

INTELIGENCIA ARTIFICIAL CON DEEP LEARNING

ING. JORGE ALBERTO CASTELLANOS



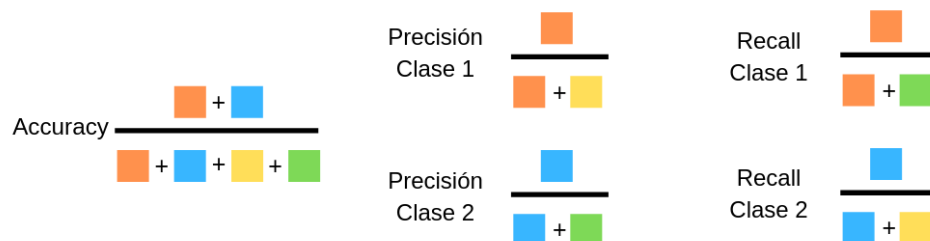
Universidad de
La Sabana

FACULTAD DE INGENIERÍA

Métricas de desempeño

	Predicción Clase 1	Predicción Clase 2
Valor real Clase 1	Aciertos True Positive Clase 1	Fallos False Positive Clase 2
Valor real Clase 2	Fallos False Positive Clase 1	Aciertos True Positive Clase 2

- Que son las métricas de desempeño:
 - Evalúan la eficacia y eficiencia de los modelos y algoritmos.
 - Forma cuantitativa de medir qué tan bien un modelo está realizando una tarea específica



Propósito

- **Evaluación de Modelos:** Las métricas de desempeño permiten a los científicos de datos y a los ingenieros de IA evaluar y comparar diferentes modelos para determinar cuál funciona mejor para una tarea dada.
- **Optimización de Modelos:** Estas métricas guían el proceso de ajuste y optimización de los hiperparámetros de un modelo para mejorar su rendimiento.
- **Toma de Decisiones:** Las métricas proporcionan información valiosa para tomar decisiones informadas sobre la implementación y el uso de modelos de IA en aplicaciones del mundo real.

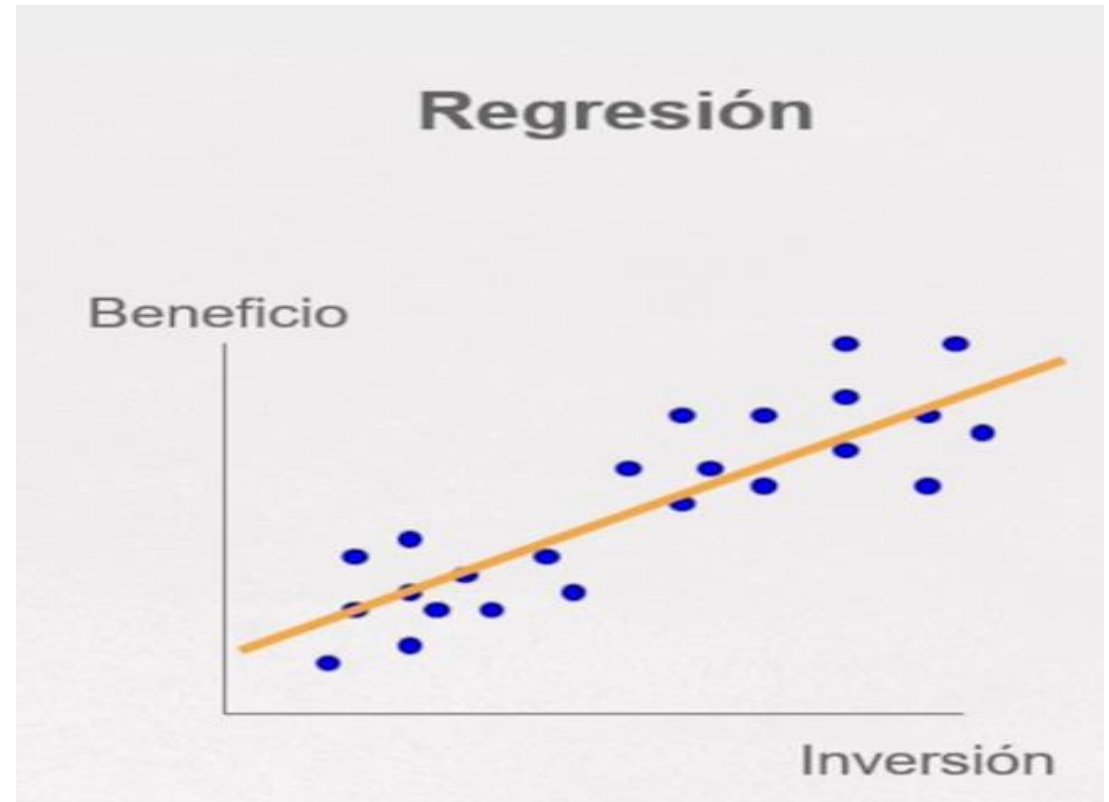
Tipos de Métricas “Clasificación”

1. Accuracy (Exactitud)
2. Precision (Precisión o Valor Predictivo Positivo)
3. Recall (Sensibilidad o Tasa de Verdaderos Positivos)
4. F1-Score
5. Matriz de Confusión
6. Curva ROC y AUC (Área Bajo la Curva)



Tipos de Métricas “Regresión”

1. Error Absoluto Medio (MAE - Mean Absolute Error)
2. Error Cuadrático Medio (MSE - Mean Squared Error)
3. Raíz del Error Cuadrático Medio (RMSE - Root Mean Squared Error)
4. Coeficiente de Determinación (R^2 Score)



Definición términos

- **TP(True Positive):**

- El modelo predice que algo es positivo, y realmente lo es.

