

Tema 01

Alumno: L.U.

15/11/21

Correo electrónico:

Pautas para la evaluación

- El examen es individual y según el tema asignado en planilla.
- El examen se desarrolla de 16:00 a 18:30 horas, con 15 minutos adicionales para preparar el documento final en formato pdf y cargarlo al Aula Virtual.
- **Prever:**
 - o **Cámara y micrófono del equipo deberá estar encendida sí o sí todo el tiempo.**
 - o **Tener a mano una identificación para mostrarla a la cámara cuando se le pida.**
 - o **Colocar su Nombre y Apellido en el ZOOM (no apodos, ni alias).**
 - o **Asegurarse en cuanto a la calidad de la/las fotos que envíen para realizar la evaluación (que sea legible).**
- Durante todo el tiempo, los alumnos deberán estar conectados a una videoconferencia que se habilitará mediante Zoom (asegurarse de contar con el equipamiento y conexión de internet adecuados).
- Desarrollo completo, agregar copias de planillas electrónicas y fotos al documento que remitan.
- En el plazo establecido, convertir el archivo a formato **PDF**, renombrarlo con el formato **BD2.Apellido.Nombre.SegundoRecuperatorio.pdf**
- Cada alumno, deberá cargar al Aula Virtual de la cátedra, en el plazo de tiempo establecido el documento final del examen parcial elaborado, quien no cumpla con este paso quedará fuera de la evaluación y deberá de recuperar en la siguiente instancia de ese parcial.

Marcar con una **X** la opción u opciones correctas, en el casillero ():

- 1- Algunas de las características de un almacén de datos respecto de un sistema operacional (OLTP), son:
 - a) Bases de datos pequeñas ()
 - b) Almacena datos históricos (X)
 - c) Soporta decisiones estratégicas ()
 - d) Ninguna es correcta ()
- 2- En las herramientas OLAP, algunos de los operadores de refinamiento de manipulación de consultas, pueden ser:
 - a) Pivot (X)
 - b) Drill (X)
 - c) Delete file ()
 - d) Ninguna es correcta ()
- 3- En el diseño y armado de un almacén de datos, para el modelado multidimensional tenemos que:
 - a) La información relevante sobre el hecho se representa por un conjunto de atributos. ()
 - b) No se puede representar una actividad que es objeto de análisis (hecho) ()
 - c) La información descriptiva de cada dimensión se representa por un conjunto de atributos (X)
 - d) Ninguna es correcta ()
- 4- Las principales formas de validación cruzada de la calidad predictiva de un árbol de decisión, son:
 - A) Validación cruzada en dos mitades (X)
 - B) Validación cruzada ad hoc ()
 - C) Validación cruzada en v partes (X)
 - D) Validación cruzada global (X)
 - E) Ninguna es correcta ()

- 5- Respecto a las herramientas para el trabajo con árboles de decisión, podemos mencionar:
- A) Familia Cart (X)
 - B) Método de tipo discriminante (X)
 - C) Método de tipo múltiple-complejo ()
 - D) Modelos híbridos (X)
 - E) Ninguna es correcta ()
- 6- Para la creación y aplicación de una red neuronal a un problema concreto se deben distinguir los siguientes pasos:
- A) Conceptualización del modelo para el estudio del modelo concreto. ()
 - B) Adecuación de la información de que se dispone a la estructura de la red a crear. (X)
 - C) Se constituirán los patrones de aprendizaje. (X)
 - D) Se constituirán los patrones de validación (X)
 - E) Ninguna es correcta. ()
- 7- Las Redes Neuronales:
- A) Tratan de emular el sistema nervioso (X)
 - B) Reflejan todas las características fundamentales del comportamiento del cerebro ()
 - C) Intenta modelizar una de las estructuras fisiológicas de soporte del cerebro, la neurona (X)
 - D) Construyen sistemas que presentan un alto grado de inteligencia ()
 - E) Ninguna es correcta ()
- 8- Respecto a los sistemas ROLAP:
- A) Se implementa sobre tecnología relacional (X)
 - B) Disponen de técnicas de compactación de datos ()
 - C) Disponen de estructuras de almacenamiento específicas como Arrays ()
 - D) Almacena físicamente los datos en estructuras multidimensionales de forma que la representación externa e interna coincidan. ()
 - E) Ninguna es correcta. ()
- 9- En el proceso KDD, algunas fases son:
- A) Estudio de la Minería ()
 - B) Evaluación y validación (X)
 - C) Normalización de datos ()
 - D) Interpretación y difusión (X)
 - E) Ninguna es correcta ()
- 10- Respecto a la fase de selección en KDD:
- A) Se implementan criterios de evaluación de hipótesis ()
 - B) Comprende la recolección e integración de la información (X)
 - C) Uso de ROLAP o MOLAP (X)
 - D) Se utiliza análisis exploratorio y grafico de los datos ()
 - E) Ninguna es correcta ()
- 11- Respecto a las técnicas descriptivas y predictivas de clasificación de minería de datos, entre las técnicas del análisis cluster, podemos mencionar:
- A) Métodos solapados-exclusivos (X)
 - B) Métodos directo-iterativos (X)
 - C) Métodos ponderados-no ponderados (X)
 - D) Métodos indirectos-iterativos ()
 - E) Ninguna es correcta ()
- 12- Respecto a las técnicas descriptivas y predictivas de clasificación de minería de datos, entre las técnicas del análisis cluster no jerárquico, podemos agrupar los métodos como:
- A) Designación ()
 - B) Reducción de dimensiones (X)
 - C) Indirectos ()
 - D) Búsqueda de la densidad (X)
 - E) Ninguna es correcta ()
- 13- Descrita el trabajo de monografía:" Fase de Selección en Minería de Datos."

