

Inteligencia Artificial 2020

Informe Actividad Práctica B3:

Validación de modelos



Facultad de Ingeniería
Ingeniería en Informática

Alumno: Iván Gabriel Guerra

Profesor: Juan Francisco Giró



Índice

Introducción	3
Objetivos	3
Requerimientos	3
Condiciones	4
Descripción del problema	4
Datos disponibles	5
Resolución del problema	6
Tecnologías utilizadas	6
Métodos utilizados	6
Curvas de distribución	6
Tablas de contingencia	7
Curvas ROC	9
Índice de Youden	10
Coeficiente de Gini	12
Índice de exactitud	12
Curvas PR	13
Comparación de resultados	15
Ejecución del proyecto terminado	15
Repositorio del proyecto terminado	16
Conclusiones	16



Introducción

Objetivos

- Utilizando un análisis ROC, comparar la capacidad que se obtiene utilizando dos procedimientos diferentes "A" y "B" de filtrado de imágenes, para clasificar y detectar fallas en piezas de fundición de aluminio a partir de su nivel de gris.
- Identificar el modelo que brinde la mayor capacidad de clasificación

Requerimientos

Comparar la aptitud de las imágenes obtenidas con los procedimientos de filtrado "A" y "B", para posibilitar la clasificación de piezas a partir de su nivel de gris, y presentar la información complementaria que justifique la decisión, según se indica:

- 1) Las curvas de distribución de ambas técnicas de filtrado
- 2) Análisis de los procedimientos de clasificación a través de sus curvas ROC e índice de Youden
- 3) Selección del procedimiento de Filtrado más conveniente.
- 4) Identificación del valor de corte más apropiado para clasificar las piezas con la mejor técnica. Considerar para ello que el costo de un "falso positivo" y de un "falso negativo" es el mismo.
- 5) Determinar y representar la curva de exactitud en función del valor de corte de ambas técnicas.
- 6) Representar los gráficos PR para ambas técnicas.



Condiciones

- Los trabajos serán realizados en forma individual.
- Presentar un informe escrito que incluya: Identidad del alumno, objetivos del trabajo, métodos usados, características del problema, particularidades del desarrollo e implementación, resultados, discusión de resultados, limitaciones comprobadas, eventuales dificultades y conclusiones.
- El formato y estructura del documento debe cumplir con las condiciones establecidas para las presentaciones de la materia.
- Nombre del archivo:

IUA-IA_2020_Actividad_B3_Apellido_nombre.pdf

- Fecha de presentación: Viernes 20 de Noviembre de 2020.

Descripción del problema

Para la detección automática de fallas en piezas de fundición de aluminio a través de ensayos no destructivos se aplica un procedimiento que incluye tres etapas:

- 1) la **captura** de una imagen de la zona estudiada
- 2) el **procesamiento** de la imagen, con una secuencia de operaciones tales como aplicación de filtros digitales, integración y mejora del contraste
- 3) la **selección del parámetro característico** que resulte más apropiado para distinguir las piezas sanas y defectuosas.



En este caso el parámetro característico adoptado es el **nivel de gris** de la zona estudiada y es necesario comparar la capacidad de clasificación que se obtiene utilizando dos secuencias diferentes de filtrado de las imágenes, que son denominados procedimientos "A" y "B".

Para evaluar los procedimientos e identificar el más apropiado se utiliza el **análisis ROC**.

Datos disponibles

Los resultados de la clasificación de piezas con los filtrados "A" y "B" se encuentran disponibles en la siguiente tabla:

Gris	Modelo A		Modelo B	
	(-)	(+)	(-)	(+)
0,11	0	0	1	0
0,13	1	0	5	0
0,15	2	0	8	0
0,17	3	0	13	0
0,19	5	0	15	0
0,21	9	0	18	0
0,23	13	0	19	0
0,25	16	0	17	1
0,27	18	0	15	1
0,29	17	1	14	2
0,31	15	3	9	3
0,33	11	4	6	6
0,35	9	5	5	9
0,37	7	6	3	12
0,39	6	8	1	16
Totales:				
	30	149	168	168

Referencia:

(-) Cantidades de piezas buenas con cada nivel de gris

(+) Cantidad de piezas con fallas o defectuosas



Resolución del problema

Tecnologías utilizadas

Para la solución del problema, se decidió utilizar la interfaz del navegador para representar el problema en una forma visual

Por lo tanto, se utilizó como lenguaje de programación Javascript y, además, para la maquetación de la página el lenguaje HTML y CSS para los estilos.

