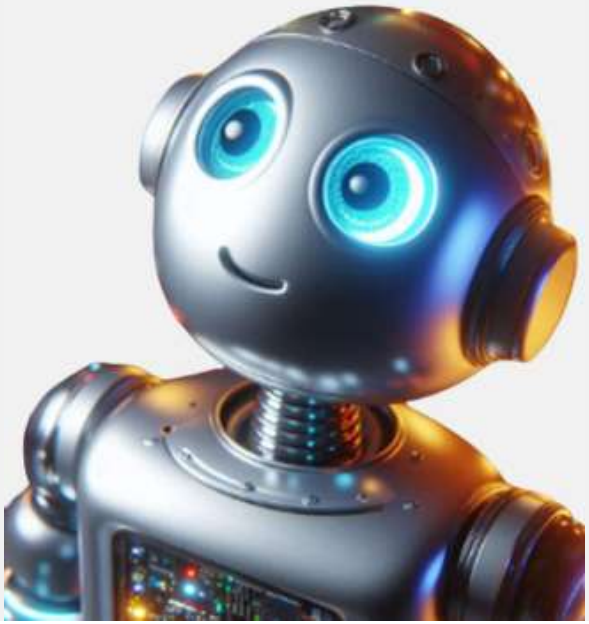
Regresión lineal y clasificación (Métricas de evaluación)

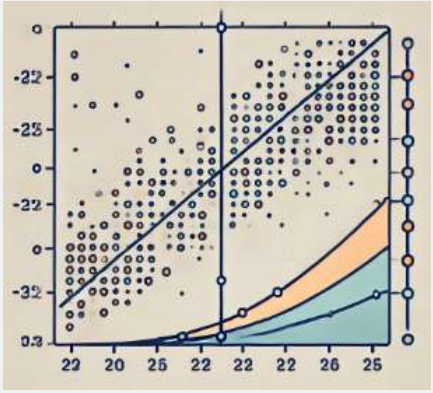
Prof. Luis Torrejón

LOGRO DE APRENDIZAJE

Conocer sobre la regresión lineal simple, múltiple y las métricas de evaluación de algoritmos de clasificación.



TEMARIO



TEMA 1

Evaluación de los algoritmos de regresión

TEMA 2

Clasificación y métricas de evaluación: Matriz de confusión, Curva ROC y AUC

Revisión de actividad grupal

REVISIÓN DE ACTIVIDAD GRUPAL:

- Instalar herramientas para realizar los desarrollos como Visual studio code, Anaconda, Miniconda, etc-

- Configurar los environment con Python y librerías base, desarrollando una aplicación que solicite una calificación (entre 0 y 100) y muestre si es suficiente para aprobar.

- Se expondrá en la siguiente clase.

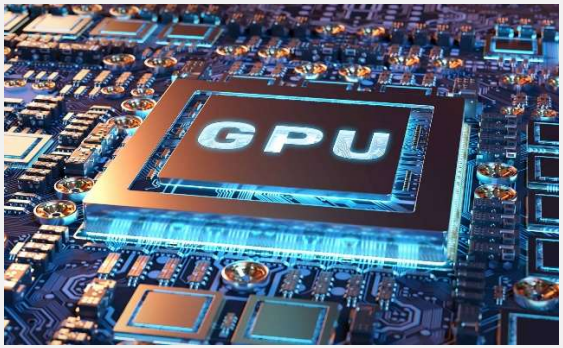
Herramientas IA

https://anaconda.org/

https://www.python.org/

https://code.visualstudio.com/

https://colab.research.google.com/

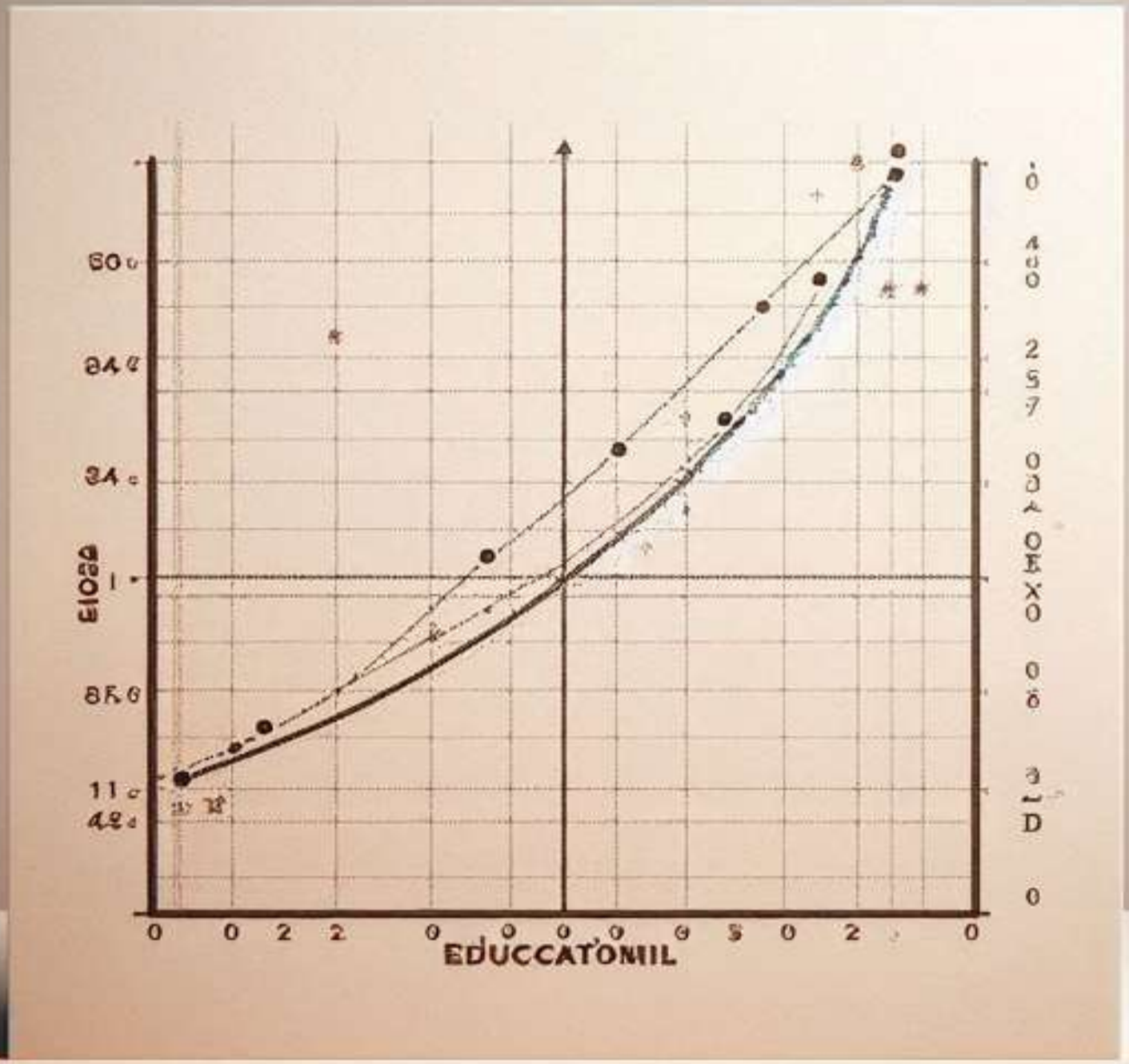


https://developer.nvidia.com/cuda-downloads

Herramientas IA

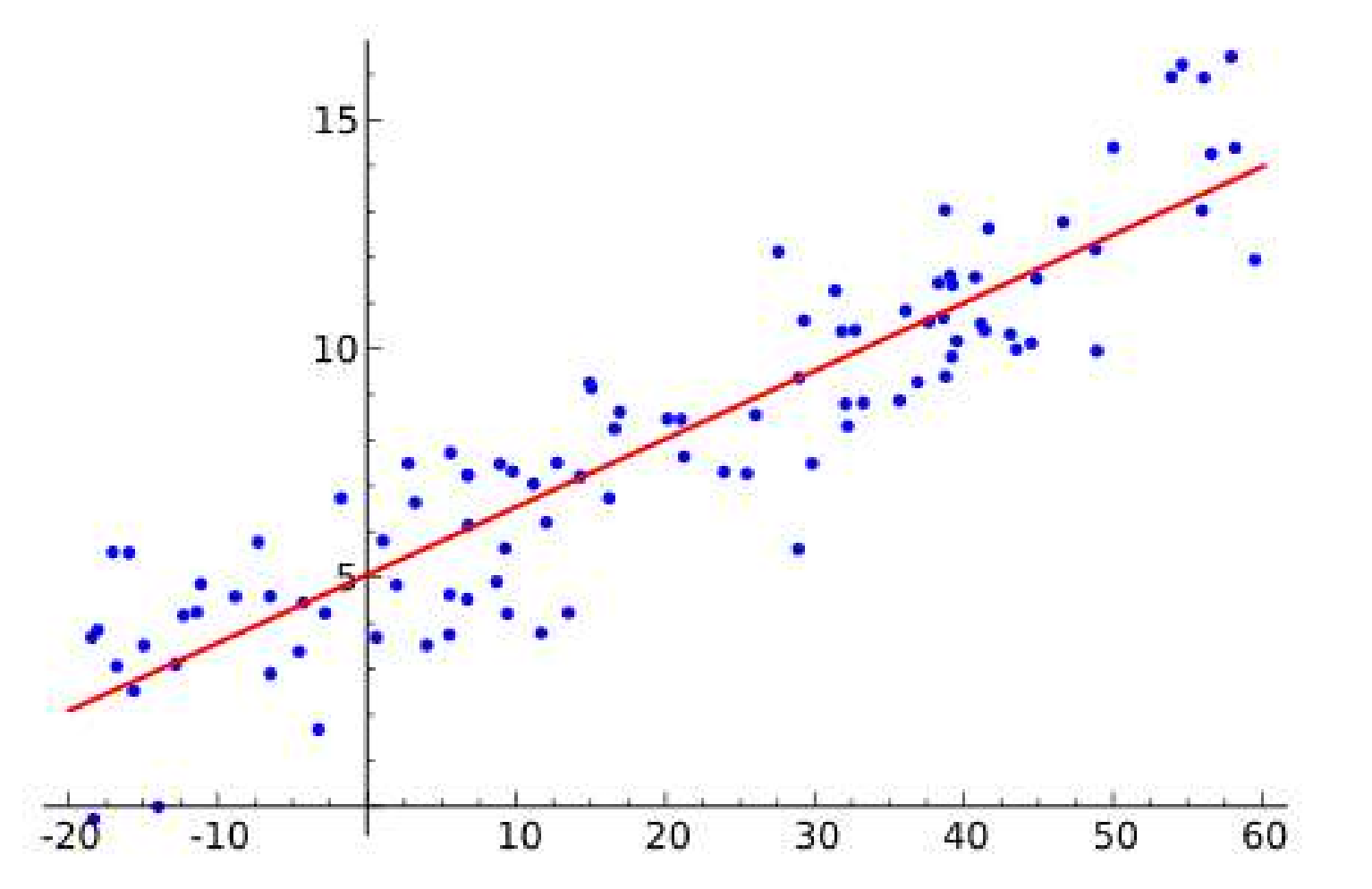
- ✓ Environment
- ✓ Framework
- ✓ Librerías