14) Considerando la siguiente consulta para el análisis de datos:

Producto	Trimestre	Kg. vendidos
Tornillos	T1	40000
Tornillos	T2	60000
Tornillos	T3	12000
Tornillos	T4	14000
Clavos	T1	80000
Clavos	T2	12000
Clavos	T3	16000
Clavos	T4	20000

- Disgregar en 2 nuevos grupos cada Producto/Trimestre de la consulta original, considerando que las sucursales del Centro y Avenida conforman el universo estudiado, y las unidades vendidas se distribuyen en igual proporción en dichas sucursales.
- Presente matricialmente los datos seleccionados.
- Realice las operaciones necesarias para expresar la consulta obtenida en el punto a) en semestres.
- Represente el punto b) luego de utilizar el operador tipo Pivot.

15) Sea la siguiente definición del esquema anidado Ministerio:

Ministerio = (Nombre, Dirección, Teléfono, Funcionarios, Secretarías, Organigrama)

Funcionarios = (Nombre autoridad, Función)

Secretarías = (Nombre secretaría, Oficinas, Correo electrónico)

Oficinas = (Jefe, Subjefe)

Organigrama = (Jerarquías)

- Represente el esquema anidado Ministerio con el formato de una tupla
- Realice la representación en árbol del esquema anidado Ministerio

Puntajes:

Item 1 a 12	Bien contestada	0,50 Puntos
	No contestada	0
Item 13	1 Punto	
Item 14 y 15	1,5 Puntos c/u	

RESPUESTAS

13) Monografía Fase de Selección en Minería de Datos

Minería de datos es el proceso de extraer conocimiento útil y novedoso de grandes volúmenes de datos para poder lograr encontrar datos o información útiles que están ocultos, lo cual no podrían ser descubiertos a través de los métodos convencionales.

Generalmente la entrada a este proceso se forma con registros que proviene de bases de datos operacionales o Data Warehouse.

El modelo KDD(Descubrimiento de Conocimiento en Bases de datos) es el primero modelo que define el descubrimiento de conocimientos en base de datos como un proceso definido en fases o etapas.

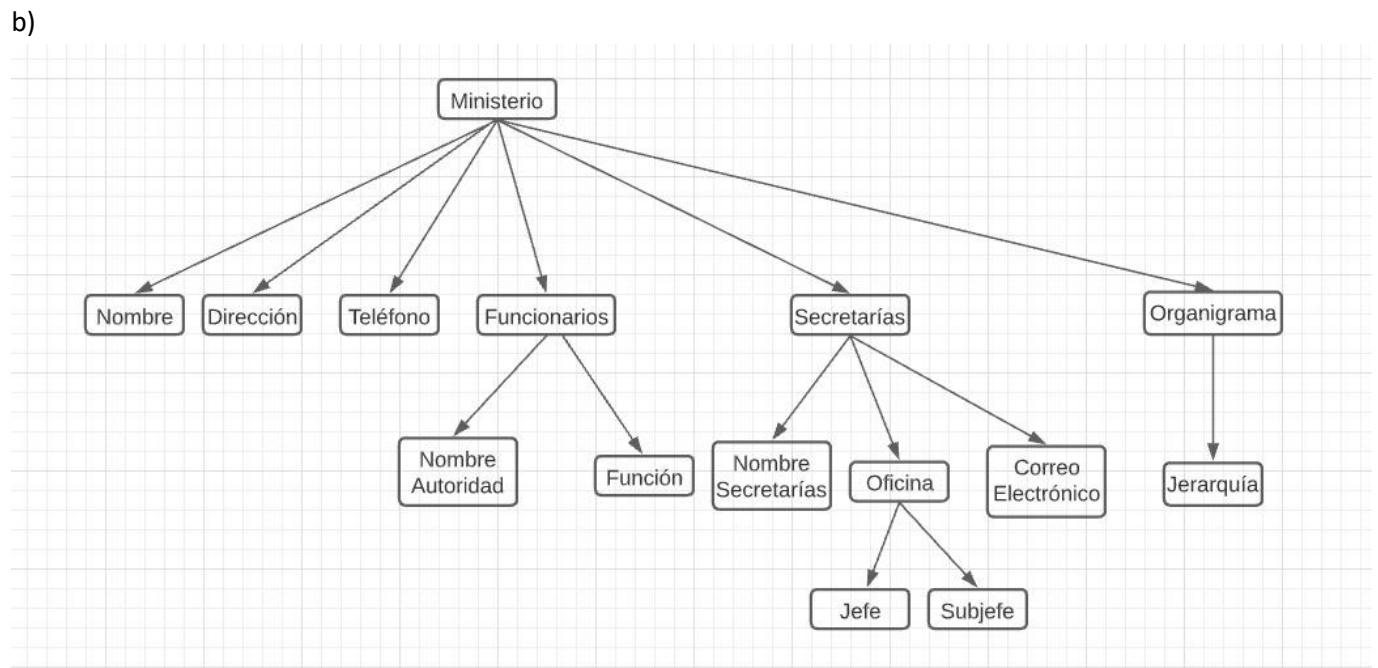
Las fases de este modelo KDD son: Selección de datos, preprocesamiento de datos, Transformación de datos y Reducción de dimensionalidad, Minería de Datos, y por último Interpretación y evaluación de los nuevos conocimientos.

