PCBDAT30CV - Diplomado en Big Data - Data Science

Módulo 5. Machine Learning

Mód 5 - PAC Prueba Autoformativa Corta

Presentaciones

Aquí están tus últimas respuestas:

Pregunta 1

Una entidad de gobierno quiere crear un modelo que prediga la calificación de la prueba ICFES para todos los estudiantes del gobierno 2018-2022. Como variables predictoras se usan variables demográficas, socioeconómicas e información del colegio donde el estudiante cursó el grado undécimo. La variable objetivo es la calificación del examen que toma valores entre 0 y 500.

Responda, ¿qué tipo de problema se quiere resolver?

Respuesta: c. Regresión

Puntuación: 10 de cada 10 Sí

Pregunta 2

Una entidad de gobierno quiere crear un modelo que prediga el desempeño de la prueba ICFES para todos los estudiantes del gobierno 2018-2022. Como variables predictoras se usan variables demográficas, socioeconómicas e información del colegio donde el estudiante cursó el grado undécimo. La variable objetivo fue discretizada y toma los siguientes valores de desempeño: Muy bajo, bajo, moderado, alto, muy alto. Responda, ¿qué tipo de problema se quiere resolver?

Respuesta: a. Clasificación **Puntuación**: 10 de cada 10 Sí

Pregunta 3

Respecto a métricas de clasificación, ¿cuál permite monitorear los falsos positivos en relación con los verdaderos positivos?

Respuesta: a. Exactitud

Puntuación: 0 de cada 10 No

Pregunta 4

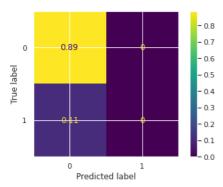
La exactitud siempre es una buena métrica para evaluar el desempeño de un modelo de clasificación.

Respuesta: Falso

Puntuación: 10 de cada 10 Sí

Pregunta 5

Usted busca crear un modelo que prediga si un paciente tiene (1, positivo) o no tiene (0, negativo) una enfermedad terminal. Usted encontró que su modelo de clasificación tiene la siguiente matriz de confusión. Escoja todas las afirmaciones correctas.



Respuesta: Los falsos negativos son más que los falsos positivos

Respuesta: Hay 0 verdaderos positivos

Puntuación: 6.67 de cada 10

Pregunta 6

Usted trabaja para una firma y en las últimas semanas ha creado 3 modelos de regresión empleando 3 algoritmos:

- Regresión lineal
- Perceptrón multicapa

· Bosque aleatorio

Usted desea comparar el desempeño de estos 3 modelos usando el conjunto de prueba. ¿Qué métrica NO debería escoger?

Respuesta: b. Coeficiente de determinación (R2)

Puntuación: 10 de cada 10 Sí

Pregunta 7

Respecto a métricas de evaluación para modelos de regresión, seleccione todas las que sean correctas.

Respuesta: c. El Error Cuadrático Medio MSE es una métrica que penaliza más fuertemente los errores grandes.

Respuesta: d. El R2 mide la fracción de varianza de "y" que es explicada por las variables predictoras "X".

Respuesta: f. El Error Absoluto Medio MAE es una métrica que penaliza de igual forma los errores grandes y pequeños.

Puntuación: 10 de cada 10 Sí

Pregunta 8

Respecto al aprendizaje no supervisado y el algoritmo de K-Means. ¿Cuál es la medida de distancia más popular?

Respuesta: b. Distancia Euclidiana

Puntuación: 10 de cada 10 Sí

Pregunta 9

Una entidad de gobierno quiere crear un modelo que prediga la calificación de la prueba ICFES para todos los estudiantes del gobierno 2018-2022. Como variables predictoras se usan variables demográficas, socioeconómicas e información del colegio donde el estudiante cursó el grado undécimo. La variable objetivo es la calificación del examen que toma valores entre 0 y 500.

Su jefe le sugiere que, al tratarse de un problema de regresión, usted utilice regresión lineal y regresión logística para crear los modelos. De acuerdo a lo anterior su respuesta sería.

Respuesta: d. Está en desacuerdo con su jefe, el problema a resolver sí es de regresión, pero la regresión logística no es un algoritmo apropiado para resolver este problema.

Puntuación: 10 de cada 10 Sí

Pregunta 10

Son métricas para evaluar el desempeño de un agrupamiento.

Respuesta: c. Inercia

Puntuación: 5 de cada 10

Bogotá Cra 18B No 106 A 15 +57 (601) 489 4783 +57 (601) 489 4786

Email: diroperaciones@udecatalunya.edu.co dir.academica@udecatalunya.edu.co