Ejemplo: El modelo predice que un paciente tiene una enfermedad, y el paciente realmente la tiene.

- **TN(True Negative):**

- El modelo predice que algo es negativo, y realmente lo es.

Ejemplo: El modelo predice que un paciente no tiene una enfermedad, y el paciente realmente no la tiene.

- **FP(False Positive):**

- El modelo predice que algo es positivo, pero realmente es negativo.

- **Ejemplo:** El modelo predice que un paciente tiene una enfermedad, pero el paciente realmente no la tiene (también conocido como "falsa alarma" o "error de tipo I").

- **FN(False Negative):**

- El modelo predice que algo es negativo, pero realmente es positivo.

- **Ejemplo:** El modelo predice que un paciente no tiene una enfermedad, pero el paciente realmente la tiene (también conocido como "error de tipo II").

1. Accuracy (Exactitud)

- Es una de las métricas de evaluación más utilizadas en problemas de clasificación. Mide la proporción de predicciones correctas sobre el total de muestras evaluadas. Se expresa matemáticamente como:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

1. Ejemplo Accuracy (Exactitud)

Supongamos que tenemos un modelo de clasificación binaria que predice si un paciente tiene una enfermedad (*positivo*) o no (*negativo*). Evaluamos el modelo con **100 pacientes**, obteniendo los siguientes resultados:

- 50 pacientes realmente tenían la enfermedad, y el modelo detectó correctamente 45 de ellos (**TP = 45**).
- 50 pacientes no tenían la enfermedad, y el modelo predijo correctamente 40 de ellos como sanos (**TN = 40**).
- El modelo clasificó erróneamente a 10 pacientes sanos como enfermos (**FP = 10**).
- El modelo no detectó la enfermedad en 5 pacientes que realmente la tenían (**FN = 5**).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

1. Limitaciones Accuracy (Exactitud)

- Clases desbalanceadas
 - Ejemplo: Fraudes en transacciones. 99% 1%
- Cuando Usar:
 - Las clases están **balanceadas** en la distribución de datos.
 - Las consecuencias de **falsos positivos y falsos negativos son similares**.
 - Es necesario tener una **visión general** del rendimiento del modelo.

2. Precision (Precisión o Valor Predictivo Positivo)

La **precisión** es una métrica fundamental en problemas de clasificación, especialmente cuando **el costo de los falsos positivos es alto**.

- Se define como la proporción de instancias correctamente clasificadas como positivas con respecto al total de instancias predichas como positivas.

$$Precision = \frac{TP}{TP + FP}$$



"De todos los casos que el modelo clasificó como positivos, ¿cuántos realmente lo eran?"

2.Ejemplo Precisión

Supongamos que estamos desarrollando un **modelo de detección de spam en correos electrónicos**. Evaluamos el modelo con **100 correos**, donde:

- 30 correos son realmente spam (**Clase Positiva**).
- 70 correos son legítimos (**Clase Negativa**).

Resultados del modelo:

- Predice 25 correos como spam, de los cuales **20 son realmente spam (TP = 20)**.
- Predice 5 correos legítimos como spam (**FP = 5**).

$$Precision = \frac{TP}{TP + FP}$$

2. Precisión



La precisión es clave cuando los **falsos positivos son costosos o peligroso**



Diagnóstico Médico: Si un test de cáncer tiene baja precisión, etiquetará erróneamente a muchas personas sanas como enfermas (FP), generando ansiedad y pruebas médicas innecesarias.



Sistemas de Recomendación: En motores de búsqueda o plataformas de streaming, una alta precisión garantiza que los resultados sean relevantes y no una lista de recomendaciones irrelevantes.



Detección de Fraude: Un sistema que detecta fraudes con baja precisión marcaría muchas transacciones legítimas como fraudulentas (FP), causando inconvenientes a los usuarios.

2. Precisión

Limitaciones:

- La precisión **no considera los falsos negativos (FN)**, lo que puede ser problemático si perder casos positivos es peligroso.

Cuando Usar la Precisión

- Cuando los falsos positivos son costosos o problemáticos.
- Cuando queremos asegurarnos de que los positivos detectados sean confiables.
- Cuando el modelo es parte de un sistema donde la intervención humana validará los resultados

3. Recall (Sensibilidad)

- La sensibilidad, también conocida como recall o tasa de verdaderos positivos (TPR - True Positive Rate), mide qué tan bien el modelo encuentra los casos positivos reales.

$$Recall = \frac{TP}{TP + FN}$$



"De todos los casos que realmente eran positivos, ¿cuántos logré detectar?"