La fase de selección de datos integra y recopila los datos y también determina las fuentes de información que pueden ser útiles.

Las fuentes pueden ser: OLAP ; OLTP; ROLAP O MOLAP; Data Warehouse.

15)a) Ministerio

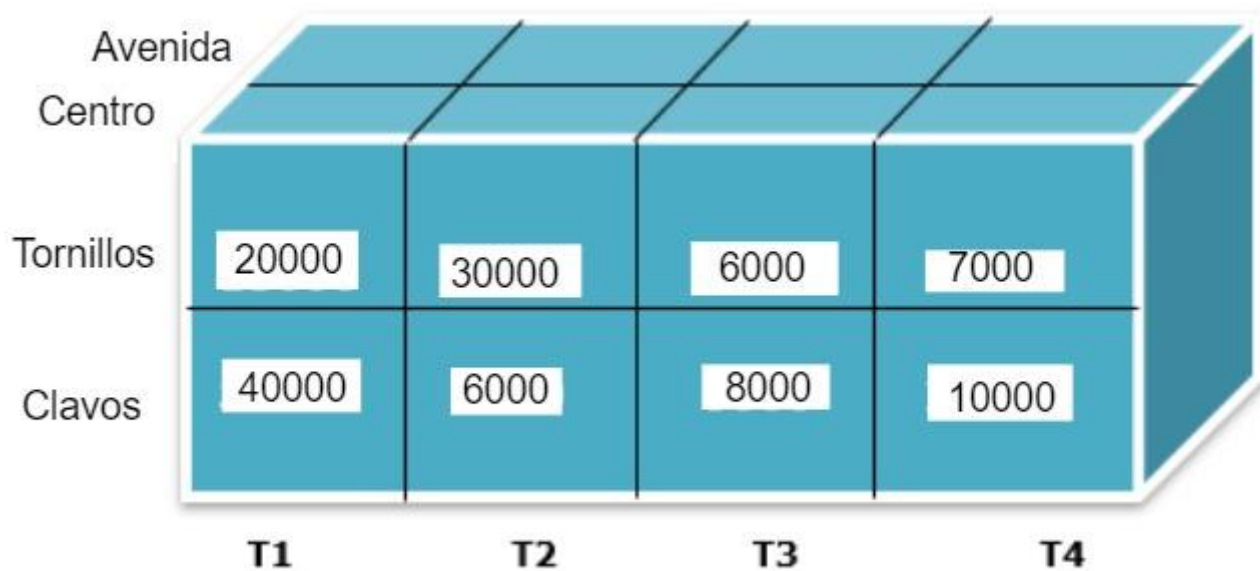
Nombre	Dirección	Teléfono	Funcionarios		Secretarías			Organigrama	
			Nombre autoridad	Función	nombre secretaría	oficinas		Correo Electrónico	Jerarquías
						jefe	subjefe		



14) a)

Producto	Trimestre	Sucursales	Kg. Vendidos
Tornillos	T1	Centro	20000
Tornillos	T1	Avenida	20000
Tornillos	T2	Centro	30000
Tornillos	T2	Avenida	30000
Tornillos	T3	Centro	6000
Tornillos	T3	Avenida	6000
Tornillos	T4	Centro	7000
Tornillos	T4	Avenida	7000
Clavos	T1	Centro	40000
Clavos	T1	Avenida	40000
Clavos	T2	Centro	6000
Clavos	T2	Avenida	6000
Clavos	T3	Centro	8000
Clavos	T3	Avenida	8000
Clavos	T4	Centro	10000
Clavos	T4	Avenida	10000

b)



c)

Producto	Semestre	Sucursales	Kg. Vendidos
Tornillos	S1	Centro	50000
Tornillos	S1	Avenida	50000
Tornillos	S2	Centro	13000
Tornillos	S2	Avenida	13000
Clavos	S1	Centro	46000
Clavos	S1	Avenida	46000
Clavos	S2	Centro	18000
Clavos	S2	Avenida	18000

d)