Métodos utilizados

Antes de empezar a analizar el problema por los distintos métodos utilizados, cabe aclarar que todos están enmarcados dentro del análisis ROC (Receiver Operating Characteristic).

Esta se trata de una técnica para evaluar y ajustar clasificadores dicotómicos.

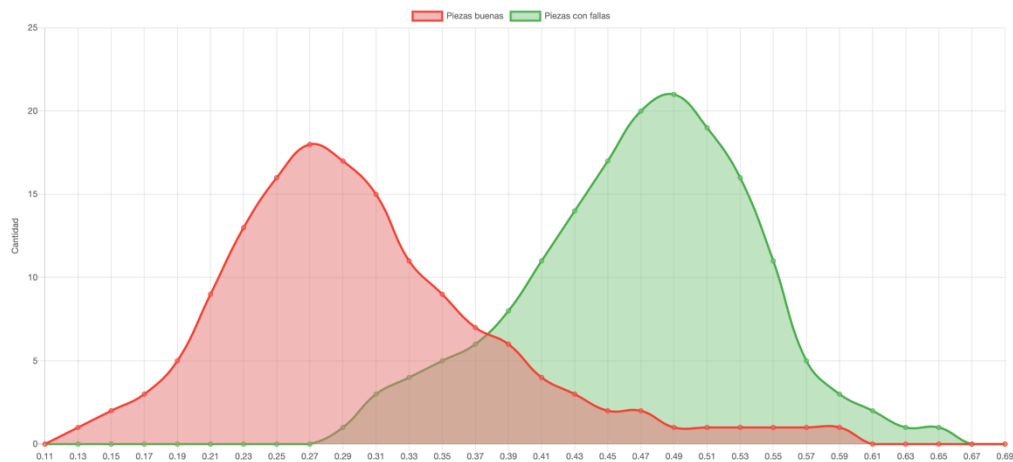
Curvas de distribución

Representa en un grafico las muestras positivas y negativas de un modelo para cada valor de corte.

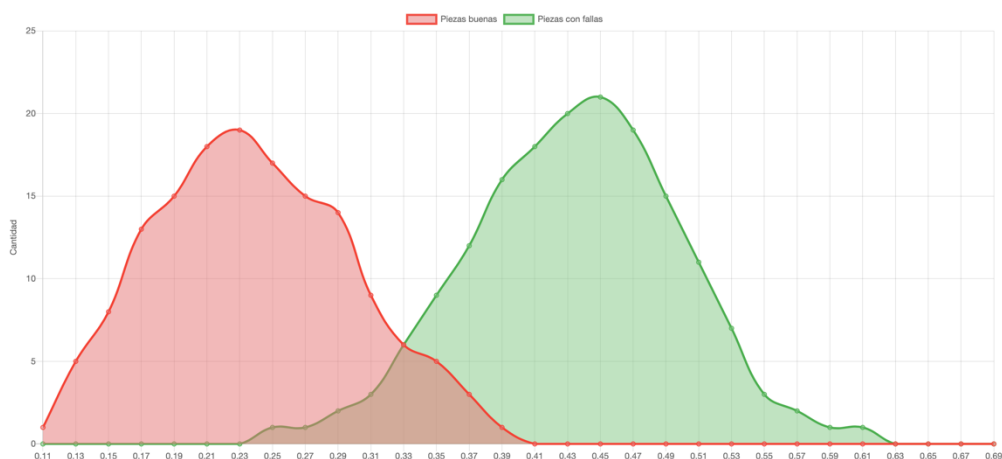
En el eje de las abscisas se colocan los diferentes valores de corte y en el eje de las ordenadas las cantidades de muestras.

En este caso las muestras positivas son las piezas con fallas y las muestras negativas son las piezas buenas.

