

# Examen 3er parcial. Examen teórico

Responde cada una de las siguientes preguntas y sube tu examen como se indica:

- 1) ¿Cuál es la diferencia entre métodos de aprendizaje supervisado y no supervisado?  
La principal diferencia es que mientras que en el aprendizaje supervisado se busca encontrar asignar una 'variable categórica' (como lo puede ser una clase), en el aprendizaje no supervisado se busca 'describir' (hablando coloquialmente) el dataset mediante ciertos métodos como el clustering, donde se .
- 2) Observa los siguientes dataset dummies:

ID	Conn_1	Conn_2	Conn_3	Conn_4	Conn_5	Conn_6	Conn_7	Conn_8	Conn_9	Conn_10
1	0.81	0.64	0.12	-0.22	0.45	0.03	0.66	0.58	0.20	0.11
2	0.35	0.57	0.14	-0.18	0.52	0.07	0.59	0.50	0.18	0.09
3	0.72	0.62	0.10	-0.31	0.40	0.01	0.64	0.55	0.22	0.15
4	0.11	0.33	0.05	-0.12	0.60	0.08	0.55	0.47	0.12	0.02
5	0.90	0.79	0.20	-0.40	0.33	-0.02	0.69	0.61	0.28	0.19
6	0.77	0.68	0.13	-0.25	0.49	0.00	0.60	0.53	0.24	0.13
7	0.36	0.41	0.07	-0.15	0.57	0.05	0.58	0.48	0.16	0.10
8	0.80	0.70	0.17	-0.28	0.42	0.02	0.63	0.56	0.26	0.18
9	0.39	0.50	0.09	-0.20	0.55	0.06	0.57	0.49	0.20	0.11
10	0.85	0.73	0.18	-0.35	0.38	0.01	0.67	0.60	0.27	0.17

ID	EEG_Δ (0.5–4Hz)	EEG_θ (4–8Hz)	EEG_α (8–12Hz)	EEG_β (12–30Hz)	EEG_γ (30–45Hz)	SleepStage
1	12.1	8.4	4.7	3.1	0.5	NREM1
2	18.2	6.2	2.3	1.2	0.3	NREM3
3	10.5	7.8	5.5	3.2	0.4	REM
4	17.9	5.1	2.0	1.1	0.2	NREM3
5	9.7	9.5	6.1	3.5	0.6	REM
6	5.4	11.2	8.3	4.9	1.2	Wake
7	14.1	7.0	4.4	2.8	0.3	NREM2
8	6.2	10.7	7.9	4.1	0.9	Wake
9	11.0	8.1	5.2	3.0	0.4	REM
10	13.3	6.5	3.9	2.2	0.2	NREM2

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

1er dataset: Métodos de aprendizaje no supervisado

2do dataset: Métodos de aprendizaje supervisado (dada la presencia de clases)

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

Principalmente la especie de variable con la que se trabaja. Mientras que en los métodos de clasificación se trabaja con variables cualitativas (clases como: sano o enfermo), en los métodos de regresión se busca encontrar un valor numérico (por ejemplo: si tenemos los datos de 100 corredores que nos dicen x: su tiempo de entrenamiento, y: sus mejores tiempos corriendo 100m. Entonces podemos buscar cuánto tiempo hará un corredor recorriendo 100 m si conocemos cuánto tiempo ha entrenado).

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

Imaginemos que tenemos los datos (X: Kilómetros recorridos, Y: Precio) de 39 carros y nosotros deseamos entrenar a nuestro modelo mediante el método de vecinos más cercanos. Si nosotros usamos una  $K=39$ , nuestro modelo realizará una promediación de los datos y nos será imposible encontrar el precio de un nuevo carro (nuevo dato) ya que el modelo será demasiado simple (lo cual demuestra un under-fitting). Sin embargo, si nosotros utilizamos una  $k=1$ , nuestro modelo será demasiado complejo lo cual será muy bueno para analizar los datos que ya tenemos, pero no los nuevos datos (osea que no es un modelo 'generalizable'). Ambos casos nos impiden realizar una correcta generalización del modelo.

- 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?

