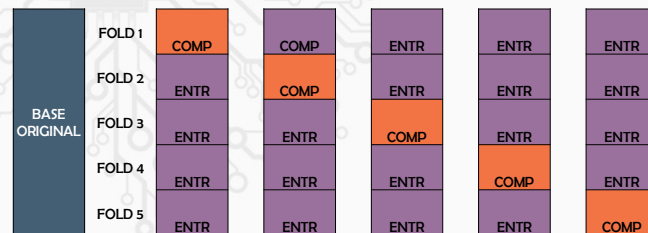
MSC RENZO CLAURE ARACENA

28

Validación cruzada

cross validation

- Es una forma más robusta de medir el rendimiento del modelo
- Se logra a través de la generación de varios grupos (folds), extrayendo de cada uno una muestra de entrenamiento y otra de validación, generando los modelos y obteniendo sus scores
- Se debe asegurar que la proporción de la variable objetivo sea la misma que en la población, sklearn ya soluciona esto



MSC RENZO CLAURE ARACENA

29



Selección de Modelos

MSC. RENZO CLAURE

30

Grid search, optimización de parámetros

- Grid search: Búsqueda de los parámetros óptimos con base en un indicador de rendimiento, usa por defecto accuracy o acierto global.
- Explora cada posible combinación sobre los hyperparametros propuestos.
- Devuelve la mejor combinación, según la métrica establecida.
- Es posible combinarlo o usarlo con Cross Validation, pero es necesario tener cuidado por el enorme consumo de recursos que puede ocasionar si se colcan demasiados hiperparámetros.

MSC. RENZO CLAURE

31

Selección de modelos

- NB_13 SELECCION GRID

MSC. RENZO CLAURE

32

Selección de modelos

conclusiones

- La exactitud, o acierto no refleja toda la información necesaria para un modelo
- En modelos binarios, por ejemplo es más que necesario utilizar en análisis de la Matriz de confusión, recall y precisión especialmente (Falsos Positivos, Falsos Negativos)
- Curvas de aprendizaje, se puede medir como va mejorando la exactitud en función al tamaño de la muestra de aprendizaje y el costo que implica
- Análisis de sensibilidad, como varia la exactitud en función de la modificación de los modelos

MSC. RENZO CLAURE