Regresión lineal simple



Regresión lineal simple

La regresión es una técnica estadística que se utiliza para analizar la relación entre dos o más variables. Permite predecir el valor de una variable dependiente en función de una o más variables independientes.



https://commons.wikimedia.org/wiki/File:Linear_regression.svg#/media/File:Linear_regression.svg

Permite responder:

¿Hay una conexión entre dos variables?
¿Qué tan intensa es esa conexión?
¿Cuál variable tiene mayor impacto?
¿Con qué grado de exactitud podemos medir el efecto de cada variable?
¿Con qué grado de exactitud podemos anticipar el resultado?
¿La relación es de tipo lineal?

Regresión lineal simple - ¿Qué puedo hacer?

- ✓ Predecir el rendimiento académico de estudiantes en función de variables como horas de estudio, asistencia a clase y calificaciones previas.
- ✓ Estimar la duración de un proyecto en función del presupuesto asignado, número de trabajadores y experiencia del equipo.
- ✓ Calcular el gasto en marketing de una empresa basado en su volumen de ventas, número de campañas publicitarias y presupuesto de marketing anterior.
- ✓ Predecir el consumo de energía de una vivienda en función de su tamaño, número de habitantes y tipo de calefacción.

Regresión lineal simple - Términos clave

- **Variable dependiente:** Es lo que queremos predecir (ej: precio de una casa).

- **Variables independientes:** Son los factores que creemos influyen en la variable dependiente (ej: tamaño de la casa, número de habitaciones).

- **Pendiente:** es un coeficiente que representa la tasa de cambio de la variable dependiente respecto a la variable independiente.

La regresión lineal simple, considera una única variable independiente que pretende explicar una variable dependiente o variable respuesta.

Regresión lineal simple - Fórmula

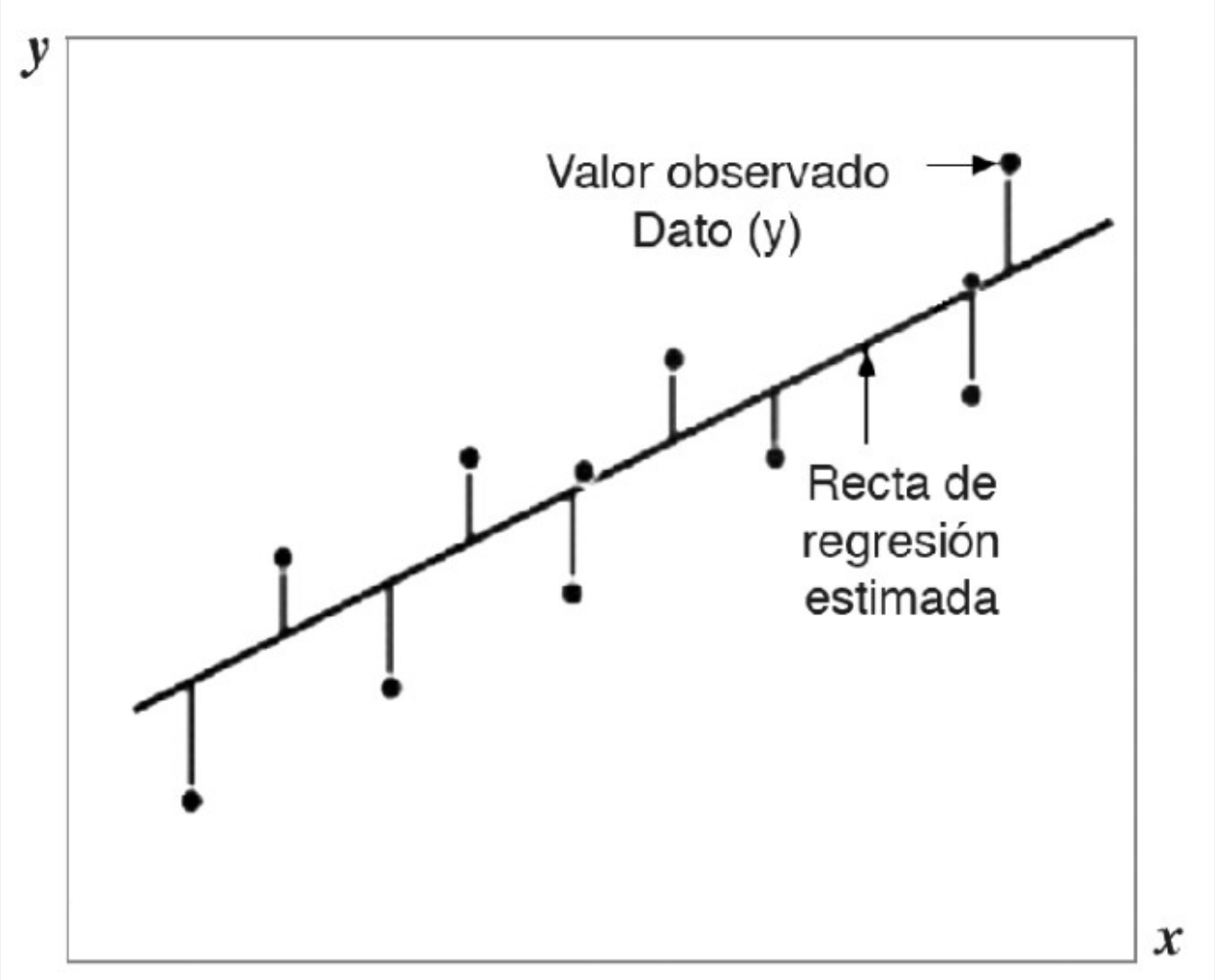
Es el caso más sencillo, en el que solo tenemos una variable independiente.

$$y = \beta_0 + \beta_1 x + \epsilon$$

- y es la variable dependiente (lo que queremos predecir).
- x es la variable independiente.
- β_0 es el **intercepto o intersección** (donde la línea cruza el eje y).
- β_1 es la **pendiente** (la tasa de cambio de y respecto a x).
- ϵ es el **error** (la diferencia entre el valor real y el valor predicho).

Regresión lineal simple - Método de Mínimos Cuadrados

Método de estimación de parámetros **Mínimos cuadrados** busca minimizar la distancia de los puntos a la recta *(cuadrado = elimina los signos)*



https://www.researchgate.net/publication/311548169_Regresion_lineal_simple_y_multiple_aplicacion_en_la_prediccion_de_variables_naturales_relacionadas_con_el_crecimiento_microalgal

Regresión lineal simple - Método de Mínimos Cuadrados

$$\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - (\beta_0 + \beta_1 X_i))^2$$

