

# Examen 3° Parcial - Aitana Orduña Lugo

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

- La principal diferencia son las **etiquetas**, por ejemplo en los **métodos de aprendizaje supervisado** los datos con el que estamos entrenando nuestro modelo ya tiene una etiqueta o grupo al que pertenecen y con esos datos lo que intenta hacer el modelo es hacer predicciones. En cambio con los **métodos no supervisados** el modelo no tiene parte de una base de respuestas ya etiquetadas y de ahí aprende, si no que más bien forma maneras de agrupar eficazmente un conjunto de datos de tal manera que queden clasificados de la mejor manera posible.

## 2. Observa los siguientes dataset dummies, ¿en cuál dataset se pueden aplicar métodos de Aprendizaje supervisado y en cuál métodos de Aprendizaje no supervisado?

### a. Dataset 1 → Aprendizaje supervisado

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

- Porque en este dataset sí existe una columna en la cuál se están clasificando de alguna manera los datos, es decir, tienen etiquetas. En este caso las etiquetas están en la columna de "SleepStages" y se dividen en NREM1, NREM2, NREM3, REM y Wake, de esta manera el modelo que entrenemos puede ayudarse de esas clasificaciones previas y así hacer predicciones.

b. Dataset 2 → Aprendizaje no supervisado

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

- En cambio en este dataset solo hay datos ahí aleatorios sin una columna con etiquetas o alguna manera ya existente de agruparlos, es por esto que sería mejor idea optar por alguno de los métodos de aprendizaje no supervisado como k-means o PCA que puedan hacer que nuestro modelo haga agrupaciones lo más eficaces posibles.

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

- La diferencia está en que para la **clasificación** el modelo aprende a asignar una categoría o clase, mientras que en la **regresión** el modelo aprende a predecir un valor numérico.

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

- La principal diferencia radica en el margen de error que tiene el modelo, esto puede deberse en parte a la cantidad de datos que se usan para training y testing.

- Por ejemplo, un modelo **sobre-entrenado (over-fitting)**, va a tender a tener márgenes de error más pequeños porque aprendió de más por así decirlo, entonces se va a intentar ajustar hasta a los outliers, provocando que ya no funcione bien con datos nuevos porque no tiene la capacidad de generalizar con respecto a los datos que ya tenía, es muy específico para eso.
- En cambio en un modelo **poco entrenado (under-fitting)**, el modelo básicamente se queda corto, no se le dieron los suficientes datos de entrenamiento para lograr entender bien los patrones de los datos y las predicciones o clasificaciones que hace con respecto a los nuevos datos están super equivocadas.

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: Over-fitting
    - Porque el accuracy del set de entrenamiento es casi perfecto pero al momento de usar los datos de prueba se desploma ese porcentaje, eso nos indica que nuestro modelo es muy estricto y específico.
  - Ensayo 2: Under-fitting

- Porque tiene un accuracy malo en ambos sets, tanto en prueba como en entrenamiento, el problema es que si desde un inicio mi modelo no aprendió bien (entrenamiento) lo va a hacer peor en datos que no son con los que se entrenó y eso indica un mal desempeño de nuestro modelo.
- Ensayo 3: Under-fitting
  - Mismo caso que el ensayo 2.
- Ensayo 4: Over-fitting
  - Mismo caso que el ensayo 1.
- Ensayo 5: Balanceado !!! 🤔
  - Porque los porcentajes son bastante buenos en ambos sets de datos, eso nos indica que el modelo que entrenamos funciona correctamente o al menos de buena manera con datos ajenos a aquellos con el que se entrenó.
- Ensayo 6: Over-fitting
  - Mismo caso que el ensayo 1.