- Un **recall alto** significa que el modelo detecta la mayoría de los casos positivos, aunque pueda generar más falsos positivos.
- Un **recall bajo** significa que el modelo está dejando pasar muchos positivos sin detectarlos.

3. Ejemplo Recall

Supongamos que estamos desarrollando un modelo para detectar pacientes con una enfermedad. Evaluamos el modelo en **100 pacientes**, donde:

- 30 pacientes **realmente tienen la enfermedad (positivos reales)**.
- 70 pacientes **están sanos (negativos reales)**.

El modelo hace las siguientes predicciones:

- Detecta correctamente **25 pacientes enfermos (TP = 25)**.
- No detecta la enfermedad en **5 pacientes enfermos (FN = 5)**.

$$Recall = \frac{TP}{TP + FN}$$

3. Recall

El recall es clave en situaciones donde los falsos negativos (FN) son peligrosos o costosos. Algunos ejemplos incluyen:

- **Diagnóstico Médico:** Si un modelo de detección de cáncer tiene bajo recall, podría no detectar muchos casos de cáncer, lo que sería crítico.
- **Detección de Fraudes:** Un sistema de seguridad con bajo recall podría dejar pasar muchas transacciones fraudulentas sin detectarlas.
- **Sistemas de Búsqueda y Recuperación de Información:** En un motor de búsqueda, un alto recall significa que se recuperan la mayoría de los documentos relevantes.

3. Recall

Limitaciones:

- Si bien el recall mide cuántos positivos detectamos, **no nos dice si esos positivos son correctos.**

Cuando Usar Recall

- Cuando es más importante encontrar todos los casos positivos que evitar falsos positivos.
- Cuando los falsos negativos tienen un impacto grave (medicina, seguridad, fraude).
- Cuando la recuperación de información es prioritaria (búsqueda en bases de datos, inteligencia artificial en textos).

4. F1-Score

El **F1-score** es una métrica que combina **precisión** (*precision*) y **recall** en una sola medida. Es particularmente útil cuando se necesita un equilibrio entre ambas métricas y cuando las clases tienen algún desbalance.

- **F1-score** se define como la media armónica de precisión y recall:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$\text{Precisión (Precision)} = \frac{TP}{TP+FP}$$

$$\text{Recall (Recall)} = \frac{TP}{TP+FN}$$

💡 "¿Cuál es el mejor equilibrio entre precisión y recall en mi modelo?"

4. Ejemplo F1-Score

Supongamos que estamos desarrollando un modelo de detección de fraudes en transacciones bancarias.

- De 1000 transacciones, 50 son fraudes reales y 950 son legítimas.
- Nuestro modelo predice 60 fraudes, de los cuales:
 - 40 son fraudes reales (TP = 40).
 - 20 son transacciones legítimas erróneamente clasificadas como fraude (FP = 20).
 - 10 fraudes reales no fueron detectados (FN = 10)

4. F1-Score

Limitaciones:

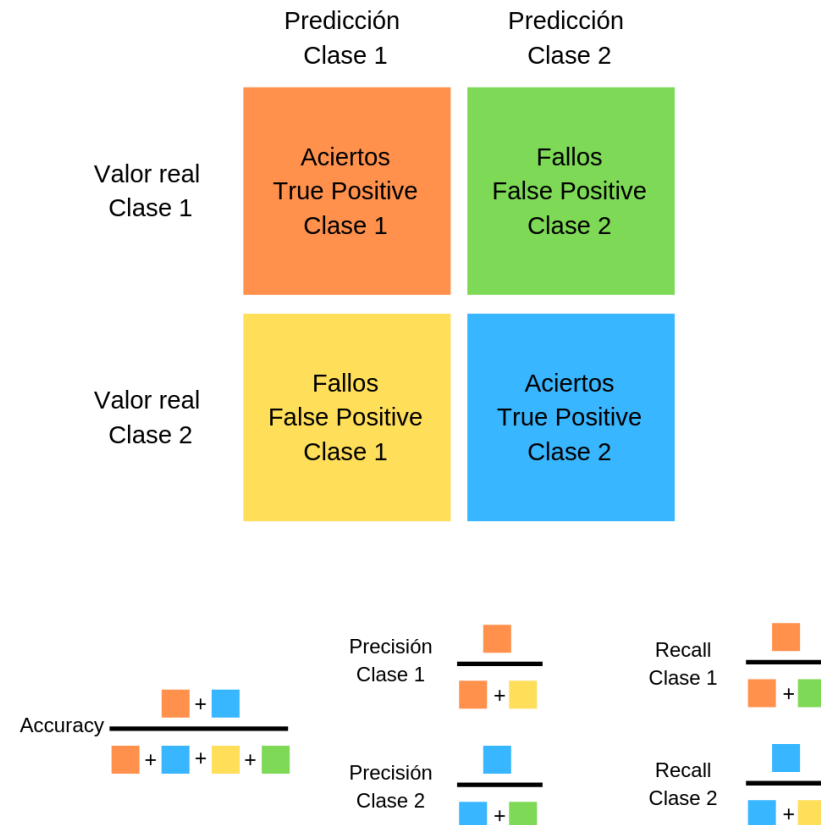
- No Considera la Distribución de Clases.
- No Diferencia la Importancia de Falsos Positivos y Falsos Negativos.
- No Mide el Rendimiento Global del Modelo.
- No Refleja Bien el Desempeño en Modelos con Umbrales Ajustables.
- No Considera la Confianza de las Predicciones

Cuando Usar Recall

- **El F1-score es útil en problemas donde la precisión y el recall son igualmente importantes.**
- Cuando las clases están desbalanceadas (ej. detección de fraudes, enfermedades raras).
- Cuando precisión y recall son igual de importantes.
- Cuando queremos una única métrica que combine ambos valores.







