



Comenzado el	viernes, 10 de mayo de 2024, 09:12
Estado	Finalizado
Finalizado en	viernes, 10 de mayo de 2024, 09:45
Tiempo empleado	32 minutos 33 segundos
Calificación	24,00 de 30,00 (80%)

Pregunta 1

Correcta  
Se puntúa 1,00 sobre 1,00

¿Qué es el machine learning?

- ☐ a. Es el estudio de cómo las computadoras pueden interactuar con los humanos a través de interfaces naturales como el habla y la escritura.
- ☐ b. Es la práctica de utilizar algoritmos para analizar y clasificar grandes cantidades de datos en una base de datos.
- ☒ c. Es el proceso mediante el cual una computadora puede mejorar su rendimiento en una tarea específica a través de la experiencia previa sin ser explícitamente programada para esa tarea.
- ☐ d. Es una rama de la ciencia de la computación que se enfoca en la creación de sistemas operativos para computadoras y dispositivos móviles.

Pregunta 2

Incorrecta  
Se puntúa 0,00 sobre 1,00

¿Cuáles son los pasos típicos que se deben realizar para generar un modelo de regresión lineal?

- ☐ a. 1 - Realizar una validación cruzada de k-fold. 2 - Crear un conjunto de datos sintético. 3 - Aplicar técnicas de aumento de datos. 4 - Ajustar los hiperparámetros del modelo.
- ☐ b. 1 - Dividir los datos en un conjunto de entrenamiento y un conjunto de prueba. 2 - Seleccionar las características del modelo. 3 - Ajustar los coeficientes del modelo. 4 - Validar el modelo.
- ☐ c. 1 - Realizar un análisis exploratorio de datos. 2 - Implementar una red neuronal profunda. 3 - Evaluar el modelo mediante una matriz de confusión. 4 - Efectuar una reducción de dimensionalidad.
- ☒ d. 1 - Elegir un algoritmo de aprendizaje no supervisado. 2 - Aplicar una técnica de clustering a los datos. 3 - Realizar una optimización de hiperparámetros. 4 - Realizar una prueba de normalidad en los residuos del modelo.

Se puntúa 0,00 sobre 1,00

¿Cuál es el propósito del código proporcionado?

```
import pandas as pd  
housing.info()
```

- ☒ a. Importar la biblioteca pandas y mostrar las primeras filas del dataframe 'housing'. ✖
- ☐ b. Crear un nuevo dataframe 'housing' y mostrar su estructura.
- ☐ c. Importar la biblioteca pandas y mostrar información sobre el dataframe 'housing', como el número de filas y columnas, tipos de datos, y uso de memoria.
- ☐ d. Realizar una operación de agregación en el dataframe 'housing'.

Respuesta incorrecta.

**Pregunta 4**

Incorrecta

Se puntúa 0,00 sobre 1,00

¿Qué sucede si en un problema de Machine Learning los datos de entrenamiento no son representativos?

- ☒ a. Los datos no representativos no afectan la capacidad del modelo para generalizar. ✖
- ☐ b. El algoritmo de Machine Learning se ajusta automáticamente a los datos no representativos y no hay ningún impacto en el rendimiento del modelo.
- ☐ c. El modelo generaliza mejor y hace predicciones más precisas en datos reales.
- ☐ d. El modelo puede aprender patrones erróneos o sesgados y tener un rendimiento deficiente en datos reales.

**Pregunta 5**

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la regresión lineal y qué tipo de problemas nos permite resolver?

- ☐ a. La regresión lineal es un método para calcular la distancia entre dos puntos en un espacio tridimensional.
- ☐ b. La regresión lineal es un método estadístico que se utiliza para resolver problemas de clasificación en el aprendizaje automático.
- ☐ c. La regresión lineal es un algoritmo de optimización utilizado para resolver problemas de procesamiento de imágenes.
- ☒ d. La regresión lineal es un método matemático que modela la relación lineal entre una variable dependiente y una o más variables independientes, y permite resolver problemas de predicción y estimación. ✔



Se puntúa 1,00 sobre 1,00

¿Cuál es una característica fundamental del Aprendizaje No Supervisado en Machine Learning?

