

Examen teórico 3er parcial

1. ¿Cuál es la diferencia entre métodos de aprendizaje supervisado y no supervisado?

El aprendizaje supervisado usa etiquetas, su objetivo es predecir salidas, lleva a cabo tareas de clasificación y regresión, el no supervisado no lleva etiquetas, su objetivo es encontrar patrones y lleva a cabo tareas de clustering y reducción dimensional.

2. Observa los siguientes dataset dummies:

¿En cuál dataset se pueden aplicar métodos de Aprendizaje supervisado y en cuál métodos de Aprendizaje no supervisado?

El dataset 1 tiene una variable de etiqueta (SleepStage), entonces se puede aplicar aprendizaje supervisado. El dataset 2 no tiene ninguna variable de etiqueta, solo tenemos características, por lo que el propósito sería descubrir patrones en los datos, entonces usaríamos el aprendizaje no supervisado.

3. ¿Cuál es la diferencia entre un problema de clasificación y uno de regresión?

La diferencia está en el tipo de salida; los de clasificación predicen categorías, su objetivo es asignar una etiqueta a cada entrada, y la salida esperada es una clase, en cambio, los de regresión predicen valores numéricos continuos, su objetivo es estimar un número a partir de las entradas, y la salida esperada por consecuencia es un número. Otra diferencia es que los de clasificación tienen clases finitas y los de regresión no.

4. ¿Cuál es la diferencia entre over-fitting y under-fitting?

El underfitting pasa cuando el dataset es tan pequeño que el modelo no puede detectar los patrones en los datos porque no puede aprender lo suficiente. El overfitting es cuando el dataset es muy complejo, quizá con demasiados parámetros o demasiados nodos, entonces pasa que el modelo memoriza los datos, así que va a funcionar muy bien con los datos que ya tiene, los que memorizó, pero no va a poder funcionar con nuevos. En el **underfitting** el error de entrenamiento y prueba son altos, y en el **overfitting**, el error de entrenamiento es muy bajo y el de prueba alto.

5. Observa la siguiente tabla con valores de accuracy en los dataset de entrenamiento y prueba

Ensayo	Accuracy Entrenamiento	Accuracy Prueba
1	99%	70%
2	60%	58%
3	65%	62%
4	95%	68%
5	92%	90%
6	100%	60%

Basado en los valores anteriores, ¿qué ensayo dirías que corresponde a casos con over-fitting, under-fitting o balanceado?

Ensayo 1 → overfitting

Ensayo 2 → Underfitting

Ensayo 3 → Un poco underfitting pero casi balanceado (un porcentaje de 65% es más aceptable que uno de 60%)

Ensayo 4 → Overfitting

Ensayo 5 → Balanceado

Ensayo 6 → Overfitting

6. Contesta lo siguiente

A. Escribe las fórmulas de accuracy, precision, recall y f1-score

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

Precisión → $\text{Precisión} = \frac{TP}{TP + FP}$

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

F1-score → $\text{F1-score} = 2 \left(\frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$

B. Con base a la siguiente tabla, calcula accuracy, precision, recall y f1-score

	Predicho Positivo	Predicho Negativo
Real Positivo	TP = 40	FN = 10
Real Negativo	FP = 5	TN = 45

Accuracy = $(40 + 45) / (40 + 45 + 5 + 10) = 85 / 100 = 0.85 = 85\%$

Precisión = $40 / (40 + 5) = 40 / 45 = 0.8889 = 88.89\%$

Recall = $40 / (40 + 10) = 40 / 50 = 0.80 = 80\%$

F1-score = $2((88.89 * 80) / (88.89 + 80)) = 2(0.7111 / 1.6889) = 2(0.4216) = 0.8432 = 84.32\%$

C. A continuación tienes 3 descripciones de métricas en un problema de clasificación, indica a qué métrica corresponde

Descripción	Métrica
Proporción de predicciones correctas (tanto positivas como negativas) sobre el total de predicciones realizadas.	Accuracy
Porcentaje de verdaderos positivos identificados correctamente, respecto al total de positivos reales.	Recall
Porcentaje de aciertos entre los casos que el modelo predijo como positivos.	Precisión