5. Matriz de Confusión

- La matriz de confusión es una herramienta fundamental para evaluar el rendimiento de un modelo de clasificación. Permite analizar cuántas predicciones fueron correctas e incorrectas y qué tipos de errores cometió el modelo. Se utiliza tanto en clasificación binaria como en clasificación multiclase.



5. Matriz de Confusión

Ejemplo

	Predicho: Perro	Predicho: Gato	Predicho: Pájaro
Real: Perro	40 	5 	3 
Real: Gato	7 	50 	2 
Real: Pájaro	6 	3 	55 


En esta tabla:

- 40 veces se clasificó bien un Perro.
- 5 veces se confundió un Perro con un Gato.
- 3 veces se confundió un Perro con un Pájaro.

5. Matriz de Confusión

Ejemplo

	Predicho: Spam	Predicho: No Spam
Real: Spam	300 	20 
Real: No Spam	50 	630 

- ✓ FN (20 correos legítimos en spam): Puede ser molesto, pero no crítico.
- ✓ FP (50 correos spam no detectados): Correo basura en la bandeja de entrada.
-  Impacto: En correo, un balance entre precisión y recall es clave.

5. Matriz de Confusión

¿Cuándo usar la matriz de confusión?



Para evaluar modelos de clasificación y analizar los tipos de errores.



Cuando se necesita visualizar el rendimiento más allá de una única métrica como la exactitud.



Para calcular métricas derivadas como precisión, recall y F1-score.



En problemas de clasificación multiclase para entender qué clases se confunden más.



Para optimizar hiperparámetros y mejorar el desempeño del modelo en clases específicas.

5. Matriz de Confusión

Limitaciones de la Matriz de Confusión

No es útil con datos desbalanceados:
La exactitud puede ser engañosa si
una clase domina.

No indica la calidad del modelo: Solo
muestra errores sin evaluar su impacto.

Escalabilidad limitada: Difícil de
interpretar en clasificación multiclase
con muchas clases.

No mide la confianza en las
predicciones: No muestra
probabilidades ni niveles de certeza.

No diferencia entre errores graves y
leves: Todos los errores se consideran
iguales.

6. Curva ROC y AUC

La Curva ROC (Receiver Operating Characteristic) es una gráfica que evalúa el rendimiento de un modelo de clasificación binaria al mostrar la relación entre la tasa de verdaderos positivos (TPR, Sensibilidad o Recall) y la tasa de falsos positivos (FPR) en diferentes umbrales de decisión.

Fórmulas clave:

Tasa de Verdaderos
Positivos (TPR o
Recall)

$$TPR = \frac{TP}{TP + FN}$$

Tasa de Falsos
Positivos (FPR)

$$FPR = \frac{FP}{FP + TN}$$

El **Área Bajo la Curva (AUC - Area Under the Curve)** mide la capacidad del modelo para diferenciar entre clases. Su valor varía entre **0 y 1**, donde:

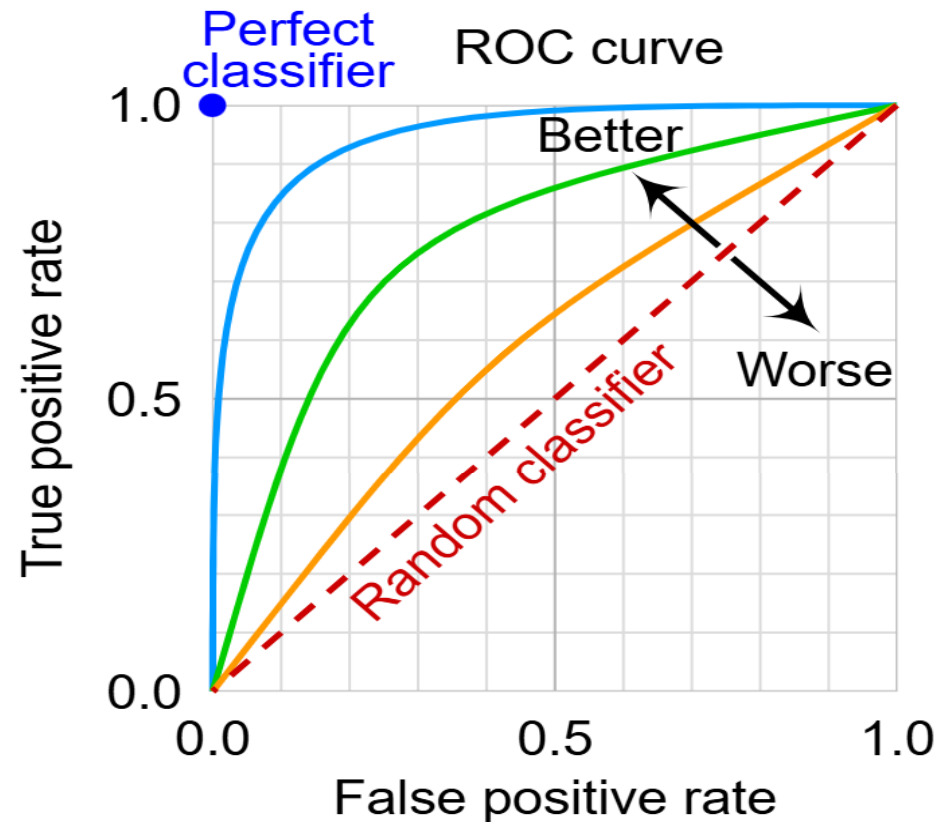
✓ **AUC = 1.0** → Modelo perfecto.

