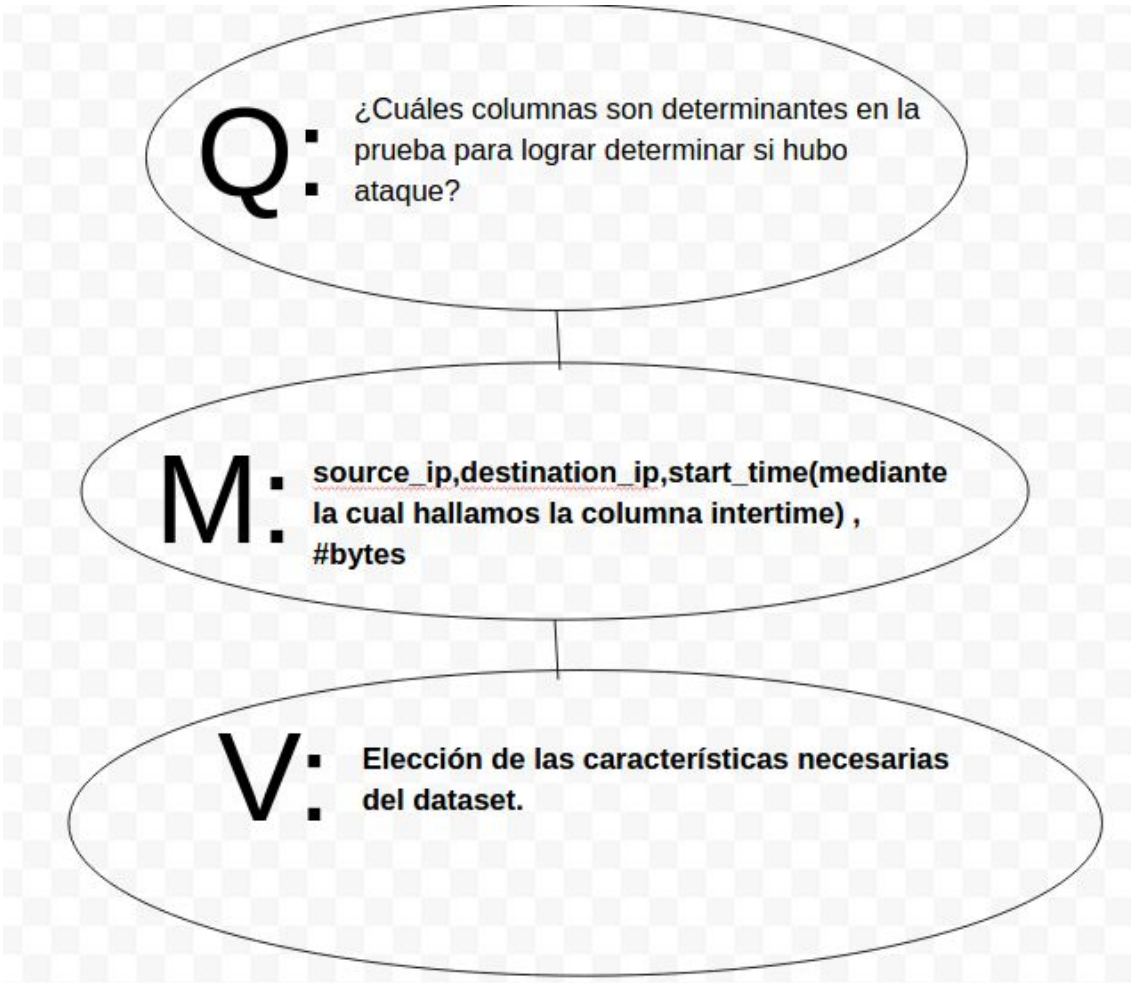
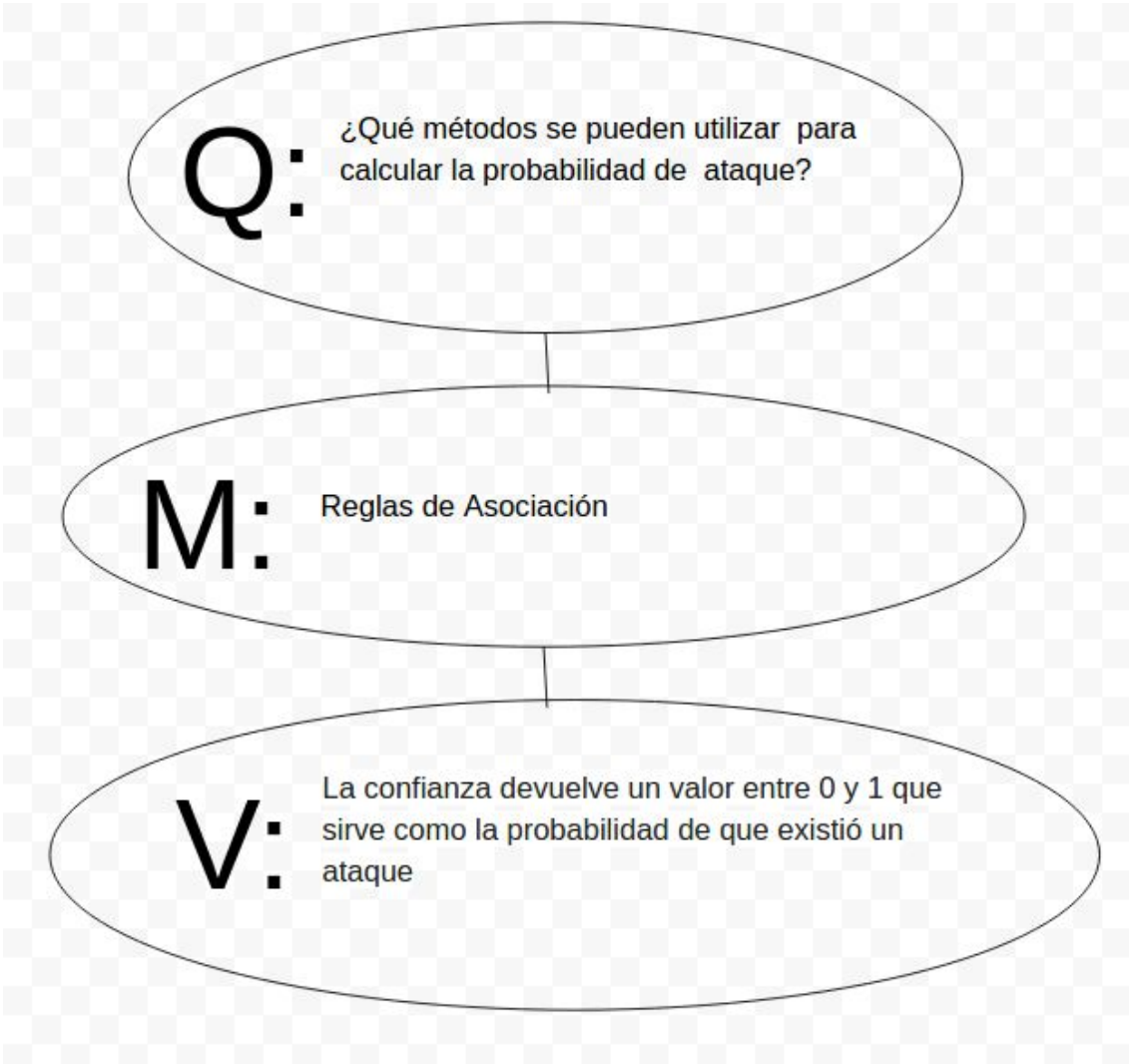