Modelo "A"



Modelo "B"



Tablas de contingencia

Representa en una tabla los resultados de un modelo vs. diagnóstico verdadero

Tabla de contingencia		Diagnóstico verdadero	
		Positivo	Negativo
Resultado del modelo	Positivo	VP	FP
	Negativo	FN	VN
Muestras en las que coinciden los valores reales y las previsiones del modelo		Total muestras positivas	Total muestras negativas

VP: Verdaderos positivos
FN: Falsos negativos

FP: Falsos positivos
VN: Verdaderos negativos



De esta tabla, se pueden extraer la **sensibilidad** (Fracción de verdaderos positivos) y la **especificidad** (Fracción de verdaderos negativos)

$$\text{Sensibilidad} = \text{FVP} = \text{VP} / (\text{VP} + \text{FN})$$

$$\text{Especificidad} = \text{FVN} = \text{VN} / (\text{VN} + \text{FP}) = 1 - \text{FFP}$$

FVP: Fracción Verdaderos Positivos FVN: Fracción Verdaderos Negativos

Modelo "A"

Indic.	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47	0.49
VP	168	168	168	168	168	168	168	168	168	167	164	160	155	149	141	130	116	99	79	58
FN	0	0	0	0	0	0	0	0	0	1	4	8	13	19	27	38	52	69	89	110
FP	149	149	148	146	143	138	129	116	100	82	65	50	39	30	23	17	13	10	8	6
VN	0	0	1	3	6	11	20	33	49	67	84	99	110	119	126	132	136	139	141	143

Indic.	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
VP	39	23	12	7	4	2	1	0	0	0
FN	129	145	156	161	164	166	167	168	168	168
FP	5	4	3	2	1	0	0	0	0	0
VN	144	145	146	147	148	149	149	149	149	149

Modelo "B"

Indic.	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47	0.49
VP	168	168	168	168	168	168	168	167	166	164	161	155	146	134	118	100	80	59	40	25
FN	0	0	0	0	0	0	0	1	2	4	7	13	22	34	50	68	88	109	128	143
FP	149	148	143	135	122	107	89	70	53	38	24	15	9	4	1	0	0	0	0	0
VN	0	1	6	14	27	42	60	79	96	111	125	134	140	145	148	149	149	149	149	149

Indic.	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
VP	14	7	4	2	1	0	0	0	0	0
FN	154	161	164	166	167	168	168	168	168	168
FP	0	0	0	0	0	0	0	0	0	0
VN	149	149	149	149	149	149	149	149	149	149