- ☐ a. En el aprendizaje no supervisado, los modelos se entrenan para predecir salidas específicas a partir de entradas dadas.
- ☐ b. En el aprendizaje no supervisado, los modelos se enfocan exclusivamente en problemas de clasificación.
- ☐ c. En el aprendizaje no supervisado, los modelos se entrenan con un conjunto de datos que contiene ejemplos etiquetados.
- ☒ d. En el aprendizaje no supervisado, los modelos exploran patrones y estructuras en datos sin la guía de etiquetas o resultados conocidos. ✓

#### Pregunta 7

Correcta

Se puntúa 1,00 sobre 1,00

Si al resolver un problema de predicción de una variable numérica se generan cuatro modelos diferentes de regresión lineal, ¿cómo seleccionamos el modelo más adecuado?

- ☐ a. Escojo el modelo que se ajusta perfectamente a los datos de entrenamiento sin importar la métrica de evaluación.
- ☐ b. Selecciono el modelo que tiene los coeficientes más pequeños, ya que son menos propensos al sobreajuste.
- ☒ c. Utilizo una métrica de evaluación, como el error cuadrático medio (MSE) o el coeficiente de determinación ( $R^2$ ), en el conjunto de prueba para comparar y seleccionar el modelo que tenga el mejor rendimiento. ✓
- ☐ d. Elijo el modelo que tiene el mayor número de características (variables independientes) para hacer predicciones más precisas.

Se puntúa 0,00 sobre 1,00

Se desea calcular la métrica MSE para un modelo de regresión lineal RLS que se acaba de entrenar. Seleccione el código correcto que permite realizar esta tarea:

- ☐ a. `y_pred = RLS.predict(X_test)`
- ```
from sklearn.metrics import mean_squared_error
MSE_M1 = mean_squared_error(y_test, y_pred)
print("Mean Squared Error (MSE) M1:", MSE_M1)
```
- ☒ b. `y_pred = RLS.predict(Y_test)` ✗
- ```
from sklearn.metrics import MSE
MSE_M1 = MSE(y_test, y_pred)
print("Mean Squared Error (MSE) M1:", MSE_M1)
```
- ☐ c. `y_pred = RLS.predict(X_train)`
- ```
from sklearn.metrics import mean_squared_error
MSE_M1 = mean_squared_error(y_train, y_pred)
print("Mean Squared Error (MSE) M1:", MSE_M1)
```
- ☐ d. `y_pred = RLS.predict(y_train)`
- ```
from sklearn.metrics import MSE
MSE_M1 = mean_squared_error(y_train, y_pred)
print("Mean Squared Error (MSE) M1:", MSE_M1)
```

Respuesta incorrecta.

**Pregunta 9**

Correcta

Se puntúa 1,00 sobre 1,00

¿Por qué es necesario disponer de una gran cantidad de datos para entrenar un algoritmo de Machine Learning de manera efectiva?

- ☒ a. Para ayudar al algoritmo a generalizar patrones y relaciones en los datos, lo que mejora su capacidad de hacer predicciones precisas en datos no vistos. ✓
- ☐ b. Para que el algoritmo sea más rápido y eficiente en sus predicciones.
- ☐ c. Para reducir la complejidad del modelo y hacerlo más sencillo de comprender.
- ☐ d. Para evitar el sobreajuste del modelo y reducir su capacidad de adaptación a nuevos datos.



Se puntúa 1,00 sobre 1,00

¿Cuál de los siguientes algoritmos es comúnmente utilizado en tareas de aprendizaje supervisado para la clasificación?

- ☐ a. Algoritmo Apriori.
- ☐ b. Regresión lineal.
- ☒ c. Regresión Logística ✓
- ☐ d. K-Means.

#### Pregunta 11

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de los siguientes describe o caracteriza mejor lo que es el Aprendizaje Supervisado en Machine Learning?

- ☒ a. En el aprendizaje supervisado, los modelos se entrenan utilizando un conjunto de datos que contiene ejemplos de entrada y las salidas deseadas correspondientes. ✓
- ☐ b. En el aprendizaje supervisado, los modelos se entrenan para realizar tareas sin supervisión de un humano.
- ☐ c. En el aprendizaje supervisado, los modelos se entrenan sin la necesidad de un conjunto de datos etiquetado.
- ☐ d. En el aprendizaje supervisado, los modelos se entrenan únicamente en datos de texto y lenguaje natural.