⚠ **AUC = 0.5** → Modelo aleatorio (sin capacidad de clasificación).

✗ **AUC < 0.5** → Modelo peor que el azar.

6. Curva ROC y AUC

Ejemplos Gráficos - Representación Visual de la Curva ROC



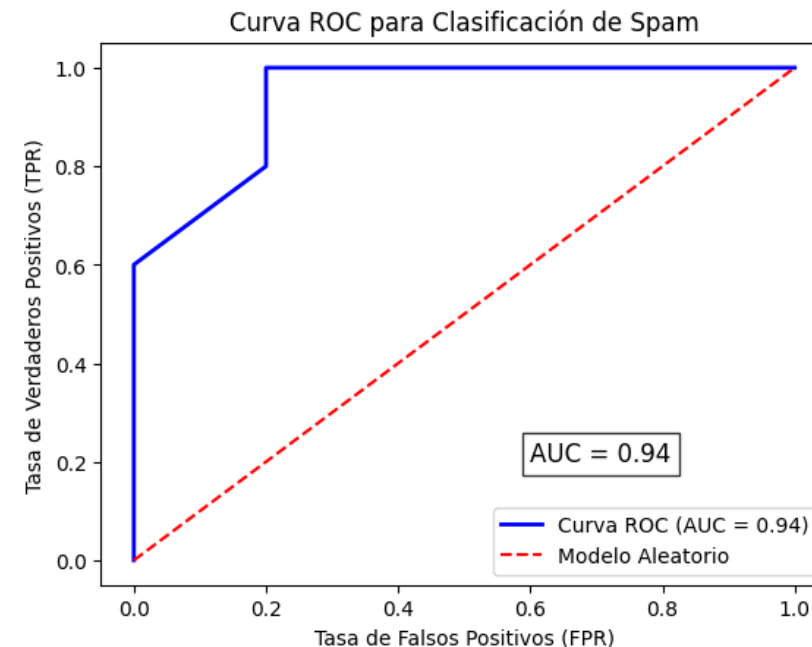
6. Curva ROC y AUC

En un caso tenemos un modelo que predice los casos de spam de un correo donde

Clasificación de correos spam = [0, 1, 0, 1, 1, 0, 0, 1, 0, 1]

Probabilidad de que sean spam= [0.3, 0.6, 0.5, 0.55, 0.45, 0.4, 0.35, 0.5, 0.4, 0.65]

Umbral	FPR (Tasa de Falsos Positivos)	TPR (Tasa de Verdaderos Positivos)
0,65	0	0,2
0,55	0	0,6
0,5	0,2	0,8
0,45	0,2	1
0,4	0,6	1
0,3	1	1



6. Curva ROC y AUC

¿Cuándo Usar la Curva ROC y AUC?

- ✓ **Cuando el balance entre Falsos Positivos y Falsos Negativos es clave** (salud, fraude, seguridad).
- ✓ **Para comparar modelos de clasificación** y elegir el más efectivo.
- ✓ **Cuando el umbral de decisión puede ajustarse** y queremos evaluar su impacto.
- ✓ **En modelos probabilísticos**, ya que mide el rendimiento en distintos umbrales.

Limitaciones

- ✗ **No es útil en datos desbalanceados:** Un AUC alto puede ser engañoso si una clase es mucho más frecuente.
- ✗ **No muestra el mejor umbral:** Se necesita otra métrica para elegir el umbral óptimo.
- ✗ **No diferencia entre tipos de errores:** No da prioridad a minimizar FN o FP según la aplicación.
- ✗ **Puede ser insensible a mejoras en TPR o FPR:** En algunos casos, un AUC similar no refleja una mejora en el modelo.

1. Error Absoluto Medio (MAE)

El Error Absoluto Medio (MAE) es una métrica utilizada para evaluar el desempeño de un modelo de regresión midiendo el error promedio entre las predicciones y los valores reales, sin considerar la dirección del error.

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

Donde:

- y_i es el valor real en la observación i .
- \hat{y}_i es el valor predicho por el modelo en la observación i .
- n es el número total de observaciones.