## 6. Contesta lo siguiente:

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

- Accuracy
  - $TP + TN / TP + TN + FP + FN$
- Precision
  - $TP / TP + FP$
- Recall
  - $TP / TP + FN$
- F1-Score
  - $2 \times ((Precision \times Recall) / (Precision + Recall))$
- Donde:
  - **TP** = Verdaderos positivos

- **TN** = Verdaderos negativos
- **FP** = Falsos positivos
- **FN** = Falsos negativos

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
  - $85 / 100 = 0.85 \rightarrow 85\%$
- Precision
  - $40 / 45 = 0.88 \rightarrow 88\%$
- Recall
  - $40 / 50 = 0.8 \rightarrow 80\%$
- F1-Score
  - $2 \times (0.704/1.68) = 0.83 \rightarrow 83\%$

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:**

#	Características	¿A qué método pertenece?
1	Asume que los datos de cada clase siguen una distribución normal (gaussiana)	1
2	No tiene fase de entrenamiento como tal; toda la predicción ocurre en tiempo real	2
3	Busca el hiperplano que maximiza el margen entre clases	3
4	Realiza múltiples divisiones binarias para clasificar los datos	4
5	Es sensible a la escala de las características, especialmente cuando se usa distancia euclidiana	2
6	Utiliza funciones núcleo (kernel) para trabajar con datos no lineales	3
7	Calcula la probabilidad de pertenencia a una clase bajo modelos estadísticos	1
8	Clasifica observaciones comparándolas con las más cercanas en el espacio de características	2
9	Utiliza múltiples árboles para reducir la varianza y mejorar la precisión	5
10	Puede sobreajustar fácilmente si el árbol es muy profundo	4
11	Su versión más simple se basa en modelos lineales con varianzas iguales por clase	1
12	Puede usar núcleos como RBF o polinomiales para separar clases de manera no lineal	3
13	Tiende a mejorar la generalización comparado con un solo árbol	5
14	Es altamente interpretativo cuando se trabaja con dos clases y pocos atributos	4

#	Características	¿A qué método pertenece?
15	Es una combinación de muchos árboles contruidos sobre subconjuntos aleatorios del conjunto de entrenamiento	5

1. Análisis Discriminante
2. Vecino más cercano (k-NN)
3. Máquina de soporte vectorial (SVM)
4. Árbol de decisión
5. Bosques aleatorios

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

- **Perceptrón:** es el equivalente a lo que vendría siendo 1 neurona, en redes neuronales es la unidad más básica existente.
- **Red Neuronal Artificial:** es un conjunto de perceptrones (el equivalente a una red neuronal anatómica considerando que los perceptrones son neuronas), están organizadas en capas de entrada (que son como los inputs sensoriales que recibe una red anatómica) y las capas de ocultas (que ahora sí tienen la función como de las neuronas). Trabaja con datos distintos a las convolucionales.
- **Red Neuronal Convolucional:** es parecida a la RNA pero en realidad se usa más en imágenes o datos que existen en un espacio. Por ejemplo de las imágenes aprende a detectar partes o pequeños componentes, y de ahí aprende a hacer aproximaciones o clasificaciones para estas, a partir de todo lo que ya aprendió de imágenes previas. Son un poco más complejas que las artificiales.

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

- Es como una inspección completa al conjunto de datos con el que estamos entrenando a nuestra red, y con cada vuelta o repaso el modelo se vuelve mejor y mejor en conocer los datos.

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

- **Principales funciones de activación:**

- ReLu, escalón, sigmoide, softmax, tanh

- **Diferencias entre sigmoide y softmax**

- Se utiliza la sigmoide cuando la clasificación solo consiste de dos diferentes clases, esto porque lo que hace esta función es convertir los valores en 1 y 0 para calcular la probabilidad.
- En cambio, softmax funciona de manera similar con respecto a las probabilidades pero esta sí permite que existan más de dos clases para clasificar, lo que hace es que divide la probabilidad de tal manera que al final todas las clases sumen 1 y obvio asigna los valores dependiendo de que tan probable sea que un dato pertenezca a una clase.