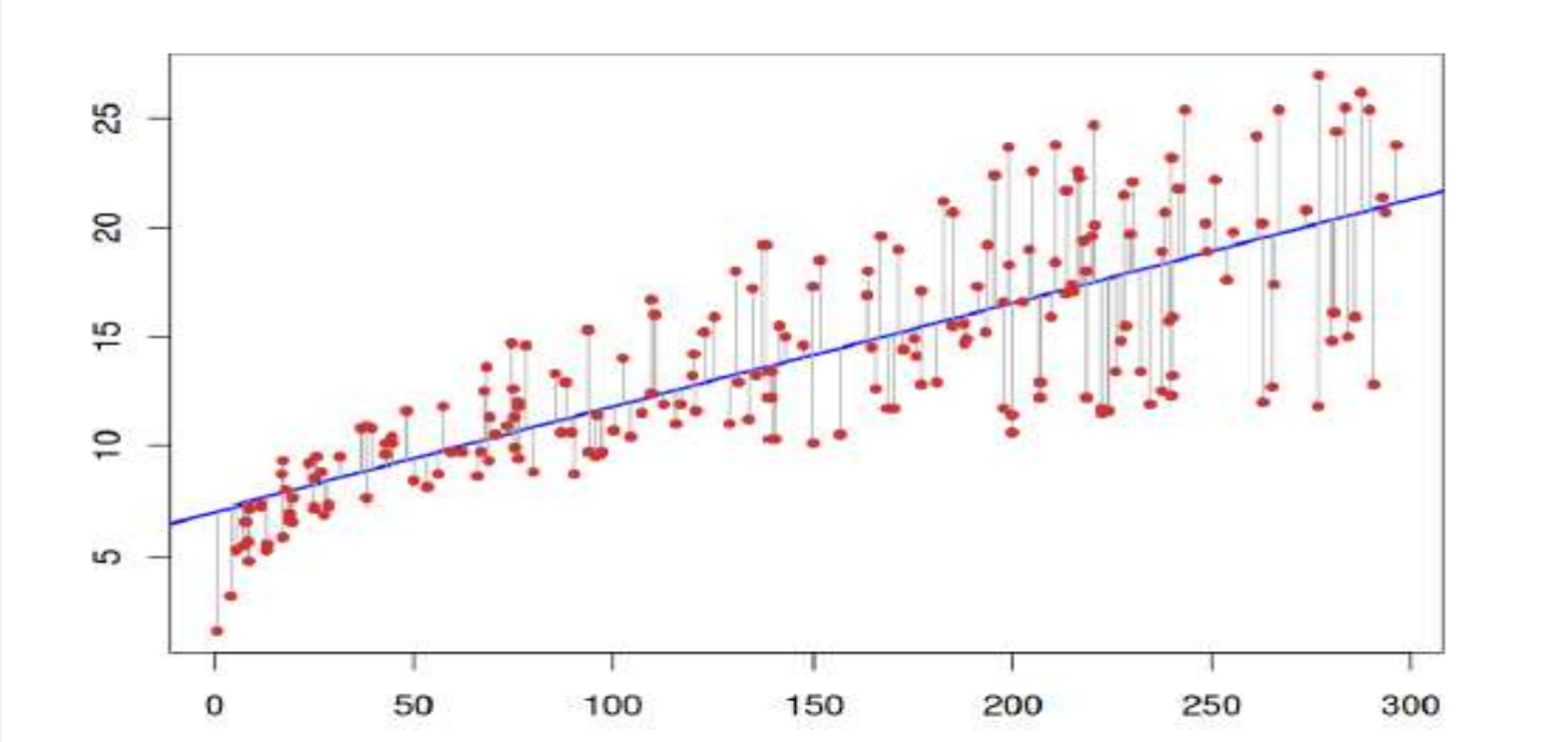
$$\beta_1 = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$\beta_0 = \frac{\sum Y - \beta_1(\sum X)}{n}$$

Regresión lineal simple – método de error

El método tendrá un error que podremos calcular restando la predicción al valor real:

$$e_i = y_i - \hat{y}_i$$



https://towardsdatascience.com/

Error reducible: es el tipo de error que se produce debido a la incapacidad del modelo para ajustarse perfectamente a los datos.

Es la diferencia entre la verdadera relación que existe entre las variables de entrada (X) y de salida (Y) y la función que nuestro modelo ha estimado.

Error irreducible: surge debido a factores fuera del control del modelo. Este tipo de error está relacionado con la variabilidad en Y que no puede explicarse por X.

Aunque tuviéramos un modelo perfecto (es decir, si f^ fuera exactamente igual a f), aún habría un error presente en nuestras predicciones porque **existen otras variables** que afectan a Y pero que no están incluidas en X.

Errores - ¿Nuestros coeficientes calculados son apropiados?

p-valor (para cada coeficiente), es una medida estadística que se utiliza para determinar la significancia. Nos permite saber si tenemos que aceptar o rechazar la hipótesis nula (afirmación que establece que no hay efecto, relación o diferencia en los datos).

- **p-valor < 0.05:** Evidencia suficiente para rechazar la hipótesis nula, sugiriendo que existe una relación significativa entre las variables.
- **p-valor ≥ 0.05:** No hay evidencia suficiente para rechazar la hipótesis nula, lo que implica que no se puede afirmar que existe una relación significativa.
- La **hipótesis nula** es lo contrario: no hay correlación entre las características y el objetivo.

Errores - ¿Nuestros coeficientes calculados son apropiados?

OLS Regression Results

Dep. Variable:	y	R-squared:	0.852
Model:	OLS	Adj. R-squared:	0.848
Method:	Least Squares	F-statistic:	224.3
Date:	Fri, 24 Oct 2024	Prob (F-statistic):	2.89e-23
Time:	12:00:00	Log-Likelihood:	-137.45
No. Observations:	80	AIC:	280.9
Df Residuals:	78	BIC:	285.4
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	4.0301	0.099	40.663	0.000	3.834	4.226
X	2.9174	0.195	14.970	0.000	2.530	3.305

Omnibus:	1.042	Durbin-Watson:	2.118
Prob(Omnibus):	0.595	Jarque-Bera (JB):	0.823
Skew:	0.269	Prob(JB):	0.364
Kurtosis:	2.233	Cond. No.	4.56

Errores - ¿Cómo podemos saber que nuestro modelo es bueno?

- **RSE (Residual Standard Error):** Mide la precisión de las predicciones del modelo; un RSE más bajo indica un mejor ajuste. El **RSE** te indica qué tan lejos, en promedio, están las predicciones del modelo de los valores reales.

- **R^2 (Coeficiente de Determinación):** Mide la proporción de la variabilidad en la variable dependiente que se explica por las variables independientes. Un R^2 alto sugiere un buen ajuste del modelo.

  - R^2 = 1: Esto significa que el modelo explica el 100% de la variabilidad de la variable dependiente. Todas las predicciones son perfectas.

  - R^2 = 0: Esto significa que el modelo no explica nada de la variabilidad de la variable dependiente. Las predicciones son tan buenas como simplemente usar la media de la variable dependiente.

  - R^2 negativo: Esto ocurre cuando el modelo se ajusta peor que la media. Esto es raro, pero puede suceder si el modelo está mal especificado.

Errores - ¿Cómo podemos saber que nuestro modelo es bueno?

El coeficiente de determinación, R2R, es un indicador estadístico que muestra qué tan bien las predicciones del modelo de regresión se alinean con los datos observados.

Errores: P-valor, RSE, R^2 - Ejemplo

Se construye un modelo de regresión lineal con las variables **habitaciones**, **metros cuadrados** y **ubicación** para predecir el precio de una casa.

1. **p-valor:**
   - **Habitaciones** es 0.02
   - **Metros cuadrados** es 0.01
   - **Ubicación** es 0.001.

2. El **RSE** es $30,000.

3. El **R^2** es 0.90.

Las variables independientes deben tener una **distribución normal** para que este tipo de algoritmos funcionen correctamente.

- Si no cumplen con esta condición, se pueden aplicar transformaciones logarítmicas (como tomar el logaritmo natural de los datos) para que se acerquen más a la distribución deseada.

- Si se utiliza esta transformación, es importante que al hacer las predicciones o estimaciones se aplique la función exponencial para volver a los valores originales.

- ✓ **Relación lineal** entre las variables (características y objetivo).
- ✓ **Baja o nula multicolinealidad** entre las variables predictoras.
- ✓ **Distribución normal** de los errores (residuos).
- ✓ **Homocedasticidad** (los errores tienen varianza constante).
- ✓ Ausencia de **autocorrelación** en los errores

Métricas de error

Métricas de error

Para evaluar la precisión y efectividad de los modelos predictivos, se utilizan diversas métricas de error que cuantifican la diferencia entre los valores reales y las predicciones generadas por el modelo.

Ayudan a **evaluar la calidad de las predicciones** de tu modelo.

- MSE (Mean Squared Error)
- MAE (Mean Absolute Error)
- RMSE (Root Mean Squared Error)
- RMSLE (Root Mean Squared Logarithmic Error**)**

Métricas de error

- **MSE (Mean Squared Error) - Error Cuadrático Medio**: mide la media de los cuadrados de las diferencias entre los valores predichos y los valores reales. Penaliza los errores grandes más que los pequeños, ya que se elevan al cuadrado.
  Un MSE bajo indica que el modelo tiene buenas predicciones en comparación con los valores reales.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- **MAE (Mean Absolute Error) - Error Absoluto Medio:** es el promedio de las diferencias absolutas entre los valores reales y predichos. A diferencia del MSE, no penaliza más los errores grandes.
  Es útil para tener una idea clara del error promedio sin dar más peso a errores grande

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Métricas de error

- **RMSE (Root Mean Squared Error) - Raíz del Error Cuadrático Medio:** es la raíz cuadrada del MSE. Tiene las mismas unidades que la variable que estamos prediciendo, lo que facilita su interpretación. Al igual que el MSE, penaliza más los errores grandes.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

- **RMSLE (Root Mean Squared Logarithmic Error) - Raíz del Error Logarítmico Medio Cuadrado:** se utiliza principalmente para evaluar el rendimiento de modelos de regresión, especialmente cuando se trabaja con datos que pueden tener grandes diferencias en magnitud.

$$\mathrm{RMSLE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log(1 + y_i) - \log(1 + \hat{y}_i))^2}$$

Métricas de error – ejemplo – cumple con expectativas

- **Rango de precios de casas**: $200,000 - $500,000
- **Expectativa**: Se espera que el modelo tenga un error bajo, dado que se cuenta con buenas características y un conjunto de datos robusto.

Resultados del Modelo:

- **MSE** = 50
- **MAE** = 5,000
- **RMSE** = 7,071
- **RMSLE** = 0.02

Interpretación:

- ✓ **MSE = 50**: Indica un error muy bajo en la predicción, ya que el modelo se está ajustando bien a los datos.
- ✓ **MAE = 5,000**: Esto sugiere que, en promedio, las predicciones están a solo $5,000 de los precios reales, lo cual es aceptable.
- ✓ **RMSE = 7,071**: Mide el error en unidades de precio y, dado el rango de precios, sugiere que el modelo es confiable.
- ✓ **RMSLE = 0.02**: Un valor bajo indica que el modelo está capturando bien las proporciones de los precios.