 **Interpretación:**

- Un **MAE bajo** indica que las predicciones están cerca de los valores reales.
- Un **MAE alto** indica que las predicciones tienen un error grande.

1. Error Absoluto Medio (MAE)

Ejemplo 1: Predicción de Temperatura

Día	Real (y_i)	Predicho(\hat{y}_i)	Error Absoluto
1	30°C	28°C	2°C
2	25°C	27°C	2°C
3	32°C	29°C	3°C
4	28°C	30°C	2°C
MAE	-	-	2.25°C

- **Conclusión:** El modelo tiene un error promedio de **2.25°C** en sus predicciones.

Ejemplo 2: Predicción de Ventas

Día	Real (y_i)	Predicho(\hat{y}_i)	Error Absoluto
1	500	480	20
2	600	620	20
3	700	650	50
4	550	530	20
MAE	-	-	27.5 ventas

- **Conclusión:** En promedio, el modelo tiene un error de **27.5 ventas por semana**.

1. Error Absoluto Medio (MAE)

¿Cuándo usar MAE?

- ✓ Cuando se necesita interpretar el error en unidades reales, lo que facilita la comprensión.
- ✓ En modelos de regresión lineal y redes neuronales para evaluar el desempeño.
- ✓ En problemas donde los errores positivos y negativos tienen igual importancia.
- ✓ Cuando se buscan métricas robustas ante grandes valores atípicos, ya que MAE no los amplifica tanto como MSE.

Limitaciones

- ✗ No penaliza los errores grandes
- ✗ No diferencia si el error es positivo o negativo, lo que puede ser importante en algunas aplicaciones.
- ✗ No es derivable en algunos puntos, lo que dificulta su uso en optimización basada en gradientes.
- ✗ No es ideal para comparar modelos con diferentes escalas de valores, ya que el MAE depende de la magnitud de los datos.

2. Error Cuadrático Medio (MSE - Mean Squared Error)

El Error Cuadrático Medio (MSE) es una métrica utilizada en problemas de regresión para medir la diferencia promedio al cuadrado entre los valores predichos y los valores reales. Penaliza los errores grandes más que los pequeños, lo que lo hace sensible a valores atípicos.

Fórmula del MSE:

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

Donde:

- y_i es el valor real en la observación i .
- \hat{y}_i es el valor predicho por el modelo en la observación i .
- n es el número total de observaciones.

 **Interpretación:**

- Un **MSE más bajo** indica que el modelo tiene predicciones más precisas.
- Un **MSE más alto** sugiere que el modelo tiene mayores errores de predicción.

2. Error Cuadrático Medio (MSE - Mean Squared Error)

Ejemplo 1: Predicción de Temperatura

Un modelo predice la temperatura diaria, y se comparan las predicciones con los valores reales:

Día	Real (y_i)	Predicho (\hat{y}_i)	Error ($y_i - \hat{y}_i$)	Error cuadrado
1	30°C	28°C	2°C	4
2	25°C	27°C	-2°C	4
3	32°C	29°C	3°C	9
4	28°C	30°C	-2°C	4
MSE	-	-	-	5.25°C ²

 **Conclusión:** El modelo tiene un MSE de 5.25, lo que indica que en promedio los errores elevados al cuadrado son 5.25°C².

2. Error Cuadrático Medio (MSE - Mean Squared Error)

¿Cuándo Usar MSE?

- ✓ Cuando se desea penalizar errores grandes más que errores pequeños, ya que el cuadrado del error aumenta el impacto de valores atípicos.
- ✓ En modelos de regresión lineal y redes neuronales, ya que es una métrica diferenciable y se puede usar en optimización.
- ✓ Cuando se necesita comparar diferentes modelos de regresión y seleccionar el que minimiza el error.
- ✓ En problemas donde los errores grandes son más costosos, como predicción de precios o pronósticos financieros.

Limitaciones

- ✗ **Penaliza fuertemente los valores atípicos**, ya que los errores grandes se elevan al cuadrado y afectan mucho la métrica.
- ✗ **No tiene una escala interpretable**, ya que el resultado está en unidades cuadradas y no en la misma escala que los datos originales.
- ✗ **Menos robusto en comparación con MAE (Mean Absolute Error)** cuando los datos contienen outliers.
- ✗ **No distingue entre errores positivos y negativos**, lo que puede ser problemático en algunos contextos.

3. Raíz del Error Cuadrático Medio (RMSE - Root Mean Squared Error)

La Raíz del Error Cuadrático Medio (RMSE) es una métrica de evaluación utilizada en problemas de regresión para medir la magnitud promedio del error en las predicciones del modelo. Se obtiene extrayendo la raíz cuadrada del Error Cuadrático Medio (MSE), lo que permite expresar el error en las mismas unidades que los datos originales.

Donde:

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

- y_i es el valor real en la observación i .
- \hat{y}_i es el valor predicho por el modelo en la observación i .
- n es el número total de observaciones.

Interpretación:

- Un **RMSE más bajo** indica que el modelo tiene predicciones más precisas.
- Un **RMSE más alto** sugiere que el modelo tiene mayores errores de predicción.

