



Examen_3er_parcial_clau

Parcial III

*Claudia Coutiño
Vásquez*

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

Esencialmente, tiene que ver con el hecho de si los datos están etiquetados o no. En el supervisado, los modelos aprenden a partir de ejemplos etiquetados (aquí, los resultados finales se conocen) y buscan predecir estos valores (etiqueta). Por otro lado, el no supervisado el objetivo en si es buscar patrones ocultos en los datos (sin etiquetas) explorando su estructura; entonces puede extraer agrupaciones, componentes.

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

- DATASET 1

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

- DATASET 2

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

El primer dataset está bonito, son registros encefalográficos para diferentes frecuencias cerebrales cuando los sujetos están durmiendo, parece que se clasifican entonces en las etapas (NREM, REM -como dato, porque el tema me apasiona en neuro→ en la etapa NREM3 está especialmente activo el sistema glinfático donde el líquido cefalorraquídeo va limpiando proteínas mal plegadas, desechos metabólicos, etc. para mantener la homeostasis o salud del cerebro, pero sus alteraciones podrían estar relacionadas con enfermedades neurodegenerativas, lo puede investigar, está muy chido-). Como se cuenta ya con las etiquetas, se podrían aplicar métodos de aprendizaje supervisado.

El segundo dataset sólo aparecen componentes y valores pero no hay una clasificación dada, entonces se podría aplicar métodos no supervisados para descubrir patrones interesantes.

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

Creo, se refiere a que las regresiones buscan predecir valores continuos, mientras que la clasificación trata de predecir las etiquetas de un conjunto de datos.

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

El overfitting ocurre cuando se quiere abarcar todo, el modelo aprende demasiado bien los datos durante el entrenamiento igual que el ruido (visto gráficamente sería como si se pusieran puntos en cada parte de una curva dada). El problema es que se vuelve demasiado "complejo" que no será bueno en la prueba, muchos parámetros darán mucha varianza.

En contraste con el underfitting, que es como si le faltara aprender al modelo, porque es muy simple (o los parámetros no son los más adecuados) o falta entrenamiento. Esto produce un alto sesgo (gráficamente como si tuvieras muchos puntos dispersos en un plano y trazarás una línea recta buscando encajar pero en realidad el patrón sigue una curva)

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?

A mi parecer, el 5 está balanceado (los valores son muy cercanos en el entrenamiento y la prueba). Underfitting en el 2 y 3 (tienen un accuracy muy bajo), y overfitting 1, 4, 6 (es más evidente porque el entrenamiento lo hace súper bien, pero el accuracy disminuye mucho en la prueba).

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

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: $\frac{TP+TN}{TP+TN+FP+FN} = 85/100 = 0.85$

Precisión: $\frac{TP}{TP+FP} = 40/45 = 0.8888$

Recall: $\frac{TP}{TP+FN} = 40/50 = 0.8$

F1-Score: $2 \cdot \frac{\text{Precisión} \cdot \text{Recall}}{\text{Precisión} + \text{Recall}} = 2(0.71104/1.6888) = 0.8420$

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.	
Porcentaje de verdaderos positivos identificados correctamente, respecto al total de positivos reales.	
Porcentaje de aciertos entre los casos que el modelo predijo como positivos.	

1. Accuracy
2. Recall
3. Precisión

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)	
2	No tiene fase de entrenamiento como tal; toda la predicción ocurre en tiempo real	
3	Busca el hiperplano que maximiza el margen entre clases	
4	Realiza múltiples divisiones binarias para clasificar los datos	
5	Es sensible a la escala de las características, especialmente cuando se usa distancia euclidiana	
6	Utiliza funciones núcleo (kernel) para trabajar con datos no lineales	
7	Calcula la probabilidad de pertenencia a una clase bajo modelos estadísticos	
8	Clasifica observaciones comparándolas con las más cercanas en el espacio de características	
9	Utiliza múltiples árboles para reducir la varianza y mejorar la precisión	

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

- 1. A
- 2. B
- 3. C
- 4. D
- 5. A
- 6. C
- 7. A
- 8. B
- 9. E
- 10. D
- 11. A
- 12. C
- 13. E
- 14. D
- 15. E

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

- Perceptrón: Es la unidad básica de la red neuronal
- Red neuronal artificial: Es un modelo de aprendizaje supervisado que se compone de nodos organizados en capas
 - Entrada → Hidden layers → salida
- Red neuronal convolutacional: Es un tipo de red neuronal que procesa datos en formato "rejilla" (utilizan datos tridimensionales), se suelen ocupar para el reconocimiento de imágenes

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

Tiene que ver con la organización de los datos de aprendizaje → Se pueden generar una cantidad de batchs que se van iterando entre sí, el conjunto completo sería el epoch. El entrenamiento busca encontrar los parámetros del modelo.

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

Las funciones de activación son relevantes en las redes neuronales, se aplican después de la suma ponderada de las entradas (después de multiplicar por el peso y añadir el bias), son para determinar si la señal se va a transmitir a la siguiente capa.

La función sigmoid se utiliza para la clasificación binaria, mientras que el softmax cuando es multiclase.

Otras funciones son Tanh, RELU (para las hidden layers), y Lineal si se trata de un problema de regresión lineal continua.