Métricas de error – ejemplo – no cumple con expectativas

- **Rango de precios de casas**: $200,000 - $500,000
- **Expectativa**: Se espera que el modelo se ajuste bien debido a la calidad de los datos.

Resultados del Modelo:

- **MSE = 50,000**
- **MAE = 10,000**
- **RMSE = 224.00**
- **RMSLE = 0.15**

Interpretación:

- ✓ **MSE = 50,000**: Indica un error alto, lo que sugiere que el modelo no se está ajustando bien a los datos.
- ✓ **MAE = 10,000**: Esto significa que las predicciones están, en promedio, a $10,000 de distancia de los valores reales, lo que es considerablemente alto.
- ✓ **RMSE = 224.00**: Un valor alto también indica que las predicciones son poco confiables.
- ✓ **RMSLE = 0.15**: Este valor también es relativamente alto, sugiriendo que el modelo no captura bien las proporciones.

Regresión lineal múltiple

Regresión lineal múltiple

Es una extensión de la regresión lineal simple que permite modelar la relación entre una variable dependiente y múltiples variables independientes.

El objetivo es encontrar una línea (o un hiperplano en dimensiones superiores) que mejor se ajuste a los datos, minimizando la suma de los errores cuadrados entre las predicciones del modelo y los valores reales.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon$$

Regresión lineal múltiple - Ejemplo de precio de una casa

1. **Variables dependientes**: Precio de la casa.

2. **Variables independientes**:
   - Número de habitaciones
   - Tamaño en m2
   - Ubicación
   - Antigüedad
   - Número de baños

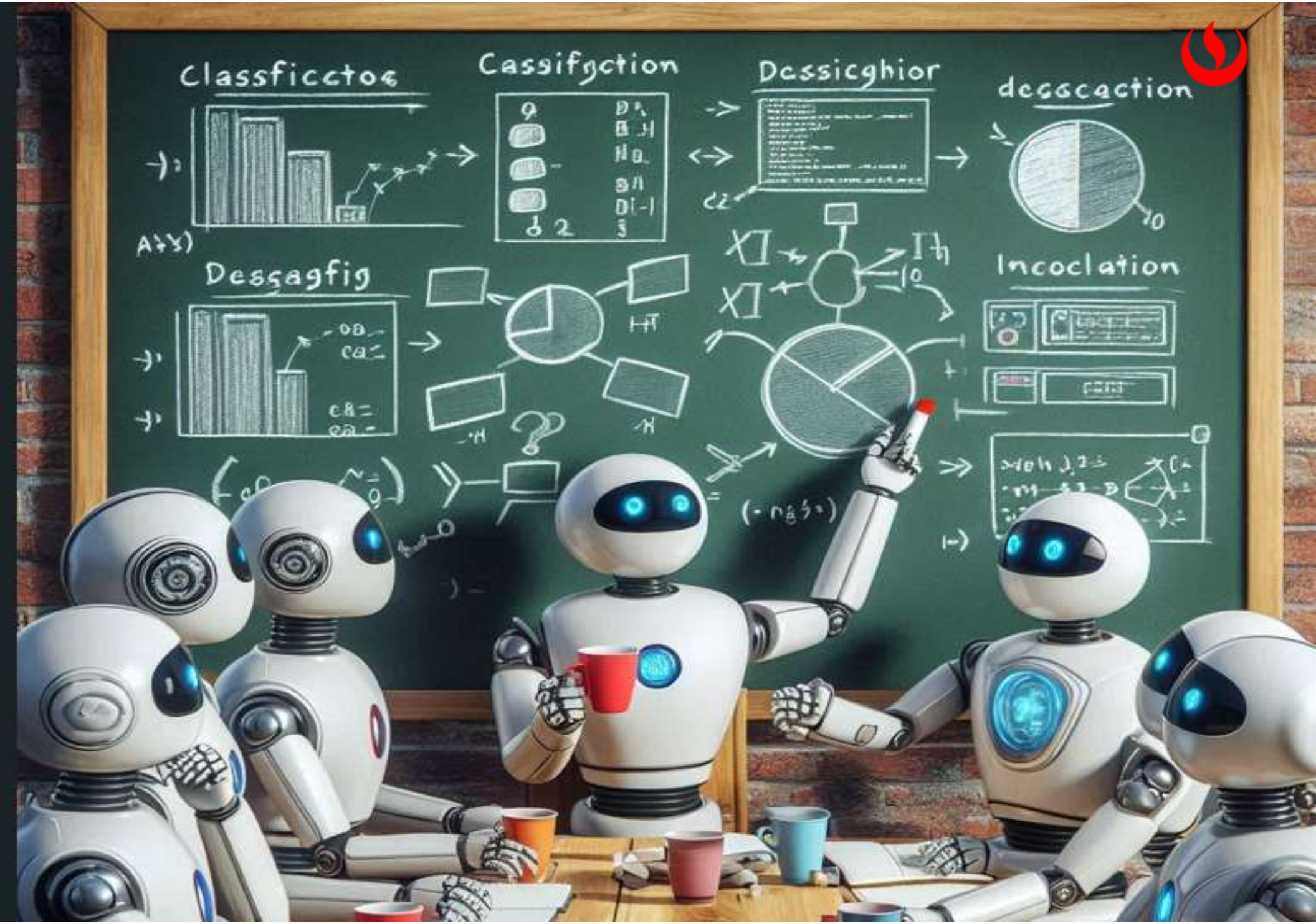
3. **Aplicación**: Un agente inmobiliario podría utilizar un modelo de regresión lineal múltiple para predecir el precio de una casa en función de estas características.

$$Precio = 5000 - 100 \times (Año) - 0.5 \times (Kilometraje)$$

Regresión lineal múltiple - Python

- Pandas - https://pandas.pydata.org/
- Matplotlib - https://matplotlib.org/
- Seaborn - https://seaborn.pydata.org/
- Sklearn - https://scikit-learn.org/stable/
- Statsmodels - https://www.statsmodels.org/stable/index.html
- Numpy - https://numpy.org/

Algoritmos de clasificación



Algoritmos de clasificación

Un algoritmo de clasificación intenta determinar a qué **categoría** o **grupo** pertenece una nueva entrada, en función de los datos de entrenamiento proporcionados.

https://www.freepik.es/vector-gratis/juego-clasificacion-basura_13146308.htm

Algoritmos de clasificación -Tipos de Clasificación

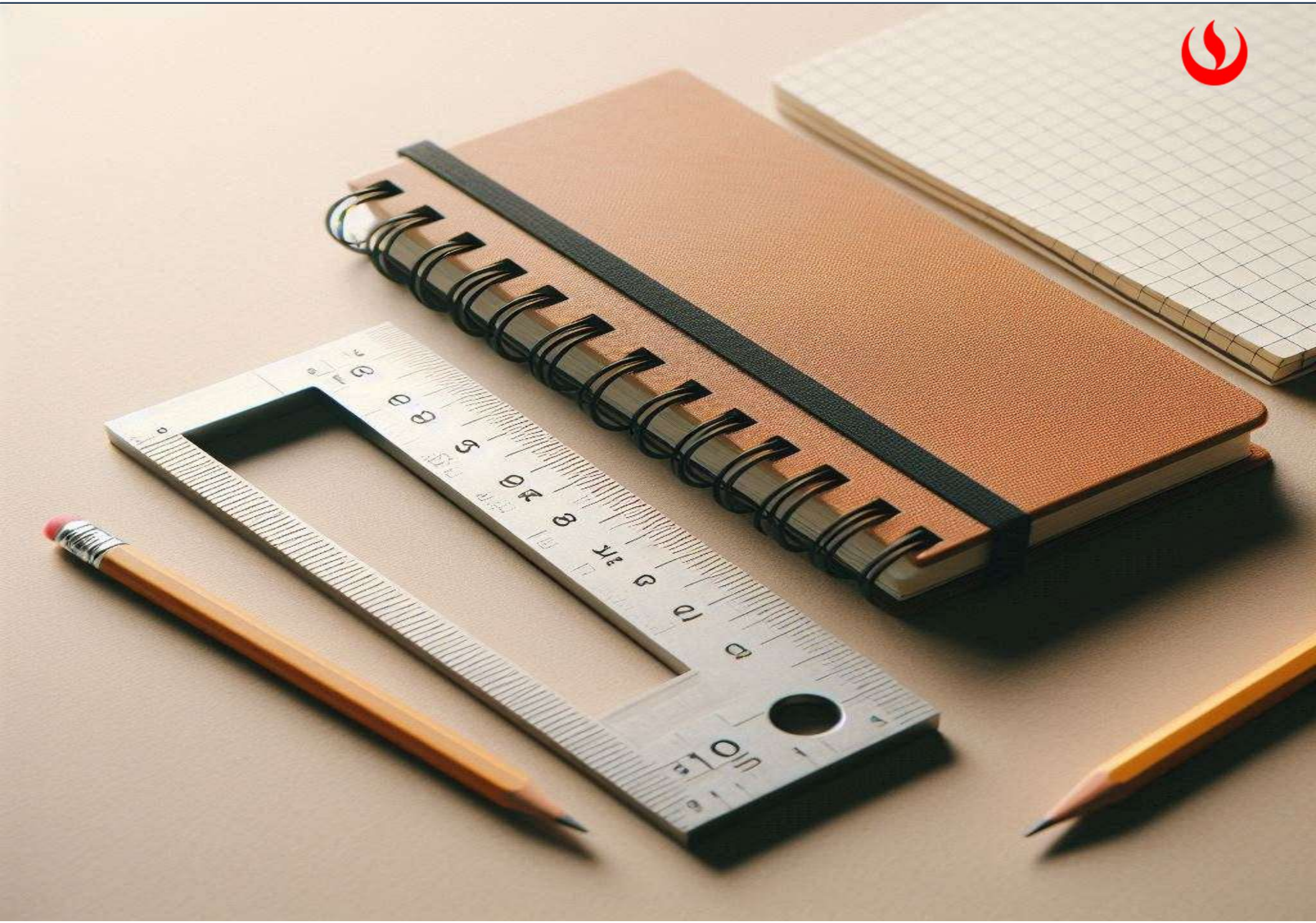
1.Clasificación Binaria: Cuando hay solo dos clases o categorías posibles.
Ejemplo: Predecir si un correo electrónico es **spam** o **no spam**.

