

# Examen 3er parcial. Examen teórico

Marcela Aguirre valdez

- Nombre del archivo Examen\_3er\_parcial\_nombre
  - En la branch **main**
  - Dentro del folder S04\_parciales/P03\_3er\_parcial
  - Subir el archivo antes de las 7pm del 19 de mayo
1. ¿Cuál es la diferencia entre métodos de aprendizaje supervisado y no supervisado?

Los métodos de aprendizaje se pueden clasificar según el tipo de datos que tenemos y lo que queremos obtener de estos datos. De manera general, al querer clasificar datos según diferentes variables podemos tener datos de los cuáles sabemos sus clasificaciones, como lo son etiquetas, y entonces dándole la información de las etiquetas reales de un dataset original podemos entrenar a los métodos a aprender a clasificar estos datos. Por el contrario, si los datos que tenemos no están ya clasificados, y queremos separarlos de alguna manera a partir de sus características o componentes, entonces podemos hacer uso de métodos de aprendizaje no supervisado.

2. Observa los siguientes dataset dummies:

- DATASET 1

I D	EEG_Δ (0.5–4Hz)	EEG_θ (4–8Hz)	EEG_α (8–12Hz)	EEG_β (12–30Hz)	EEG_γ (30–45Hz)	SleepSta ge
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

1	13.3	6.5	3.9	2.2	0.2	NREM2
0						

#### - DATASET 2

I D	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

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

El primer dataset, por tener ciertas métricas y finalmente una etiqueta (del tipo de ondas que se están registrando) se pueden aplicar métodos de aprendizaje supervisado. El segundo dataset que contiene únicamente métricas o características, puede ser utilizado en aprendizaje no supervisado.

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

El tipo de etiqueta o output final que se desea obtener. Los problemas de regresión tienen que ver con medidas numéricas, generalmente continuas, mientras que la clasificación tiene que ver con categorías o etiquetas discretas, como discriminar entre si ciertos datos pertenecen a un grupo control o a algún grupo experimental según las características extraídas.

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

El over-fitting ocurre cuando entrenamos a nuestro modelo a identificar a nuestro training dataset con muy poca varianza y alta precisión y exactitud, admitiendo menor cantidad de errores en este entrenamiento. Sin embargo este entrenamiento resulta en un modelo poco flexible que tiene alta tasa de error al evaluarse con datos que no pertenecen al training set. El under-fitting por otro lado es un entrenamiento más laxo y flexible, que admite más errores en el entrenamiento pero que puede resultar en mejores resultados al evaluarse el modelo con set de datos. El over fitting es altamente sensible al training set, y tiene muy poca precisión con datos nuevos. Se debe entonces buscar un punto dulce para que tanto comprometemos a nuestro modelo, qué tan estricto es en el entrenamiento para que sea más flexible. Esto podemos evaluarlo con cross-validation u u

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?

Los ensayos 1, 4 y 100 son over-fitting porque a pesar de que el entrenamiento tiene altísima precisión, el modelo falla al evaluarse. Diría que el ensayo 5 es un entrenamiento balanceado. Los 2 y 3 podrían ser un under-fitting, pero podemos ver que resultan balanceados en el sentido de que la precisión para ambas fases son muy similares, sin embargo no significa que sean buenos modelos, en realidad son muy malos.

6. Contesta lo siguiente

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

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN})$$

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

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

$$\text{F1-score} = \text{Media armónica precision y recall} = 2 * (\text{Precision} * \text{recall}) / (\text{Precision} + \text{recall})$$

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

$$\text{Accuracy} = 40+45 / 100 = 0.95$$

$$\text{Precision} = 40 / 40+5 = 0.88$$

$$\text{Recall} = 40 / 50 = 0.8$$

$$\text{F1} = 2*0.8*0.888/(0.8+0.888888) = 0.8421$$

	<b>Predicho Positivo</b>	<b>Predicho Negativo</b>
<b>Real Positivo</b>	TP = 40	FN = 10
<b>Real Negativo</b>	FP = 5	TN = 45

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

<b>Descripción</b>	<b>Métrica</b>
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	B
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	C
7	Calcula la probabilidad de pertenencia a una clase bajo modelos estadísticos	C
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	D
11	Su versión más simple se basa en modelos lineales con varianzas iguales por clase	A
12	Puede usar núcleos como RBF o polinomiales para separar clases de manera no lineal	C
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	D
15	Es una combinación de muchos árboles contruidos sobre subconjuntos aleatorios del conjunto de entrenamiento	E

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

El perceptrón es un modelo idealizado de una neurona (biológica) que se usa en modelos de redes neuronales para manejar información. Recibe entradas y mediante funciones de activación transmite información transformada (output). Estos perceptrones pueden conectarse entre ellos para procesar información :p Una red neuronal artificial es el resultado de estas conexiones, puede construirse bajo diferentes parámetros como número de neuronas, número y tipo de capas de neuronas, funciones de activación, etc. Una red neuronal convolutacional es específicamente una red neuronal artificial que procesa información de tipo imagen (matrices de píxeles), para extraer características y finalmente (generalmente) clasificar el tipo de imagen que se recibe. Se caracterizan por el uso de matrices convolucionales (más pequeñas que la imagen) y que buscan desplazándose por toda la imagen para transformarla y extraer información de dónde en la imagen aparece el atributo que corresponde al de la matriz convolutacional usada.

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

Una época es un ciclo completo de entrenamiento a partir de un training set y test set original. Generalmente en modelos de deep learning se configuran los modelos para hacer múltiples épocas donde se permutan los grupos de entrenamiento y de prueba para darle mayor robustez a los modelos y mejorar su eficiencia.

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

escalón, sigmoide, ReLu → rectificador lineal unitario, atan, softmax

La función sigmoide (en forma de S) se utiliza principalmente para clasificaciones(? de probabilidad acotada entre 0 y 1 => para valores muy negativos el valor tiende asintóticamente a 0, y para valores muy positivos tiende a 1.

La función softmax por otro lado es de clasificación de etiquetas. Se encarga de dar un output de probabilidades, donde la suma da 1 => es realmente útil en problemas de clasificación donde evaluamos cuál es la probabilidad de una entrada de pertenecer a cualesquiera de los grupos/etiquetas de salida => mediante otra neurona especificamos cómo asignamos la etiqueta, generalmente se asigna la etiqueta de probabilidad más alta :p