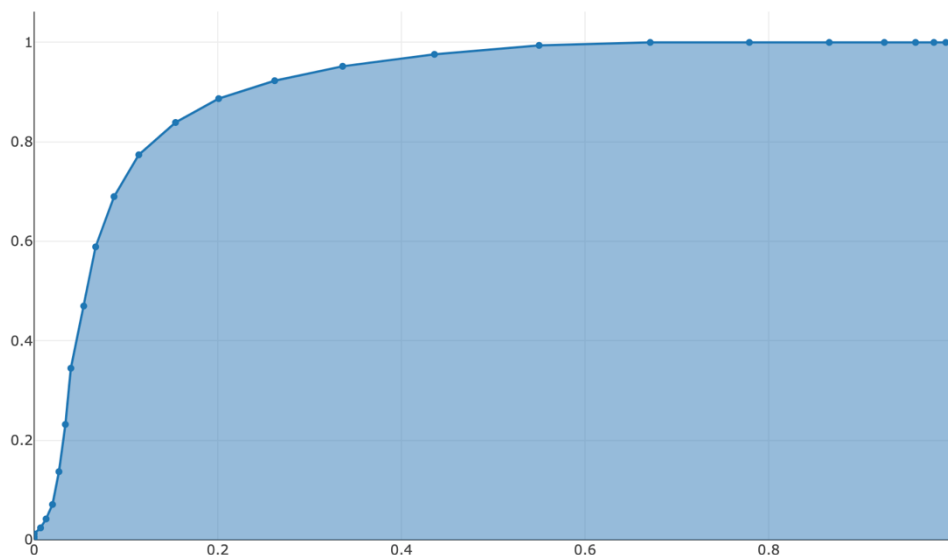


Curvas ROC

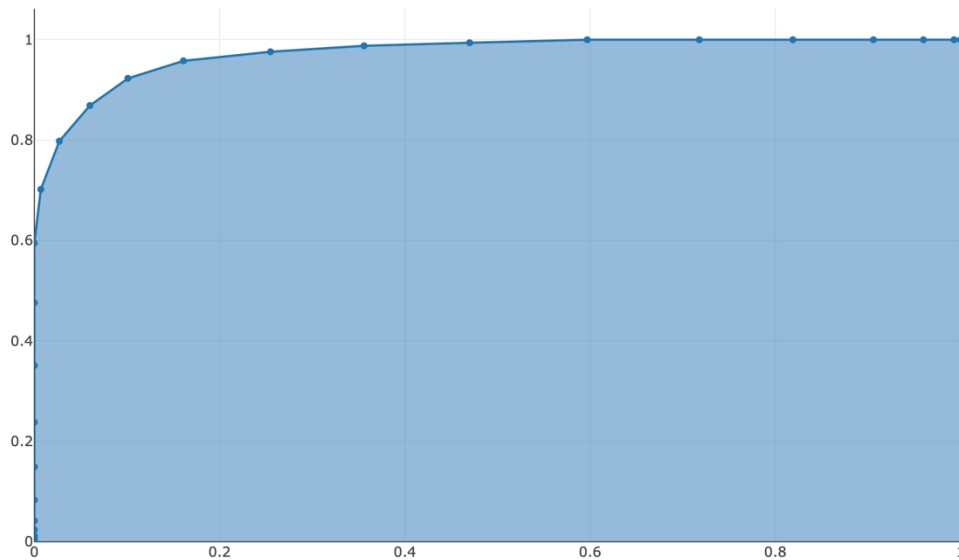
Este gráfico representa la sensibilidad, o fracción de verdaderos positivos (FVP), frente a la especificidad, o fracción de falsos positivos (FFP), para un sistema clasificador binario según se varía el umbral de discriminación (valor de corte).

- Se representa a partir de las tablas de contingencia del desempeño de un modelo.
- Proporciona una representación global de la exactitud de un diagnóstico.
- El área bajo la curva (AUC) sirve como indicador de la exactitud global del modelo.
- La exactitud máxima corresponde a un $AUC = 1$ y la mínima a un $AUC = 0,5$.
- Tiene el mejor desempeño cuando ambas poblaciones son similares en número.

Modelo "A"



Modelo "B"



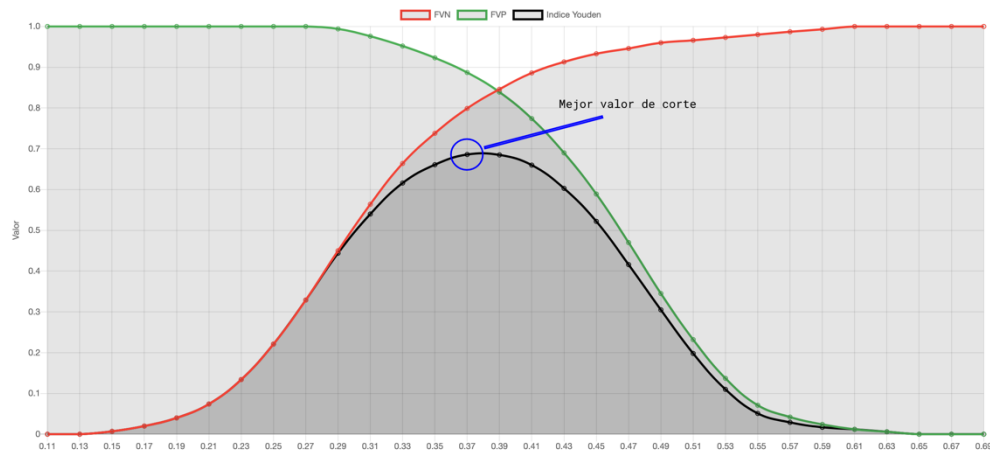
Índice de Youden

Este índice es tradicionalmente utilizado para comparar herramientas de diagnóstico

- Considera conjuntamente la sensibilidad y especificidad de un modelo.
- Establece el **mejor desempeño posible** de un clasificador y el **punto de corte** correspondiente.
- Mayores valores de γ están asociados a mejores desempeños de un modelo
- $\gamma = 1$ representa un clasificador perfecto ($0 \leq \gamma \leq 1$)
- Para calcular el índice de Youden se utiliza la siguiente formula:

