UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES ARAGÓN

INSTRUCCIONES GENERALES:

Lee con atención cada sección y resuelve los 25 reactivos interactivos sobre minería de datos.

Revisa tus respuestas antes de hacer clic en "Terminar".

SECCIÓN I. OPCIÓN MÚLTIPLE

Instrucciones: Haz clic sobre la letra correspondiente a la respuesta correcta.

1. ¿Qué diferencia a la minería de datos de las consultas tradicionales o los reportes descriptivos básicos?

a) Recupera únicamente información conocida de las bases transaccionales.

b) Resume la información conocida en tablas de filas y columnas.

c) Explora conjuntos amplios y complejos para revelar comportamientos no evidentes y anticipar escenarios futuros.

d) Elimina la necesidad de contar con datos históricos.

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

2. ¿Qué caracteriza a los modelos descriptivos?

a) Resumen, agrupan o caracterizan patrones existentes en los datos históricos sin requerir una variable objetivo predefinida.

b) Estiman el valor de una variable en nuevos registros a partir de una etiqueta previa.

c) Analizan únicamente datos ordenados en el tiempo.

d) Solo funcionan con datos no estructurados.

3. ¿Qué hace el modelo predictivo de clasificación?

a) Estima valores numéricos continuos en función de variables explicativas.

b) Divide los registros en grupos sin una etiqueta previa.

c) Localiza registros que se apartan de la norma estadística.

d) Asigna observaciones a categorías o clases discretas y predefinidas a partir de sus atributos de entrada.

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

4. ¿Qué busca evitar la prevención del sobreajuste (overfitting)?

a) Que el modelo se entrene con datos históricos.

b) Que el modelo "memorice" el ruido o las particularidades del conjunto de entrenamiento y pierda capacidad para predecir escenarios reales futuros.

c) Que los resultados sean explicables ante los responsables del área.

d) Que la información se divida en conjuntos de entrenamiento y prueba.

SECCIÓN II. SELECCIÓN Y SEÑALAMIENTO

Instrucciones: Marca la casilla de verificación correspondiente a la opción correcta.

5. Marca la casilla del modelo descriptivo que identifica relaciones de co-ocurrencia o afinidad entre ítems, como en el análisis de canasta de mercado:

[ ] Agrupamiento (Clustering)

[ ] Reglas de asociación

[ ] Detección de anomalías

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

6. Marca la casilla del modelo predictivo que estima valores numéricos continuos, como el monto de ventas esperado o la demanda futura de un insumo:

[ ] Clasificación

[ ] Patrones secuenciales

[ ] Regresión

7. Marca la casilla del modelo que examina mediciones sucesivas a intervalos regulares (diarios, semanales, mensuales) para proyectar tendencia, estacionalidad y variaciones cíclicas:

[ ] Series temporales

[ ] Reglas de asociación

[ ] Agrupamiento (Clustering)

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

SECCIÓN III. COMPLETAR MEDIANTE LISTA DESPLEGABLE

Instrucciones: Despliega el menú en cada espacio y selecciona el término correcto.

8. Los modelos descriptivos no requieren una variable [ ] predefinida o etiqueta previa para resumir, agrupar o caracterizar patrones en los datos históricos.

9. La detección de [ ] o valores atípicos (outliers) localiza registros que se apartan de manera significativa de la norma estadística, y es fundamental para identificar posibles fraudes financieros.

10. Para comprobar que un modelo no dependa de coincidencias particulares de los datos históricos, suele dividirse la información en conjuntos de entrenamiento y [ ] (o aplicarse validación cruzada).

11. Los datos no estructurados carecen de un modelo de datos o [ ] fijo predefinido, lo que dificulta su consulta directa mediante estructuras relacionales convencionales.

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

SECCIÓN IV. TEXTO EXACTO

Instrucciones: Escribe la palabra o siglas correctas en el recuadro asignado.

12. Escribe la palabra en inglés que designa la fuga o abandono de clientes, cuya probabilidad puede estimarse con modelos de clasificación:

Respuesta: [ ]

13. Escribe la palabra en inglés que designa el sobreajuste de un modelo a las particularidades del conjunto de entrenamiento (una sola palabra):

Respuesta: [ ]

14. Escribe las siglas, en español, del Procesamiento de Lenguaje Natural:

Respuesta: [ ]

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

SECCIÓN V. SOPA DE LETRAS INTERACTIVA

Instrucciones: Encuentra en la sopa de letras las 4 palabras clave que corresponden a los conceptos definidos en los reactivos 15, 16, 17 y 18. Haz clic sobre cada letra para seleccionarla.

R	R	T	Z	T	T	S	K	O	T	A	Z
H	U	F	R	E	S	E	F	C	Z	R	Z
I	B	V	C	N	C	N	A	O	A	Y	Y
I	H	I	D	D	Z	T	T	F	L	J	C
F	F	I	Q	E	F	I	V	I	U	W	J
S	O	W	K	N	P	M	P	D	C	A	J
E	M	K	N	C	Z	I	G	I	A	D	I
S	X	Q	G	I	T	E	N	A	N	H	A
G	M	E	B	A	X	N	F	O	A	W	Q
O	V	N	R	H	U	T	Z	W	S	Q	O
S	H	Q	U	A	M	O	V	S	T	Z	K
V	U	N	B	X	J	S	E	G	A	B	J

15. Palabra que completa el nombre del caso clásico de reglas de asociación: análisis de _____ de mercado (market basket analysis), que detecta qué productos suelen adquirirse juntos.

16. Componente de largo plazo que se descompone y proyecta en una serie temporal, junto con la estacionalidad recurrente y las variaciones cíclicas o irregulares.

17. Palabra que completa lo que permiten evaluar la minería de textos y el PLN respecto a clientes y usuarios: "evaluar _____ o percepciones".

18. Palabra que completa lo que es indispensable reducir al trabajar con textos, registros de usuarios o archivos multimedia: "la reducción de _____ en los datos y modelos".

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

SECCIÓN VI. ORDENAR SECUENCIA

Instrucciones: Ordena las etapas de preprocesamiento de datos no estructurados, según el orden en que la guía las presenta, arrastrando cada etiqueta a la posición numerada correcta.

[Extracción de características y representación]

[Limpieza y depuración]

[Normalización y estandarización]

19. Etapa 1: [ ]

20. Etapa 2: [ ]

21. Etapa 3: [ ]

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS

SECCIÓN VII. ARRASTRAR Y SOLTAR

Instrucciones: Arrastra la etiqueta de la caja de conceptos hacia su definición o recuadro correspondiente.

[Interpretabilidad y alineación con el negocio] **[Agrupamiento]**

[Patrones secuenciales] **[Análisis de imágenes, audio y video]**

22. Divide un conjunto de registros, clientes o productos en grupos con alta similitud interna, por ejemplo, agrupar clientes por frecuencia de compra o ticket promedio:

Respuesta: [ ]

23. Modelo que analiza el orden en que ocurren los eventos, por ejemplo, la secuencia de pasos que realiza un usuario antes de cancelar una suscripción:

Respuesta: [ ]

24. Revisión de si los resultados de un modelo tienen sentido operativo, son explicables ante los responsables del área y pueden integrarse en los procesos de decisión:

Respuesta: [ ]

25. Técnica que facilita la detección y clasificación de objetos o escenas, la transcripción de grabaciones y la extracción de metadatos o señales útiles para procesos operativos:

Respuesta: [ ]

Equipo 13
Marlon Giovanni Cortés Sepúlveda
Iván roberto Cruz Trejo
Salvador Flores Hernandez

LIVEWORKSHEETS