7. A continuación se presentan varias características de algoritmos de Machine Learning supervisado. Indica a qué método corresponde cada una de las siguientes afirmaciones:

- A) Análisis Discriminante
- B) Vecino más cercano (k-NN)
- C) Máquina de soporte vectorial (SVM)
- D) Árbol de decisión
- E) Bosques aleatorios

Nº	Característica	¿A qué método pertenece?
1	Asume que los datos de cada clase siguen una distribución normal (gaussiana)	A) Análisis Discriminante
2	No tiene fase de entrenamiento como tal; toda la predicción ocurre en tiempo real	B) Vecino más cercano (k-NN)
3	Busca el hiperplano que maximiza el margen entre clases	C) Máquina de soporte vectorial (SVM)
4	Realiza múltiples divisiones binarias para clasificar los datos	D) Árbol de decisión
5	Es sensible a la escala de las características, especialmente cuando se usa distancia euclidiana	B) Vecino más cercano (k-NN)
6	Utiliza funciones núcleo (kernel) para trabajar con datos no lineales	C) Máquina de soporte vectorial (SVM)
7	Calcula la probabilidad de pertenencia a una clase bajo modelos estadísticos	A) Análisis Discriminante
8	Clasifica observaciones comparándolas con las más cercanas en el espacio de características	B) Vecino más cercano (k-NN)
9	Utiliza múltiples árboles para reducir la varianza y mejorar la precisión	E) Bosques aleatorios

10	Puede sobreajustar fácilmente si el árbol es muy profundo	D) Árbol de decisión
11	Su versión más simple se basa en modelos lineales con varianzas iguales por clase	A) Análisis Discriminante
12	Puede usar núcleos como RBF o polinomiales para separar clases de manera no lineal	C) Máquina de soporte vectorial (SVM)
13	Tiende a mejorar la generalización comparado con un solo árbol	E) Bosques aleatorios
14	Es altamente interpretativo cuando se trabaja con dos clases y pocos atributos	D) Árbol de decisión
15	Es una combinación de muchos árboles construidos sobre subconjuntos aleatorios del conjunto de entrenamiento	E) Bosques aleatorios

8.¿Cuál es la diferencia entre Perceptron, Red Neuronal Artificial y Red Neuronal Convolutacional?

- Perceptrón: es la unidad básica de las redes neuronales, tiene una sola capa de salida, solo aprende problemas linealmente separables, no requiere preprocesamiento espacial, su sensibilidad a la posición del input es alta y su función de activación común es el escalón
- Red neuronal Artificial: es la red formada por varias capas de perceptrones conectados entre sí, tiene múltiples capas de entrada y salida, aprende relaciones no lineales, no requiere preprocesamiento espacial, su sensibilidad a la posición del input es alta y sus funciones de activación común son el ReLU, sigmoide y tanh
- Red Neuronal Convolutacional: es la red especializada en procesamiento de datos con estructura espacial, tiene capas convolucionales, de pooling y totalmente conectadas,

aprende patrones espaciales, requiere de preprocesamiento espacial, su sensibilidad a la posición del input es baja y sus funciones de activación común son el ReLU y softmax

9.En el contexto de aprendizaje de una red neuronal, ¿qué es un epoch?

Es una pasada completa por todo el conjunto de datos de entrenamiento, durante un epoch, la red toma todos los datos de entrenamiento, calcula la pérdida para cada batch y ajusta sus pesos y biases mediante retropropagación y descenso del gradiente.

10.Lista las principales funciones de activación y describe la diferencia entre las funciones sigmoid y softmax

Principales funciones de activación:

- Sigmoid → se usa en clasificación binaria, el rango de salida es (0,1)
- Tanh
- ReLU
- Leaky y ReLU
- Softmax → se usa en clasificación multiclase, capa de salida, el rango de salida es (0,1), suma = 1
- Swish