$$\gamma = |\text{sensibilidad} + \text{especificidad} - 1|_{\max}$$

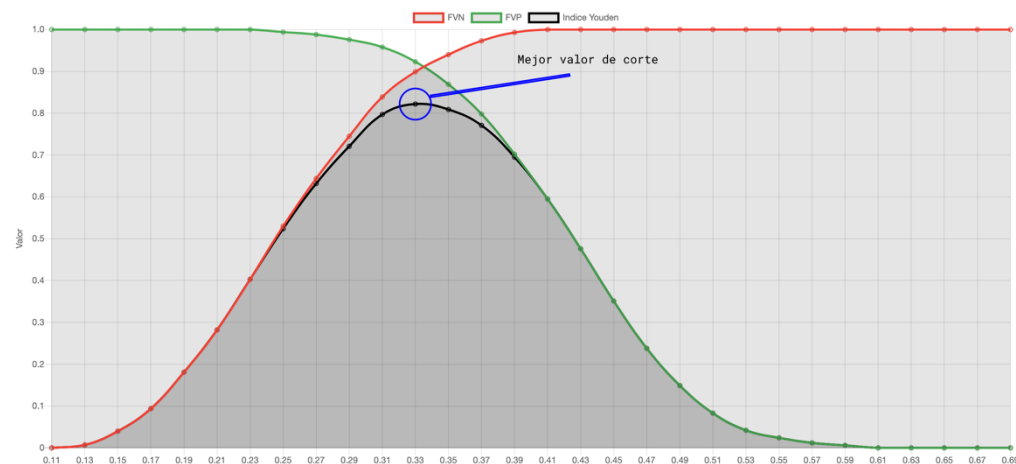
Modelo "A"



Indic.	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47	0.49
FVN	0	0	0.007	0.02	0.04	0.074	0.134	0.221	0.329	0.45	0.564	0.664	0.738	0.799	0.846	0.886	0.913	0.933	0.946	0.96
FVP	1	1	1	1	1	1	1	1	0.994	0.976	0.952	0.923	0.887	0.839	0.774	0.69	0.589	0.47	0.345	
Youden	0	0	0.007	0.02	0.04	0.074	0.134	0.221	0.329	0.444	0.54	0.616	0.661	0.686	0.685	0.66	0.603	0.522	0.416	0.305

Indic.	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
FFP	0.966	0.973	0.98	0.987	0.993	1	1	1	1	1
FVP	0.232	0.137	0.071	0.042	0.024	0.012	0.006	0	0	0
Youden	0.198	0.11	0.051	0.029	0.017	0.012	0.006	0	0	0

Modelo "B"



Indic.	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47	0.49
FVN	0	0.007	0.04	0.094	0.181	0.282	0.403	0.53	0.644	0.745	0.839	0.899	0.94	0.973	0.993	1	1	1	1	1
FVP	1	1	1	1	1	1	1	0.994	0.988	0.976	0.958	0.923	0.869	0.798	0.702	0.595	0.476	0.351	0.238	0.149
Youden	0	0.007	0.04	0.094	0.181	0.282	0.403	0.524	0.632	0.721	0.797	0.822	0.809	0.771	0.695	0.595	0.476	0.351	0.238	0.149

Indic.	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
FFP	1	1	1	1	1	1	1	1	1	1
FVP	0.083	0.042	0.024	0.012	0.006	0	0	0	0	0
Youden	0.083	0.042	0.024	0.012	0.006	0	0	0	0	0



Coeficiente de Gini

Es una medida de desigualdad en el modelo, ideada por el estadístico italiano Corrado Gini.