2.Clasificación Multiclase: Cuando hay más de dos categorías posibles.
Ejemplo: Predecir el tipo de flor (iris setosa, iris virginica, iris versicolor) basado en características como el tamaño de los pétalos.

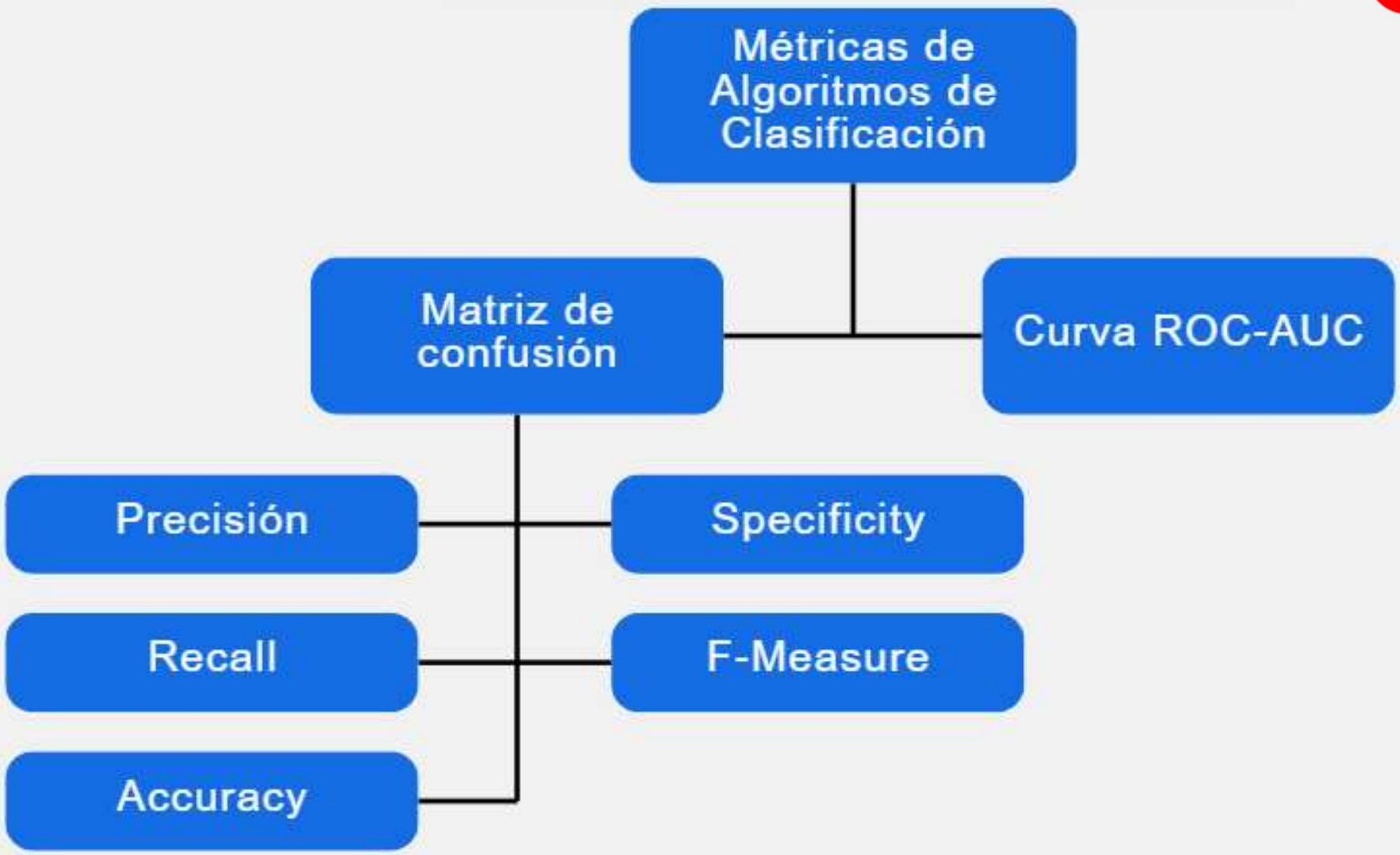
Algoritmos de clasificación - Ejemplos

https://flic.kr/p/2eGtv1C

Algoritmos de clasificación - Métricas



Algoritmos de clasificación - Métricas



Algoritmos de clasificación – Métricas – Matriz de confusión

Es una herramienta que muestra el número de predicciones para clasificadores binarios, por ejemplo, correctas e incorrectas.

Se clasifican en una tabla con las categorías "Predicción Positiva" y "Predicción Negativa"

	Predicción positiva	Predicción negativa
Real: Positivo	TP	FN
Real: Negativo	FP	TN

Algoritmos de clasificación – Métricas – Matriz de confusión - Ejemplo

Supongamos que tienes un modelo que clasifica correos electrónicos como "Spam" o "No Spam". Luego de evaluar el modelo con 100 correos.

	Predicción positiva	Predicción negativa
Real: Positivo	50 (TP)	10 (FN)
Real: Negativo	5 (FP)	35 (TN)

- ✓ True Positives (TP): 50 correos correctamente identificados como spam.
- ✓ False Negatives (FN): 10 correos que eran spam, pero el modelo los clasificó como no spam.
- ✓ False Positives (FP): 5 correos que no eran spam, pero el modelo los clasificó incorrectamente como spam.
- ✓ True Negatives (TN): 35 correos correctamente identificados como no spam.

Algoritmos de clasificación – Métricas – Matriz de confusión – Accuracy (Exactitud)

Es la proporción de predicciones correctas (tanto verdaderos positivos como verdaderos negativos) entre todas las predicciones realizadas. Es una métrica común que mide qué tan a menudo el clasificador predice correctamente.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}}$$

$$\text{Exactitud} = \frac{50 + 35}{50 + 35 + 5 + 10} = 0.85\,(85\%)$$

Algoritmos de clasificación – Métricas – Matriz de confusión - Precisión

Es la proporción de verdaderos positivos entre todas las instancias que fueron predichas como positivas. Indica qué tan precisa es una clase específica al evitar falsos positivos.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Precisión} = \frac{TP}{TP + FP} = \frac{50}{50 + 5} = 0.91\,(91\%)$$

Algoritmos de clasificación – Métricas – Matriz de confusión – Recall (Sensibilidad)

Es la proporción de verdaderos positivos entre todas las instancias que realmente son positivas. Mide la capacidad del modelo para detectar correctamente los casos positivos, evitando falsos negativos

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{50}{50 + 10} = 0.83\,(83\%)$$

Algoritmos de clasificación – Métricas – Matriz de confusión - Specificity

También llamada tasa de **verdaderos negativos,** es una métrica que se define como la proporción de verdaderos negativos que se identifican correctamente entre todos los casos negativos.

$$\text{Specificity} = \frac{\text{Verdaderos Negativos (TN)}}{\text{Verdaderos Negativos (TN)} + \text{Falsos Positivos (FP)}}$$

$$\text{Specificity} = \frac{35}{35+5} = \frac{35}{40} = 0.875$$

Algoritmos de clasificación – Métricas – Matriz de confusión – F1-Score

Es la media armónica entre la precisión y el recall. Es útil cuando se busca un balance entre ambas métricas, ya que un modelo con alta precisión, pero bajo recall o viceversa puede no ser deseable.

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

$$\text{F1-Score} = 2 \times \frac{0.91 \times 0.83}{0.91 + 0.83} = 0.87$$

Algoritmos de clasificación – Métricas – resumen de significado

- **Accuracy(Exactitud):** nos da una idea general de cuántas predicciones fueron correctas, pero no nos dice si el modelo se está desempeñando bien con respecto a las clases individuales.
- **Precision** es crítica cuando el costo de los falsos positivos es alto.
- **Recall(Sensibilidad)** es importante cuando queremos minimizar los falsos negativos.
- **Specifity(Especificidad)** es útil cuando debemos evitar falsos positivos.
- **F1-Score** equilibra la precisión y el recall, siendo útil cuando ambas métricas son igual de importantes.

Ejemplo:

- **Accuracy (Exactitud)**: 85%
- **Precision**: 88.8%
- **Recall**: 80%
- **Especificidad**: 90%
- **F1-Score**: 84.2%

Algoritmos de clasificación – Métricas – Ejemplo

Imagina que estamos intentando detectar un caso raro como fraude en transacciones bancarias. Digamos que en un conjunto de datos hay **990 transacciones normales** y **10 fraudes**.

	Predicción Positiva (Fraude)	Predicción Negativa (No Fraude)
Real Positivo (Fraude)	8	2
Real Negativo (No Fraude)	115	875

- **Accuracy (Exactitud): 88.3%**
- **Precision: 6.5%**
- **Recall: 80%**
- **Especificidad: 88.4%**
- **F1-Score: 12%**

Algoritmos de clasificación – Métricas – ¿Cuál utilizamos?