#### Pregunta 12

Correcta

Se puntúa 1,00 sobre 1,00

Indique qué es lo que realiza el siguiente código:

```
import pandas as pd  
housing = pd.read_csv('housing.csv')
```

- ☐ a. El código importa la biblioteca `pandas` y muestra el contenido de un archivo CSV llamado "housing.csv".
- ☒ b. El código importa la biblioteca `pandas` y lee un archivo CSV llamado "housing.csv" en un DataFrame. ✓
- ☐ c. El código importa la biblioteca `pandas` y crea un DataFrame vacío.
- ☐ d. El código importa la biblioteca `pandas` y guarda un DataFrame en un archivo CSV llamado "housing.csv".

Respuesta correcta



Se puntúa 1,00 sobre 1,00

Usted tiene tres modelos de regresión lineal con los siguientes valores de MSE: 40, 25 y 100. ¿Cuál de los tres modelos es mejor?

- ☐ a. El modelo con un MSE de 100.
- ☐ b. El modelo con un MSE de 40.
- ☒ c. El modelo con un MSE de 25. ✓
- ☐ d. No se puede determinar cuál de los modelos es mejor solo con la información proporcionada.

#### Pregunta 14

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué define mejor al aprendizaje semi supervisado en Machine Learning?

- ☒ a. En el aprendizaje semi supervisado, los modelos se entrenan utilizando un conjunto de datos que contiene tanto ejemplos etiquetados como no etiquetados. ✓
- ☐ b. En el aprendizaje semi supervisado, los modelos se centran en tareas de regresión y predicción numérica.
- ☐ c. En el aprendizaje semi supervisado, los modelos se entrenan exclusivamente con un pequeño conjunto de datos etiquetados.
- ☐ d. En el aprendizaje semi supervisado, los modelos se entrenan sin ningún tipo de supervisión o guía.

#### Pregunta 15

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuáles son los principales tipos de Machine Learning?

- ☐ a. Aprendizaje Automático, Aprendizaje de Máquinas y Aprendizaje Cognitivo.
- ☒ b. Aprendizaje Supervisado, Aprendizaje No Supervisado y Aprendizaje Profundo. ✓
- ☐ c. Aprendizaje de Reforzamiento, Aprendizaje de Regresión y Aprendizaje de Clasificación.
- ☐ d. Aprendizaje Lineal, Aprendizaje No Lineal y Aprendizaje por Refuerzo.



Se puntúa 1,00 sobre 1,00

¿Qué sucede si en un problema de Machine Learning los datos de entrenamiento contienen características o variables irrelevantes?

- ☐ a. Las características irrelevantes no afectan el rendimiento del modelo de manera significativa.
- ☐ b. El modelo se vuelve más robusto y generaliza mejor en una variedad de problemas.
- ☒ c. El modelo puede verse afectado negativamente, volviéndose menos preciso y eficiente. ✓
- ☐ d. El modelo de Machine Learning se entrena más rápido debido a la menor cantidad de características a considerar.

#### Pregunta 17

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de los siguientes algoritmos es comúnmente utilizado en tareas de aprendizaje no supervisado, especialmente en la agrupación (clustering) de datos?

- ☐ a. Naive Bayes.
- ☐ b. Regresión logística.
- ☐ c. Support Vector Machine (SVM).
- ☒ d. K-Means. ✓

#### Pregunta 18

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de las siguientes afirmaciones describe mejor el error de generalización en Machine Learning?

- ☒ a. El error de generalización se refiere a la discrepancia entre el rendimiento del modelo en los datos de entrenamiento y su capacidad para hacer predicciones precisas en datos nuevos o no vistos. ✓
- ☐ b. El error de generalización se refiere a la diferencia entre la precisión de un modelo en la clasificación de datos y su capacidad para regresar valores numéricos precisos.
- ☐ c. El error de generalización es el error cometido al comparar el desempeño de diferentes modelos en datos de prueba.
- ☐ d. El error de generalización es la diferencia entre el rendimiento del modelo en los datos de entrenamiento y su capacidad para ajustarse perfectamente a esos datos.



Se puntúa 1,00 sobre 1,00

¿Qué describe mejor el fenómeno de underfitting en machine learning?