La formula de calculo es: $G = 2 \times AUC - 1$

Para un clasificador aleatorio se da: $AUC = 0,5$ y un $G = 0$

Para un clasificador perfecto se da: $AUC = 1$ y $G = 1$

Los valores posibles de G son: $0 \leq G \leq 1$

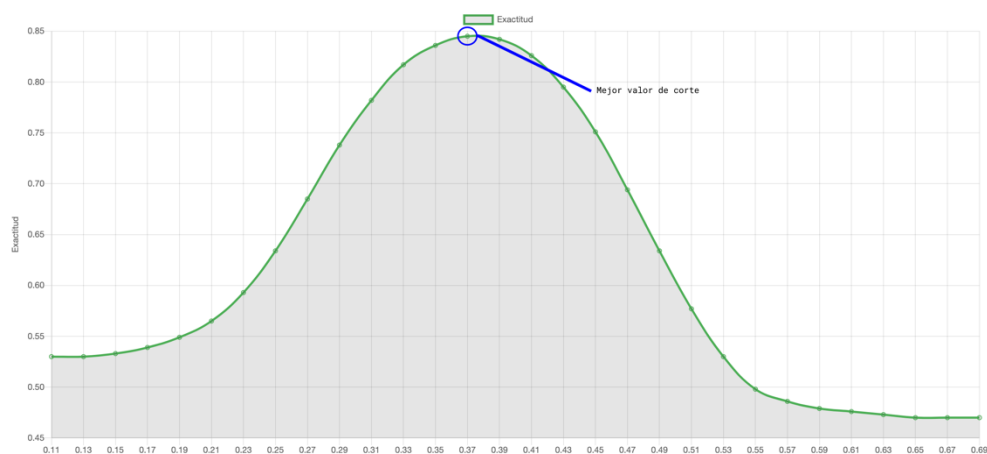
Índice de exactitud

Este índice da una relación de la cantidad de muestras clasificadas correctamente, ya sea en forma positiva o negativa, respecto al total de muestras tomadas en el modelo.

La formula de calculo de la exactitud es:

$$\text{exactitud} = (VP + VN) / (VP + FN + VN + FP)$$

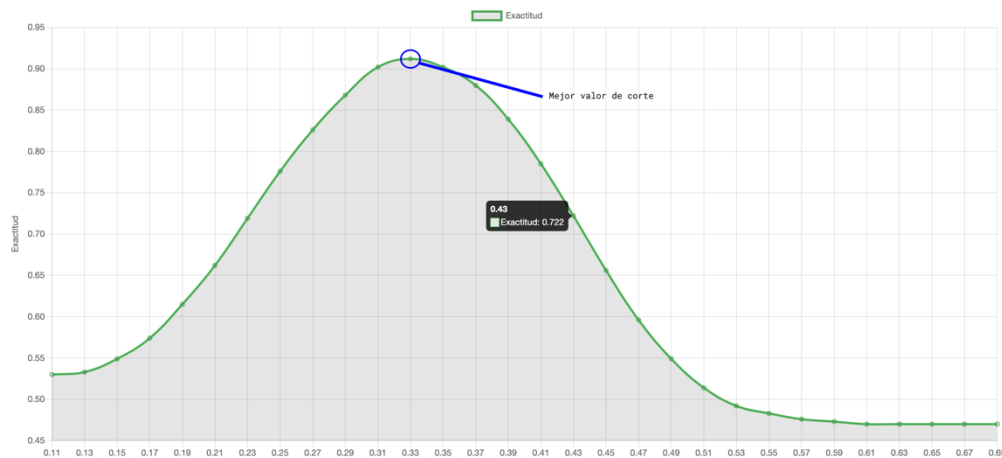
Modelo "A"



Indic.	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47	0.49
Exact.	0.53	0.53	0.533	0.539	0.549	0.565	0.593	0.634	0.685	0.738	0.782	0.817	0.836	0.845	0.842	0.826	0.795	0.751	0.694	0.634

Indic.	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
Exact.	0.577	0.53	0.498	0.486	0.479	0.476	0.473	0.47	0.47	0.47

Modelo "B"



Indic.	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47	0.49
Exact.	0.53	0.533	0.549	0.574	0.615	0.662	0.719	0.776	0.826	0.868	0.902	0.912	0.902	0.88	0.839	0.785	0.722	0.656	0.596	0.549

Indic.	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.65	0.67	0.69
Exact.	0.514	0.492	0.483	0.476	0.473	0.47	0.47	0.47	0.47	0.47

Mejor valor de corte

Curvas PR

Son útiles en los casos en que se dan fenómenos cuyas curvas de distribución se alejan considerablemente de la "normal", donde, el análisis ROC no es el más apropiado.