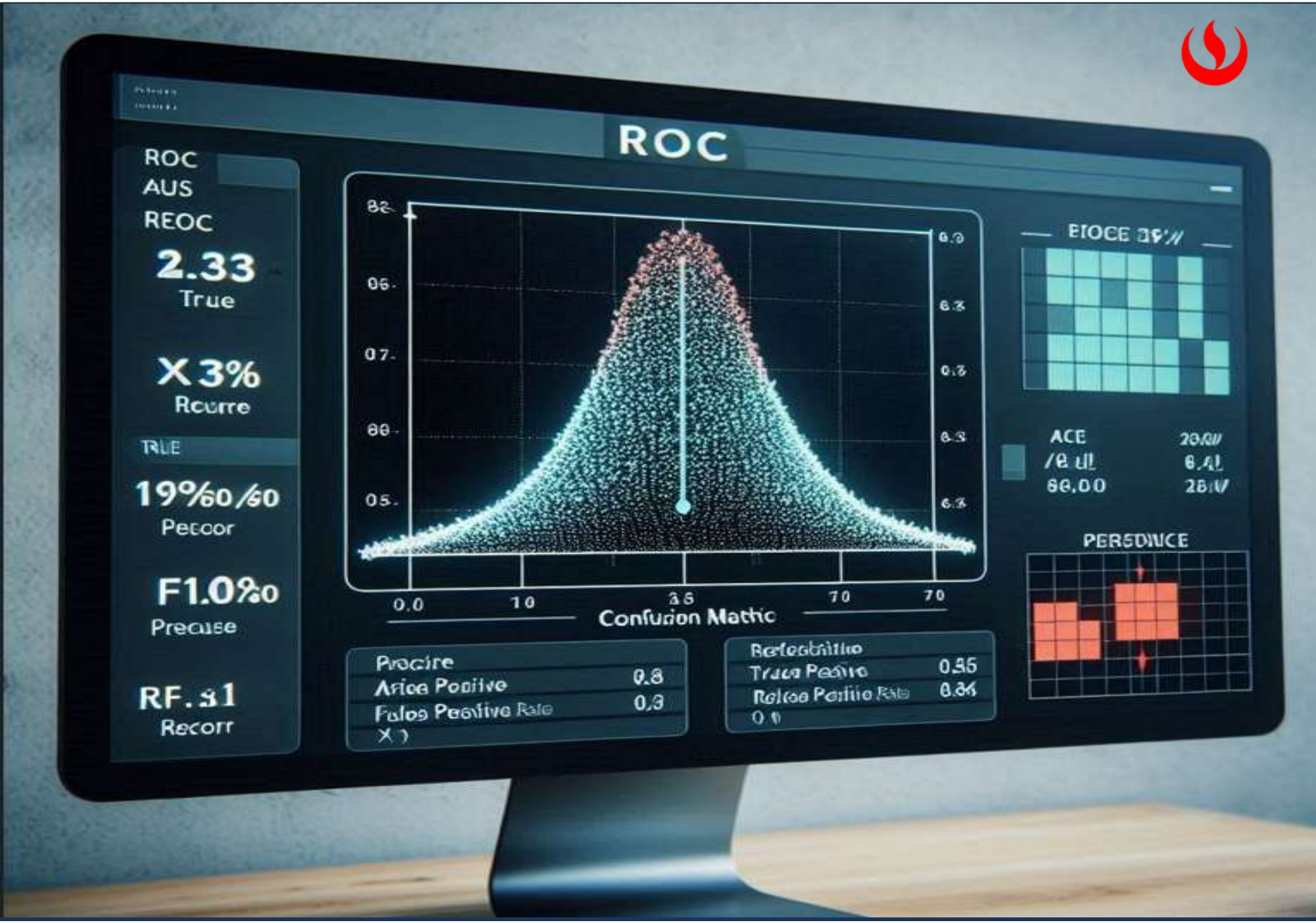
- Todas las métricas (precisión, recall, accuracy, especificidad, etc) nos brindan información valiosa sobre la calidad de nuestro modelo.

- Es fundamental analizar todas ellas. Por ejemplo, si ignoramos la especificidad, podríamos desarrollar un modelo que tenga alta precisión y recall, pero que simplemente clasifique todos los casos como verdaderos.

- El recall nos indica el rendimiento de un clasificador en relación con los falsos negativos (es decir, cuántos casos no se detectaron), mientras que la precisión nos ofrece información sobre su rendimiento en cuanto a los falsos positivos (es decir, cuántos fueron incorrectamente identificados como positivos).

- Además, no debemos pasar por alto la especificidad, especialmente en contextos médicos donde es crucial minimizar los falsos positivos.

Algoritmos de clasificación – Curva ROC - AUC



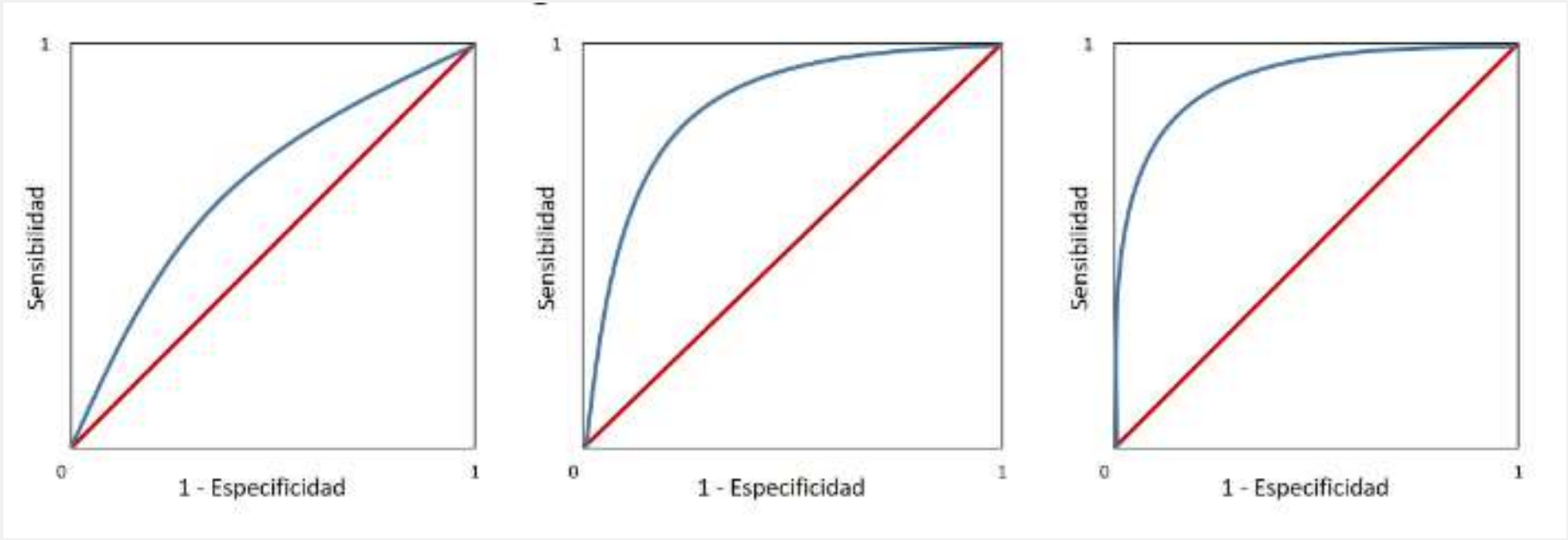
Algoritmos de clasificación – Curva ROC

- Un gráfico **ROC** (Receiver Operating Characteristic o Curva de Característica Operativa del Receptor) es una visualización de la **sensibilidad** frente a la **especificidad** en un sistema de clasificación binaria.

- Otra forma de entender este gráfico es como la representación de la **tasa de verdaderos positivos** frente a la **tasa de falsos positivos**, ajustando el **umbral de decisión** (el valor a partir del cual clasificamos un caso como positivo).

