

Examen 3 Rafael

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

En el aprendizaje supervisado el modelo ya sabe quien es quien. A qué categoría pertenece cada uno. Algunos son regresión lineal y clasificación con árboles de decisión. En el aprendizaje no supervisado es lo opuesto. Usualmente se buscan patrones o estructuras que subyacen en la información. Algunos ejemplos son K-means o un PCA.

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

El dataset 1 es para aprendizaje supervisado porque incluye una etiqueta que clasifica la información (Sleep Stage). El dataset 2 permite es para aprendizaje no supervisado debido a que la información aún no está clasificada en grupos.

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

Uno de clasificación de grupos, y uno de regresión es para sistemas numéricos (respuesta uno a uno).

4. ¿Cuál es la diferencia entre overfitting y underfitting

Overfitting es cuando los datos están sobre catalogados en demasiadas categorías que no hay grupos concretos porque no se dejó de podar el árbol o el modelo. Underfitting es lo opuesto, son pocas categorizaciones debido a que no se pudo el árbol antes.

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

1 Overfitting

2 Underfitting

3 Underfitting

4 Overfitting

5 Balanceado

6 Overfitting

6. ¿Qué es un modelo de regresión lineal y para qué sirve?

Accuracy: 0.85

Precision: 0.8889

Recall: 0.8

F1-Score: 0.8421

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:

1. Asume que los datos de cada clase siguen una distribución normal (gaussiana): Análisis Discriminante
2. No tiene fase de entrenamiento como tal; toda la predicción ocurre en tiempo real: Vecino más cercano (k-NN)
3. Busca el hiperplano que maximiza el margen entre clases: Máquina de soporte vectorial (SVM)
4. Realiza múltiples divisiones binarias para clasificar los datos: Árbol de decisión
5. Es sensible a la escala de las características, especialmente cuando se usa distancia euclidiana: Vecino más cercano (k-NN)
6. Utiliza funciones núcleo (kernel) para trabajar con datos no lineales: Máquina de soporte vectorial (SVM)

7. Calcula la probabilidad de pertenencia a una clase bajo modelos estadísticos: Análisis Discriminante
8. Clasifica observaciones comparándolas con las más cercanas en el espacio de características: Vecino más cercano (k-NN)
9. Utiliza múltiples árboles para reducir la varianza y mejorar la precisión: Bosques aleatorios
10. Puede sobreajustar fácilmente si el árbol es muy profundo: Árbol de decisión
11. Su versión más simple se basa en modelos lineales con varianzas iguales por clase: Análisis Discriminante
12. Puede usar núcleos como RBF o polinomiales para separar clases de manera no lineal: Máquina de soporte vectorial (SVM)
13. Tiende a mejorar la generalización comparado con un solo árbol: Bosques aleatorios
14. Es altamente interpretativo cuando se trabaja con dos clases y pocos atributos: Árbol de decisión
15. Es una combinación de muchos árboles construidos sobre subconjuntos aleatorios del conjunto de entrenamiento: Bosques aleatorios

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

El perceptrón es el primer modelo hecho de una red neuronal, es más simple de red neuronal porque es de una sola capa y se usa solo para clasificar datos linealmente. La red neuronal artificial tiene múltiples capas ocultas, y es no lineal por las ecuaciones de función. La red neuronal convolutacional es para analizar imágenes y usa capas de convolución para detectar patrones.

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

Es cada vez que el algoritmo hace una run por la red neuronal

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

Las principales funciones de activación en redes neuronales son ReLU, sigmoid, tanh y softmax. ReLU es la más utilizada por su eficiencia y evita el problema del gradiente desvanecido. La función sigmoid convierte cualquier valor en un rango entre 0 y 1, y se usa en tareas de clasificación binaria. Softmax también transforma valores en un rango de 0 a 1 y regresa una distribución de probabilidad que es multiclase. La diferencia entre sigmoid y softmax es que sigmoid usa una sola salida, mientras que softmax se aplica a vectores y permite seleccionar entre múltiples clases de la red neuronal.