

CUESTIONARIO METODOLOGÍA CRISP-DM

Computación Blanda – 2020-02

AGOSTO 28 DE 2020

NOMBRE: JULIAN GIRALDO CARDONA CÓDIGO: 1004752912

1. ¿Cuál es la fase que se considera probablemente la más importante de la metodología?

La fase que se considera probablemente la más importante de la metodología es:
La fase de comprensión del negocio o problema.

2. Haga una lista de las tareas generales de la fase de comprensión del negocio

- a. **Determinar los objetivos del negocio**

Tiene como metas, determinar cuál es el problema que se desea resolver, por qué la necesidad de utilizar Data Mining y definir los criterios de éxito.

- b. **Evaluación de la situación**

En esta tarea se debe calificar el estado de la situación antes de iniciar el proceso de DM.

- c. **Determinación de los objetivos de DM**

Esta tarea tiene como objetivo representar los objetivos del negocio en términos de las metas del proyecto de DM.

3. ¿En qué fase se establece el plan del proyecto?

El plan de proyecto se establece en la fase: ***Fase de comprensión de los datos.*** Esta comprende la recolección inicial de datos, con el objetivo de establecer un primer contacto con el problema.

4. ¿En qué fase se define el coste y beneficios del negocio?

Se define en la fase ***1. Fase de comprensión del negocio o problema,*** en la evaluación de la situación se definen los requisitos del problema, tanto en términos de negocio como en términos de Data Mining.

5. Definir los criterios de éxito del negocio, ¿es una tarea general o específica?
Es una tarea específica de la fase **1. Fase de comprensión del negocio o problema.**
6. ¿Cuáles son las fases que demandan mayor esfuerzo y tiempo en la metodología DM?
Las fases **2. Fase de comprensión de los datos, 3. Fase de preparación de los datos, 4. Fase de modelado.**
7. Enuncie las tareas específicas que deben realizarse en la verificación de la calidad de los datos.
En esta tarea, se efectúan verificaciones sobre los datos, para determinar la consistencia de los valores individuales de los campos, la cantidad y distribución de los valores nulos, y para encontrar valores fuera de rango, los cuales pueden constituirse en ruido para el proceso. La idea en este punto, es asegurar la completitud y corrección de los datos.
8. ¿En qué fase se realiza la limpieza de los datos?
En la fase **3. Fase de preparación de los datos**
9. ¿En qué fase se encuentra la necesidad (potencial) de la creación de nuevas estructuras de datos?
Se encuentra en la fase **3. Fase de preparación de los datos.**
10. ¿En qué fase se realiza la selección de la técnica de DM más apropiada?
Se realiza en **4. Fase de modelado**
11. Si se determina que un modelo produce resultados no satisfactorios en relación con los objetivos del negocio, ¿En qué fase se ha realizado dicho descubrimiento?
En la fase **5. Fase de evaluación.**
12. ¿A qué tarea general corresponde la actividad específica: “Plan de monitorización y mantenimiento?”
A la tarea **6. Fase de implementación.**