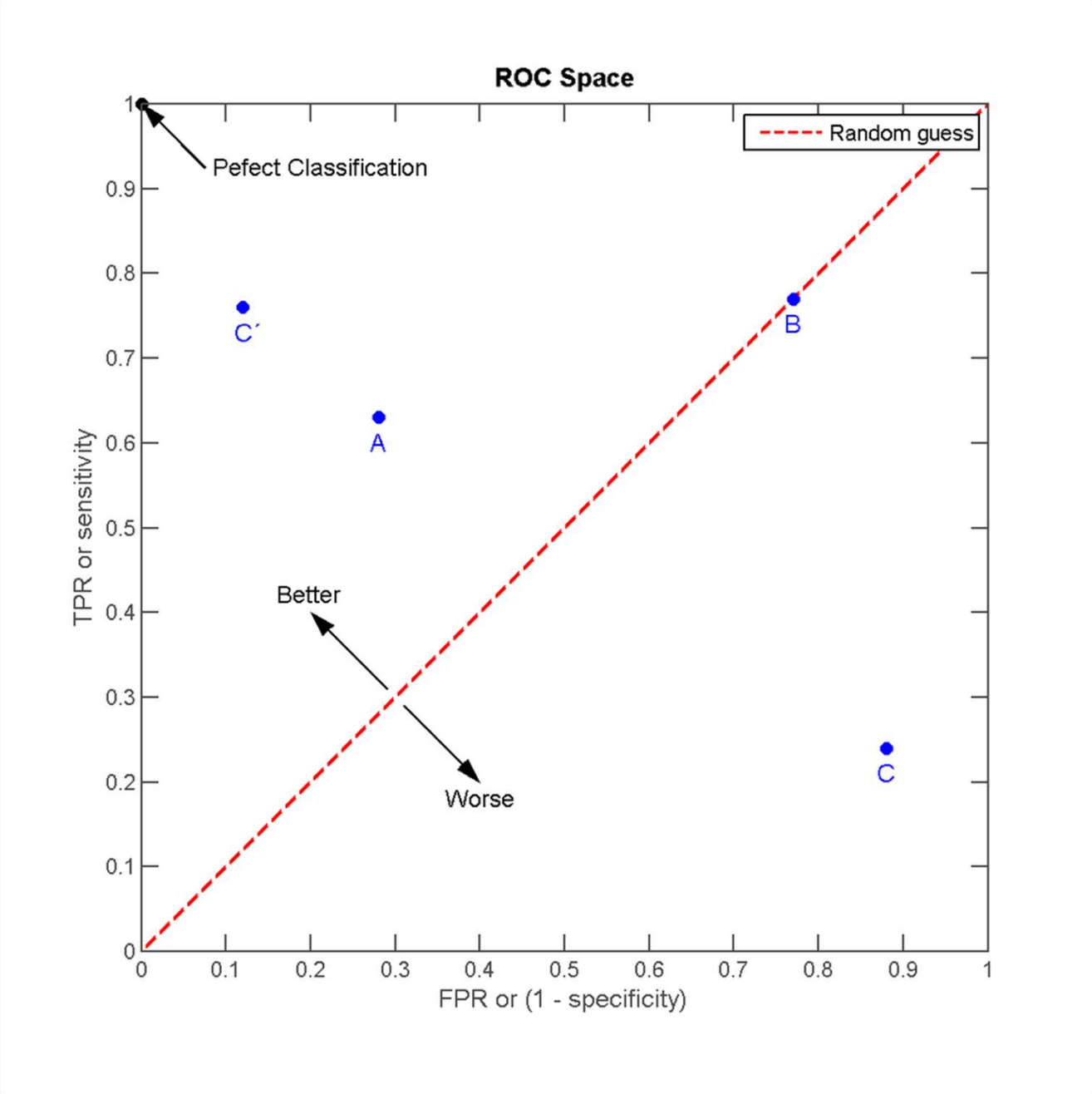
- **Ratio de verdaderos positivos (TPR)** = **Sensibilidad = Recall** = TP / (TP + FN)

- **Ratio de falsos positivos (FPR)** = FP / (FP + TN)



https://www.debatesiesa.com/una-herramienta-para-afinar-criterios-de-decision/

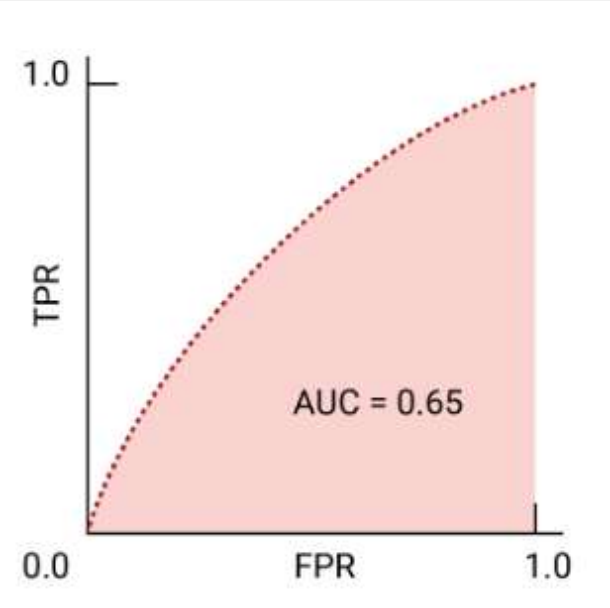
Algoritmos de clasificación – Curva ROC

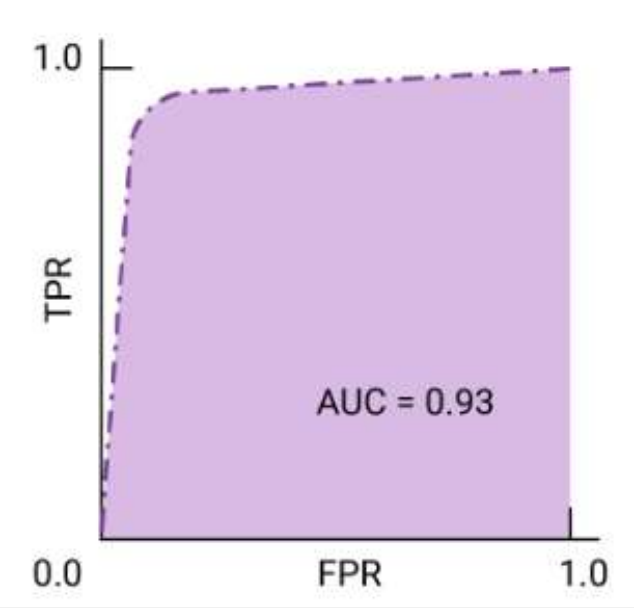


https://es.wikipedia.org/wiki/Curva_ROC#/media/Archivo:ROC_space-2.png

Algoritmos de clasificación – Curva ROC - AUC

AUC (Area Under the Curve o Área Bajo la Curva) es una métrica que se utiliza para evaluar el rendimiento de un modelo de clasificación binaria.

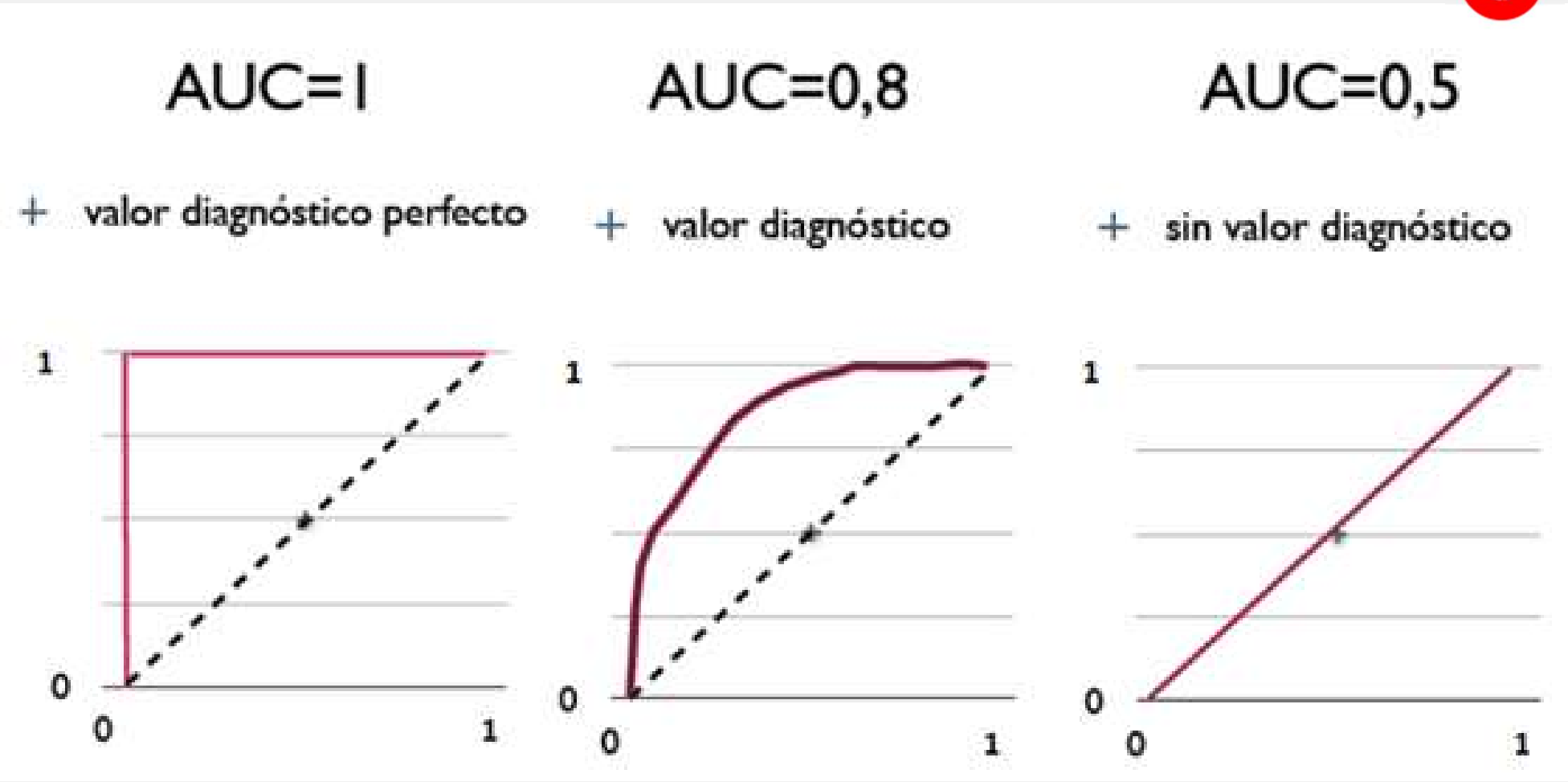




https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=es-419

El AUC es una métrica valiosa para evaluar el desempeño de dos modelos distintos, especialmente cuando el conjunto de datos es equilibrado. Por lo general, el modelo que presenta una mayor área bajo la curva se considera el más efectivo.