Las variables que se utilizan para esta curva son:

- Precisión (P):

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

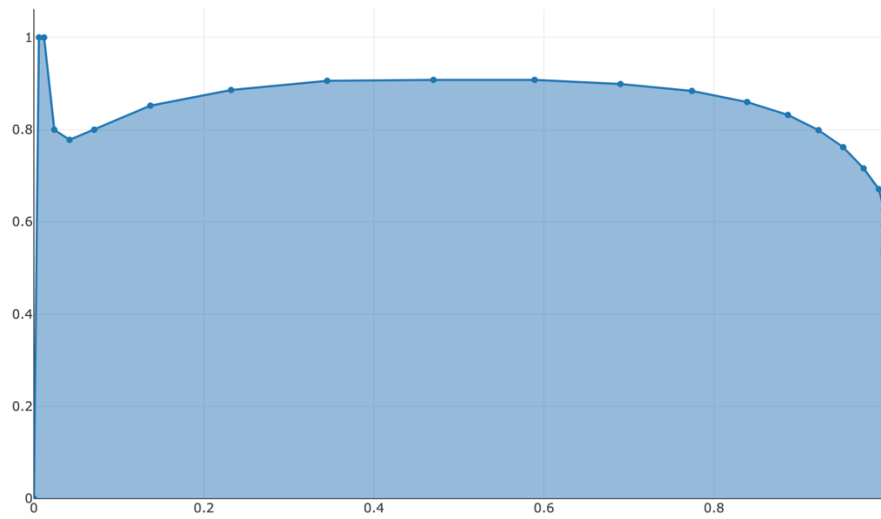
- Alcance o "recall" (R):

$$\text{Alcance} = \text{Sensibilidad} = \frac{VP}{VP + FN}$$

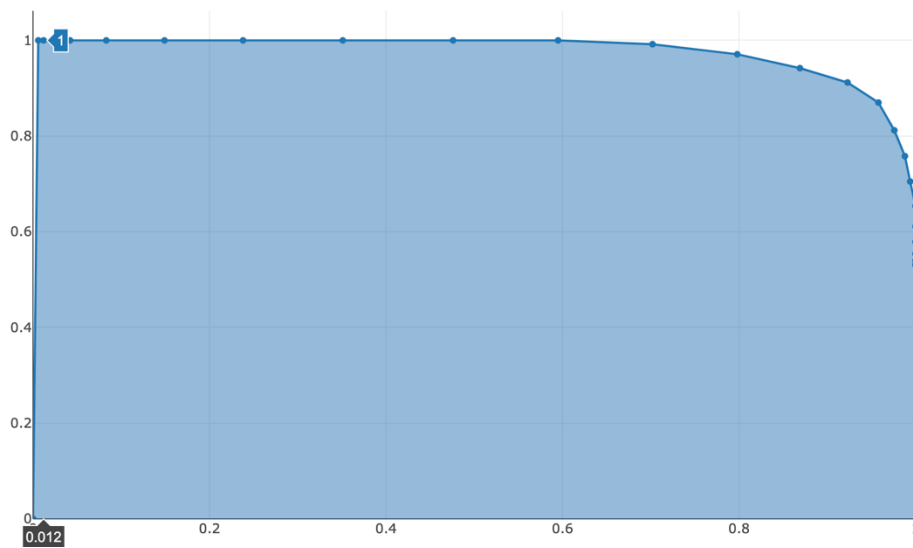
- Es preferible cuando el numero de muestras negativas (falsas positivas) es muy grande y las verdaderas positivas escasas.



Modelo "A"



Modelo "B"





Comparación de resultados

Indicadores evaluados	Modelos comparados	
	A	B
Máxima exactitud del modelo	0.845	0.912
Posición del valor de corte en que la máxima exactitud del modelo es alcanzada	0.37	0.33
Área bajo la curva ROC	0.905	0.973
Coefficiente de Gini	0.81	0.946
Mejor índice de Youden	0.686	0.822
Valor de corte del mejor índice de Youden	0.37	0.33

Como se puede ver:

- En la comparación de máximas exactitudes alcanzadas por los dos modelos, la mayor corresponde al modelo "B" (0.912).
- El índice de Youden permite llegar a las mismas conclusiones que con la máxima exactitud: el mejor clasificador es el modelo "B".
- Considerando el área bajo la curva ROC (AUC), el mejor modelo es el "B" también.

En conclusión, el **mejor modelo** considerado es el "B".

Ejecución del proyecto terminado

Para la ejecución del proyecto, dirigirse al repositorio del proyecto indicado al final de este informe y seguir las instrucciones detalladas en el archivo README.md.



Repositorio del proyecto terminado

- Enlace al repositorio del proyecto terminado:

<https://github.com/iguerra94/ModelValidation>

Conclusiones

Luego de realizar el análisis ROC como herramienta de comparación de los resultados obtenidos, se pudo concluir que el **modelo mas conveniente** para clasificar las muestras es el **modelo "B"**.