- ☐ a. Cuando un modelo es demasiado complejo, lo que resulta en un alto rendimiento en los datos de entrenamiento, pero una baja precisión en datos nuevos o no vistos.
- ☒ b. Cuando un modelo es demasiado simple, lo que impide capturar la complejidad subyacente o las tendencias de los datos, resultando en un bajo rendimiento tanto en los datos de entrenamiento como en los de prueba. ✓
- ☐ c. Cuando un modelo logra un equilibrio perfecto entre sesgo y varianza, generalmente proporcionando un rendimiento óptimo en ambos, los datos de entrenamiento y de prueba.
- ☐ d. Cuando un modelo se entrena con una cantidad insuficiente de datos, lo que lleva a conclusiones incorrectas sobre las tendencias de los datos debido a la falta de variedad en el conjunto de entrenamiento.

**Pregunta 20**

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué caracteriza mejor al aprendizaje reforzado en Machine Learning?

- ☐ a. El aprendizaje reforzado se basa únicamente en la supervisión de un experto humano para entrenar modelos.
- ☒ b. El aprendizaje reforzado implica que un agente de toma decisiones interactúa con un entorno para maximizar una recompensa acumulativa a lo largo del tiempo. ✓
- ☐ c. El aprendizaje reforzado se utiliza principalmente para clasificar datos en categorías predefinidas.
- ☐ d. El aprendizaje reforzado se enfoca en la identificación de patrones y relaciones en grandes conjuntos de datos no etiquetados.

**Pregunta 21**

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es el propósito del código proporcionado?

```
from sklearn.model_selection import train_test_split
train_set, test_set = train_test_split(housing, test_size=0.2, random_state=42)
```

- ☐ a. Importar la biblioteca pandas y realizar una operación de filtrado en un DataFrame.
- ☐ b. Importar la biblioteca pandas y dividir el dataframe 'housing' en conjuntos de entrenamiento y p
- ☐ c. Importar la biblioteca pandas y dividir el dataframe 'housing' en grupos basados en una columna específica.
- ☒ d. Importar la biblioteca scikit-learn y dividir el dataframe 'housing' en conjuntos de entrenamiento y prueba para tareas de aprendizaje automático. ✓

Respuesta correcta





Se puntúa 0,00 sobre 1,00

¿Cuál es el propósito del código proporcionado?

```
import pandas as pd  
housing.describe()
```

- ☐ a. Importar la biblioteca pandas y describir la estructura del dataframe 'housing', incluyendo estadísticas descriptivas como la media, la desviación estándar y los percentiles de las columnas numéricas.
- ☐ b. Realizar una operación de filtrado en un DataFrame.
- ☒ c. Importar la biblioteca pandas y mostrar las primeras filas del dataframe 'housing'. ✖
- ☐ d. Crear un nuevo dataframe 'housing' y mostrar su estructura.

Respuesta incorrecta.

### Pregunta 23

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué sucede si en un problema de Machine Learning los datos de entrenamiento son de baja calidad?

- ☐ a. Los datos de baja calidad no afectan el rendimiento del modelo de Machine Learning.
- ☐ b. El modelo se entrena más rápido debido a la menor cantidad de datos.
- ☐ c. El modelo generaliza mejor y hace predicciones más precisas en datos reales.
- ☒ d. Los datos de baja calidad pueden resultar en un modelo inexacto y poco confiable con un rendimiento deficiente en datos reales. ✔

### Pregunta 24

Correcta

Se puntúa 1,00 sobre 1,00

¿En qué consiste el entrenamiento de un modelo de regresión lineal?

- ☐ a. El entrenamiento se refiere a la etapa en la que se diseñan las características del modelo, como su forma y tamaño.
- ☒ b. El entrenamiento implica ajustar los coeficientes del modelo (ecuación) para que se ajuste de la mejor manera posible a los datos y pueda hacer predicciones precisas. ✔
- ☐ c. El entrenamiento consiste en dividir los datos en un conjunto de entrenamiento y un conjunto de prueba.
- ☐ d. El entrenamiento se trata de seleccionar el mejor conjunto de variables independientes para el modelo.



Se puntúa 1,00 sobre 1,00

Usted tiene tres modelos de regresión lineal con los siguientes valores de  $R^2$  (coeficiente de determinación): 0.5, 0.3 y 0.8. ¿Cuál de los tres modelos es mejor?