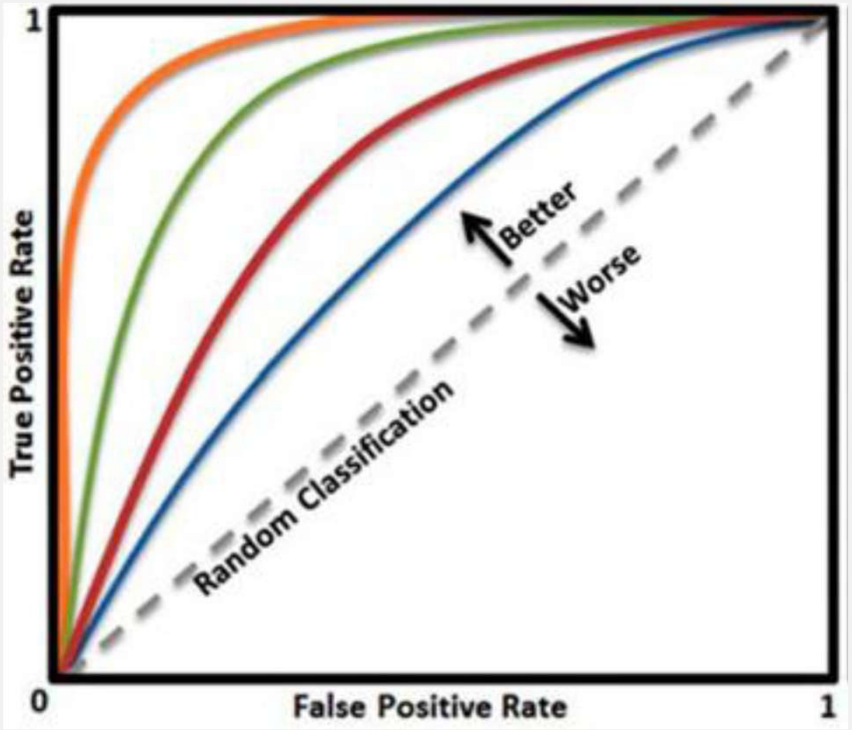
Algoritmos de clasificación – Curva ROC - AUC

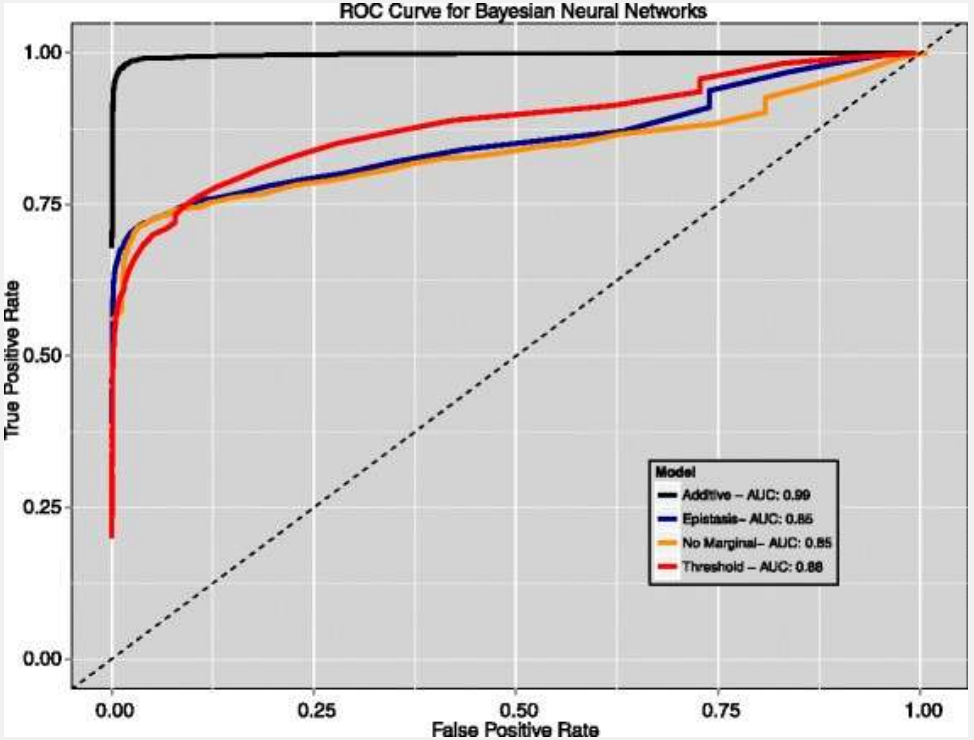


https://es.wikipedia.org/wiki/Curva_ROC

Algoritmos de clasificación – Curva ROC - AUC

El principal objetivo de la curva ROC y AUC es evaluar el rendimiento de un modelo de clasificación binaria en función de distintos umbrales de decisión y comparar varios modelos para determinar cuál es mejor en términos de su capacidad para distinguir entre clases positivas y negativas.

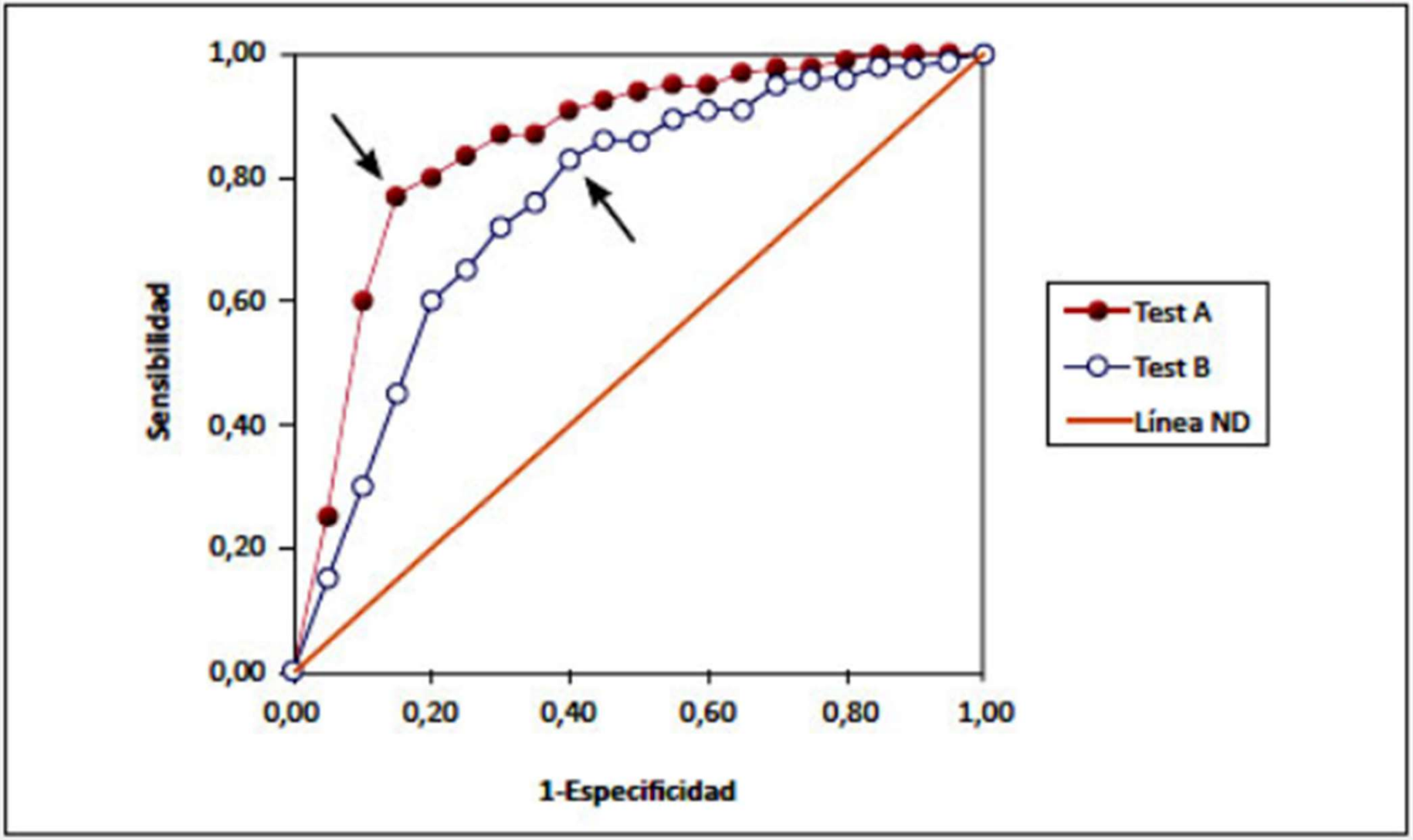




https://openi.nlm.nih.gov/

https://pmc.ncbi.nlm.nih.gov/articles/PMC4256933/figure/Fig6/

Algoritmos de clasificación – Curva ROC – AUC - Ejemplos



https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-10182012000200003

ALGORITMOS – ACTIVIDAD GRUPAL

- Por grupo revisar el archivo Laboratorio_Actividad_1.ipynb que se encuentra en la sección aplica / Recursos de aprendizaje.
- Realizar una presentación donde se explique lo resuelto.

CONCLUSIONES

01 La regresión lineal predice variables continuas, mientras que las métricas de clasificación se utilizan para variables categóricas.

02 Se evalúa la regresión lineal mediante MSE y R^2, mientras que la clasificación se mide con precisión, recall y F1-score.

03 El objetivo de la regresión lineal es establecer relaciones lineales, mientras que la clasificación busca asignar etiquetas correctas.

04 La regresión lineal proporciona coeficientes interpretables, mientras que las métricas de clasificación indican la precisión del modelo en sus predicciones.

BIBLIOGRAFÍA

- Ebac. Regresión Lineal: teoría y ejemplos. https://ebac.mx/blog/regreson-lineal

- Regresión lineal simple
  https://en.wikipedia.org/wiki/Linear_regression
  https://towardsdatascience.com/linear-regression-detailed-view-ea73175f6e86
  https://www.youtube.com/watch?v=zPG4NjIkCjc

- Regresión lineal múltiple
  http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm
  https://psychstat3.missouristate.edu/Documents/MultiBook3/Mlt06.htm

- Brownlee, J. (diciembre, 2017). Difference Between Classification and Regression in Machine Learning.