3. Raíz del Error Cuadrático Medio (RMSE - Root Mean Squared Error)

Ejemplo: Predicción de Ventas

Un modelo predice el número de ventas diarias en una tienda, y se comparan las predicciones con los valores reales:

Día	Real (y_i)	Predicho (\hat{y}_i)	Error ($y_i - \hat{y}_i$)	Error cuadrado
1	100	95	5	25
2	120	125	-5	25
3	110	105	5	25
4	115	120	-5	25
MSE	-	-	-	25
RMSE	-	-	-	5


Conclusión: El modelo tiene un RMSE de 5, lo que significa que en promedio el error de predicción es 5 ventas.

3. Raíz del Error Cuadrático Medio (RMSE - Root Mean Squared Error)

¿Cuándo Usar RMSE?

- ✓ Cuando se necesita un error en las mismas unidades que los datos originales, facilitando su interpretación.
- ✓ En problemas de regresión donde se desea evaluar la precisión del modelo, como predicción de precios, temperatura o demanda.
- ✓ Cuando se quiere penalizar más los errores grandes, ya que el cuadrado de los errores amplifica su impacto.
- ✓ En modelos de machine learning como redes neuronales y regresión lineal, donde es una métrica comúnmente utilizada.

Limitaciones

- ✗ **Es sensible a valores atípicos**, ya que los errores grandes se elevan al cuadrado antes de tomar la raíz.
- ✗ **No indica la dirección del error**, lo que significa que no diferencia si las predicciones son sistemáticamente más altas o más bajas que los valores reales.
- ✗ **Puede sobreestimar el impacto de errores grandes**, lo que no siempre es deseable en ciertos contextos.
- ✗ **No es ideal para comparar modelos en conjuntos de datos con escalas diferentes**, ya que el RMSE depende de las unidades de los datos originales.

Comparación con otras métricas

Métrica	Interpretación	Sensibilidad a errores grandes
MAE	Error promedio absoluto	Baja
MSE	Error cuadrático medio (eleva los errores al cuadrado)	Alta
RMSE	Raíz del MSE, devuelve el error en las mismas unidades que los datos	Alta

4. Coeficiente de Determinación (R² Score)

El Coeficiente de Determinación (R² Score) es una métrica utilizada en modelos de regresión para medir qué tan bien el modelo explica la variabilidad de los datos. Evalúa la proporción de la variabilidad total en la variable dependiente que puede ser explicada por las variables independientes del modelo.




Fórmula del R² Score:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Donde:

- $SS_{res} = \sum (y_i - \hat{y}_i)^2$ es la suma de los errores al cuadrado (error residual).
- $SS_{tot} = \sum (y_i - \bar{y})^2$ es la suma de las desviaciones al cuadrado (total de variabilidad en los datos).
- y_i son los valores reales.
- \hat{y}_i son los valores predichos por el modelo.
- \bar{y} es el valor promedio de los datos reales.

Interpretación:

- $R^2=1$  → El modelo explica **toda la variabilidad** de los datos.
- $R^2=0$  → El modelo **no explica nada** de la variabilidad de los datos.
- $R^2<0$  → El modelo es **peor que una predicción basada solo en la media**.

4. Coeficiente de Determinación (R^2 Score)

Ejemplo: Predicción de Ingresos Mensuales

Supongamos que tenemos un modelo de regresión que predice los ingresos mensuales de trabajadores en función de su experiencia en años. Se comparan los valores predichos con los valores reales:

Trabajador	Ingreso Real (y_i)	Ingreso Predicho (\hat{y}_i)	Error Cuadrático ($y_i - \hat{y}_i$) ²
1	3000	3100	10000
2	4000	3900	10000
3	5000	5100	10000
4	6000	5900	10000
Total (SS_{res})	-	-	40000

Supongamos que la variabilidad total (SS_{res}) en los datos es 100,000.

$$R^2 = 1 - \frac{40,000}{100,000} = 0.6$$


Conclusión: El modelo explica el 60% de la variabilidad de los ingresos mensuales con base en la experiencia.

4. Coeficiente de Determinación (R^2 Score)

¿Cuándo Usar R^2 ?

- ✓ Cuando se necesita evaluar la calidad del ajuste de un modelo de regresión.
- ✓ Para comparar distintos modelos de regresión, eligiendo aquel con el mayor R^2 .
- ✓ Cuando el objetivo es explicar la relación entre variables, midiendo cuánta variabilidad de la variable dependiente es explicada por las independientes.
- ✓ En econometría, predicción financiera y análisis de tendencias, donde se usan modelos de regresión lineal.

Limitaciones

- ✗ **No mide la precisión del modelo:** Un R^2 alto no implica que el modelo haga buenas predicciones individuales.
- ✗ **No detecta sobreajuste:** Un modelo con demasiadas variables puede tener un R^2 alto pero ser ineficiente en datos nuevos.
- ✗ **No siempre mejora con más variables:** Algunas variables pueden no aportar información útil y solo inflar artificialmente el R^2 .
- ✗ **No es adecuado para modelos no lineales:** Puede ser engañoso cuando se usa en modelos de regresión no lineal sin modificaciones.