Over-fitting casos: 1, 6

Underfitting casos: 2, 3

Balanceado: 4, 5

6) Contesta lo siguiente:

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

$$\text{Accuracy} = \frac{\text{Predicciones Correctas}}{\text{Total de predicciones}}$$

$$\text{Precision} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Positivos}}$$

$$\text{Recall} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos positivos} + \text{Falsos Negativos}}$$

$$\text{f1-score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}}$$

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

$$\text{Accuracy} = 0.85$$

$$\text{Precision} = 0.909$$

$$\text{Recall} = 0.83$$

$$\text{f1-score} = 2.304$$

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.	Precision

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
2	No tiene fase de entrenamiento como tal; toda la predicción ocurre en tiempo real	D
3	Busca el hiperplano que maximiza el margen entre clases	C
4	Realiza múltiples divisiones binarias para clasificar los datos	D
5	Es sensible a la escala de las características, especialmente cuando se usa distancia euclidiana	B
6	Utiliza funciones núcleo (kernel) para trabajar con datos no lineales	D
7	Calcula la probabilidad de pertenencia a una clase bajo modelos estadísticos	B
8	Clasifica observaciones comparándolas con las más cercanas en el espacio de características	B
9	Utiliza múltiples árboles para reducir la varianza y mejorar la precisión	E
10	Puede sobreajustar fácilmente si el árbol es muy profundo	E
11	Su versión más simple se basa en modelos lineales con varianzas iguales por clase	B
12	Puede usar núcleos como RBF o polinomiales para separar clases de manera no lineal	B
13	Tiende a mejorar la generalización comparado con un solo árbol	E
14	Es altamente interpretativo cuando se trabaja con dos clases y pocos atributos	A
15	Es una combinación de muchos árboles construidos sobre subconjuntos aleatorios del conjunto de entrenamiento	E

8. ¿Cuál es la diferencia entre Perceptrón, Red Neuronal Artificial y Red Neuronal Convolutiva?

Perceptrón: Se trata de 'la neurona digital' si la quisiéramos definir de forma coloquial, el cual recibe un conjunto de valores de entrada  $x$ , cada uno con su propio valor  $w$  (peso) que da lugar a una salida  $y$ .

Red Neuronal Artificial: Cuando hablamos de una red neuronal artificial podemos hacer referencia al porqué se llama *Deep Learning*, ya que estas poseen múltiples capas ocultas entre las propias capas de entrada y salida (inspiradas en redes neuronales biológicas). La capa de entrada puede recibir una gran variedad de tipos distintos de datos.

Red Neuronal Convolutiva: Está más relacionada al análisis de imágenes (un ejemplo famoso son aquellas para jugar piedra papel o tijera en base a las fotos

de las manos de un jugador, así identificar si se trata de piedra, papel o tijera en base a una fotografía).

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

Es un procedimiento de perfeccionamiento de los pesos de una red neuronal artificial, mediante el cual se van ajustando de manera iterativa al menor error posible (o tolerable).

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

HAY MUCHÍSIMAS: Escalonadas, lineales, sigmoide, ReLu, arcotangente, etc. Cada una nos da diferentes ventajas dependiendo del conjunto de datos el cual estamos analizando (ej. las escalonadas al tener una salida 1 o 0 nos pueden ayudar para clases binarias).

La función sigmoide nos da lugar a múltiples salidas pero aun así se utiliza para 'clases binarias', lo cual nos puede dar un amplio conjunto de salidas para un amplio conjunto de entradas. Sin embargo *softmax* es para multiclase.

**NO ACABE DE ESTUDIAR TODO CON SUS  
VIDEOS, PERO EL SEÑOR OCTAVIO  
GUTIÉRREZ ES MI HÉROE,**



<https://www.youtube.com/@CodigoMaquina>

AUNQUE REPRUEBE  
SEGUIRE  
APRENDIENDO MÁS  
PORQUE LOS  
TEMAS ESTAN  
CHIDOS (PERO NO  
ME DIO TIEMPO) :D  
PURO RIFARSE SIN  
APUNTES.....NO ME  
REPRUEBE :/