- ☐ a. El modelo con un  $R^2$  de 0.3.
- ☒ b. El modelo con un  $R^2$  de 0.8. ✓
- ☐ c. El modelo con un  $R^2$  de 0.5.
- ☐ d. No se puede determinar cuál de los modelos es mejor solo con la información proporcionada.

#### Pregunta 26

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de las siguientes afirmaciones es cierta en el contexto de la clasificación en Machine Learning?

- ☐ a. La clasificación se utiliza únicamente en aplicaciones médicas para diagnosticar enfermedades.
- ☒ b. La clasificación se refiere a la tarea de predecir una etiqueta o categoría a un objeto o conjunto de datos en función de sus características. ✓
- ☐ c. La clasificación es una técnica utilizada exclusivamente en tareas de procesamiento de imágenes y visión por computadora.
- ☐ d. La clasificación se refiere a la tarea de predecir valores numéricos continuos a partir de datos de entrada.

Se puntúa 0,00 sobre 1,00

Se desea entrenar un modelo de regresión lineal para predecir la variable 'y' en función de una variable X. Seleccione el código correcto para realizar esta tarea:

- ☐ a. 

```
from sklearn.linear_model import LinearRegression
RLS = LinearRegression()
RLS.fit(X_train, y_train)
```
- ☒ b. 

```
from sklearn.linear_model import LinearRegression
RLS = LinearRegression()
RLS.fit(y_train, X_train)
```

 ✖
- ☐ c. 

```
from sklearn.linear_model import LinearRegression
RLS = LinearRegression()
RLS.fit(X_test, y_test)
```
- ☐ d. 

```
from sklearn.linear_model import LinearRegression
RLS = LinearRegression()
RLS.fit(X, y)
```

Respuesta incorrecta.

**Pregunta 28**

Correcta

Se puntúa 1,00 sobre 1,00

Cuando se trabaja con variables numéricas en aprendizaje automático, ¿cuál es el propósito principal del escalado o normalización?

- ☒ a. Alinear y ajustar las variables numéricas para que estén en una escala común y comparable. ✔
- ☐ b. Hacer que todas las variables numéricas tengan la misma media.
- ☐ c. Reducir la cantidad de datos en un conjunto de variables numéricas.
- ☐ d. Ajustar las variables numéricas para que tengan una desviación estándar de 1.



Se puntúa 1,00 sobre 1,00

¿Qué es el overfitting o sobre-ajuste en el contexto del Machine Learning?

- ☐ a. El overfitting es una técnica que se utiliza para mejorar la generalización de los modelos de Machine Learning.
- ☐ b. El overfitting es un término que se aplica solo a problemas de regresión en Machine Learning.
- ☐ c. El overfitting es un proceso necesario para entrenar modelos de Machine Learning de alta precisión.
- ☒ d. El overfitting ocurre cuando un modelo se ajusta demasiado a los datos de entrenamiento y captura el ruido en lugar de los patrones reales, lo que resulta en un mal rendimiento en datos no vistos. ✓

### Pregunta 30

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es la diferencia principal entre el aprendizaje por lotes y el aprendizaje en línea en Machine Learning?

- ☐ a. En el aprendizaje por lotes, los modelos se entrenan utilizando una única muestra de datos, mientras que en el aprendizaje en línea se utilizan múltiples muestras a la vez.
- ☐ b. En el aprendizaje por lotes, los modelos se actualizan de manera continua a medida que se reciben nuevos datos, mientras que en el aprendizaje en línea se entrenan una sola vez y no se actualizan.
- ☒ c. En el aprendizaje por lotes, los modelos se entrenan con todo el conjunto de datos de entrenamiento de una sola vez, mientras que en el aprendizaje en línea se actualizan de forma iterativa con datos a medida que llegan. ✓
- ☐ d. En el aprendizaje por lotes, los modelos se entrenan utilizando pequeños subconjuntos de datos a la vez, mientras que en el aprendizaje en línea se utilizan todos los datos de entrenamiento a la vez.

Usted se ha identificado como [ESTEBAN MARTIN ROMERO RODRIGUEZ](#) (Cerrar sesión)

[Política de privacidad](#)

[Resumen de retención de datos](#)

[Descargar la app para dispositivos móviles](#)

Los derechos de diseño, contenido y metodología son de uso exclusivo de la Pontificia Universidad Católica del Ecuador, queda prohibida su reproducción parcial o total.