Preprocesamiento de datos en Lenguaje natural NLP

- **Que es el lenguaje:**

Un Lenguaje es un conjunto potencialmente infinito de oraciones y sentencias de palabras construidas mediante reglas gramaticales, fonéticas y de significación que rigen el propio lenguaje.



Tipos de Lenguaje

01



Nace de manera espontanea por la necesidad de comunicarse. (Idiomas, ingles, español, chino)

Lenguaje Natural

Lenguaje
Formal

Lenguajes creados para una situación particular.
(Matemático, lógico, musical, programación)



02

03



Lenguajes creados antes de ser usado por los parlantes, como una mezcla de natural y formal.

Lenguaje Artificial

¿Qué es el Preprocesamiento en NLP?

01

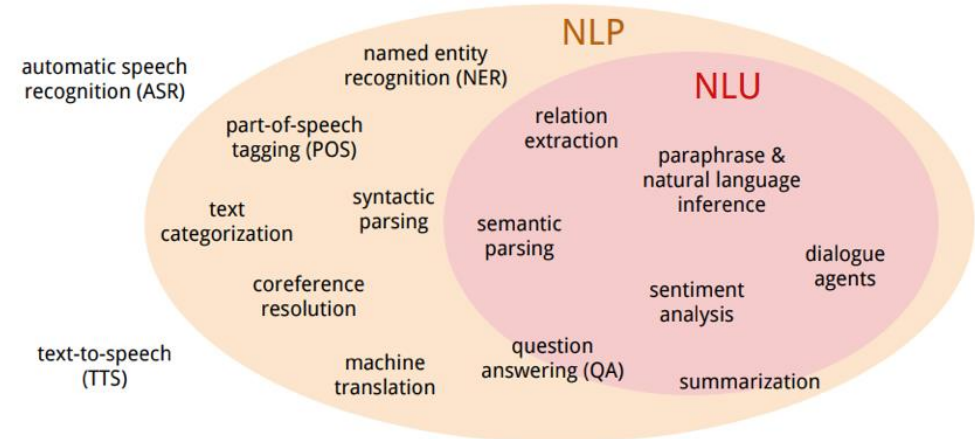
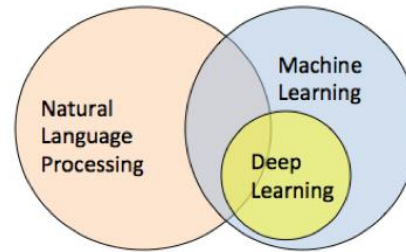
Definición

Transformación de texto en un formato adecuado para su análisis.

02

Objetivo

Tratar la interacción entre los lenguajes humanos (lenguajes naturales) y los dispositivos informáticos.



Campo que combina la **Informática**, la **Inteligencia Artificial** y la **Lingüística**;

Donde se puede encontrar

Recuperación de
información

Extracción y
categorización de
información

Análisis automático
de texto subjetivo
(Análisis de
sentimientos)

Traducción
automática

Generación del
lenguaje

Questions &
Answering
(Chatbots)

Donde se puede encontrar

- Recuperación de información
- Extracción y categorización de información
- Análisis automático de texto subjetivo (Análisis de sentimientos)
- Traducción automática
- Generación del lenguaje
- Questions & Answering (Chatbots)

Conceptos para el NLP

- Corpus: Colección de textos como puede ser un conjunto de artículos científicos, libros, tweets, críticas.
- Bag of Words (BoW): modelo para simplificar el contenido de un documento(s)
 - Sin gramática, ni orden de palabras. +ocurrencias de palabras
- Normalización: Poner el texto en igual de condiciones:
 - Convertir Mayúsculas o minúsculas
 - Eliminar puntuación
 - Convertir números a palabras
 - Eliminar palabras que no aporten al texto (Stop Word)

Conceptos para el NLP

- Tokenización: dividir el texto en unidades mas pequeñas llamadas **tokens** que pueden ser palabras, frases, símbolos u otros elementos significativos.
 - Segmentación: dividir en oraciones o párrafos
 - Tokenización: dividir grandes cadenas de texto en palabras.
- Stemming: Proceso de eliminar los afijos (sufijos, prefijos, infijos, circunflejos) de una palabra para obtener un tallo de palabra.
 - Caminando → Caminar
- Lematización: Proceso lingüístico, sustituy palabra flexionada (Plurales, verbos conjugados o femeninos) por su lema; como una palabra valida en el idioma.

Conceptos para el NLP

- Stop Word: Son palabras que no aportan nada al significado de las frases como las preposiciones, determinantes, etc.
- Part of speech (POS) Tagging: Asignar una etiqueta de categoría a las partes tokenizadas de una oración. El etiquetado POS más popular sería identificar palabras como sustantivos, verbos, adjetivos, etc.
 - Artículo o determinante
 - Sustantivo o nombre
 - Pronombre
 - Verbo
 - Adjetivo
 - Adverbio
 - Preposición
 - Conjunción
 - Interjección

Conceptos para el NLP

- n-grammas: A diferencia de la representación sin orden de una bolsa de palabras (bag of words), el modelado de n-gramas está interesado en preservar secuencias contiguas de N elementos de la selección de texto.