

## Your grade: 71.42%

Your latest: 71.42% • Your highest: 71.42% • We keep your highest score. Review your overall course grades [here](#).

1. Con respecto a los métodos de reducción de la dimensionalidad ¿Cuáles de las siguientes afirmaciones son correctas? Selección múltiple.

1 / 1 point

La reducción de dimensionalidad, ya sea por selección o por transformación, puede ayudar a prevenir el sobreajuste de los modelos.

☒ Correct

Correcto. Reducir el espacio de características para el aprendizaje ayuda a prevenir el sobreajuste de los modelos.

La reducción de la dimensionalidad por transformación preserva el espacio original de características.

Los métodos de reducción de la dimensionalidad por selección facilitan la interpretación de los modelos.

☒ Correct

Correcto. En la reducción de la dimensionalidad por selección se elige un subconjunto de variables del conjunto de datos original, lo que facilita la interpretación de los modelos.

Todos los métodos por transformación pueden preservar la estructura de vecindad del conjunto de datos.

2. ¿Cuál de las siguientes afirmaciones describe el método de análisis de componentes principales (PCA)?

1 point

PCA busca preservar la similitud coseno entre los datos.

PCA preserva La varianza máxima al proyectar los datos en la dirección del vector propio principal.

☒ Correct

Correcto. PCA busca maximizar la varianza de los datos proyectados en las direcciones de los vectores propios principales.

PCA está diseñado para encontrar relaciones lineales entre las características.

PCA se considera un método de reducción de la dimensionalidad por selección.

You didn't select all the correct answers

3. El PCA con SVD (descomposición en valores singulares) se basa en la idea de que cualquier matriz puede descomponerse en un producto de matriz ortonormal, matriz diagonal y otra matriz ortonormal.

1 / 1 point

Verdadero

Falso

☒ Correct

Correcto. Una descomposición en valores singulares de una matriz de datos  $X$  es una factorización del tipo  $X=U\Sigma V'$ , donde  $U$  y  $V'$  son matrices ortonormales y  $\Sigma$  es una matriz diagonal que contiene los valores singulares. Si los datos están centrados en la media SVD y PCA suministran la misma información.

4. Se quiere investigar la relación entre diversos factores clínicos y el riesgo de desarrollar diabetes tipo 2 en una población. Inicialmente, se recopilaron datos de seis (6) variables diferentes, que incluyen: índice de masa corporal (BMI), niveles de glucosa en ayunas (FastingGlucose), presión arterial sistólica (SystolicBP), edad (Age), historial familiar de diabetes (FamilyHistory) y actividad física (PhysicalActivity). Se aplicó un análisis de componentes principales (PCA) para reducir la dimensionalidad del conjunto de datos y explorar la relación entre las variables. Los primeros tres componentes obtenidos se muestran en la siguiente Tabla:

1 / 1 point