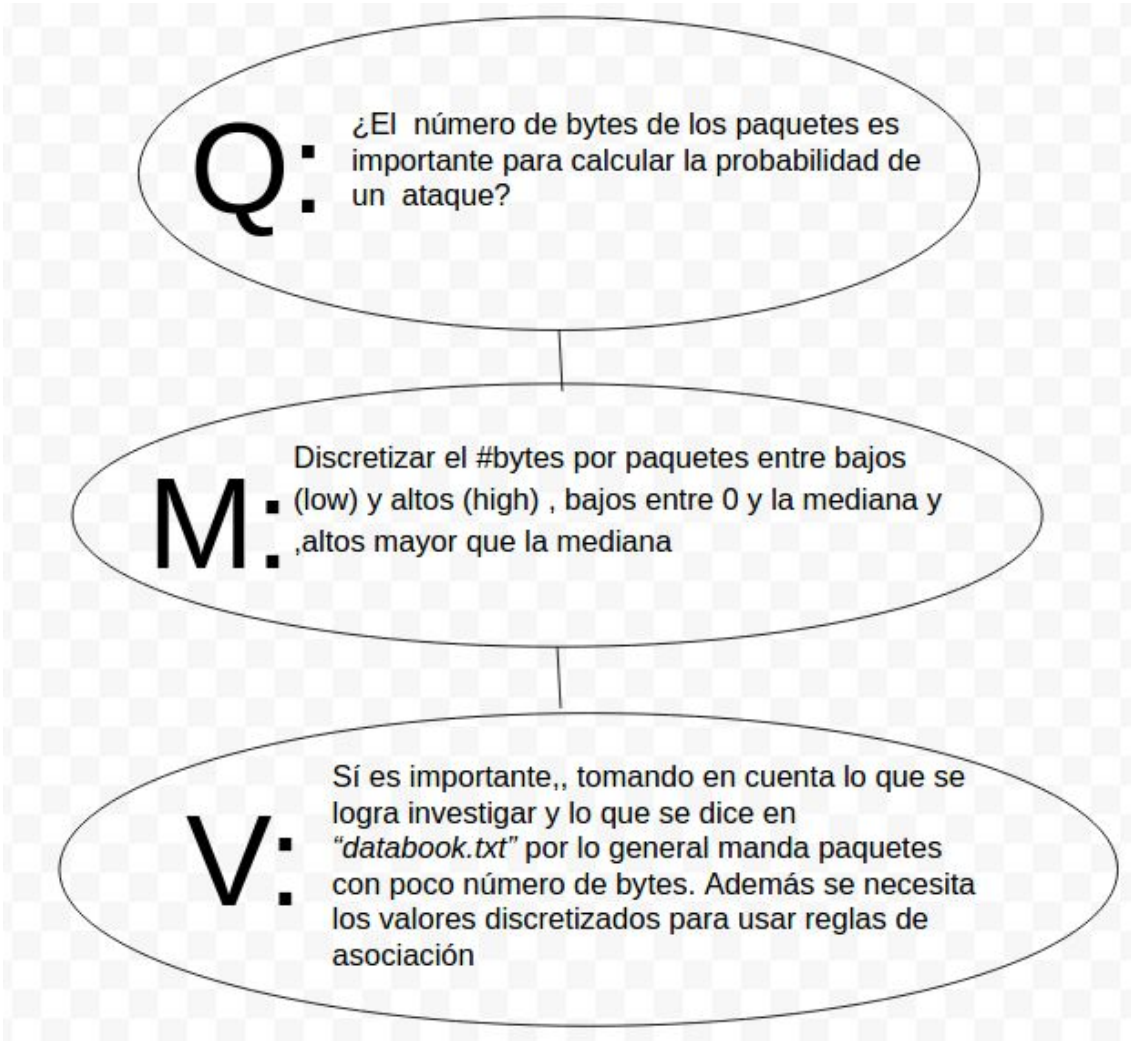
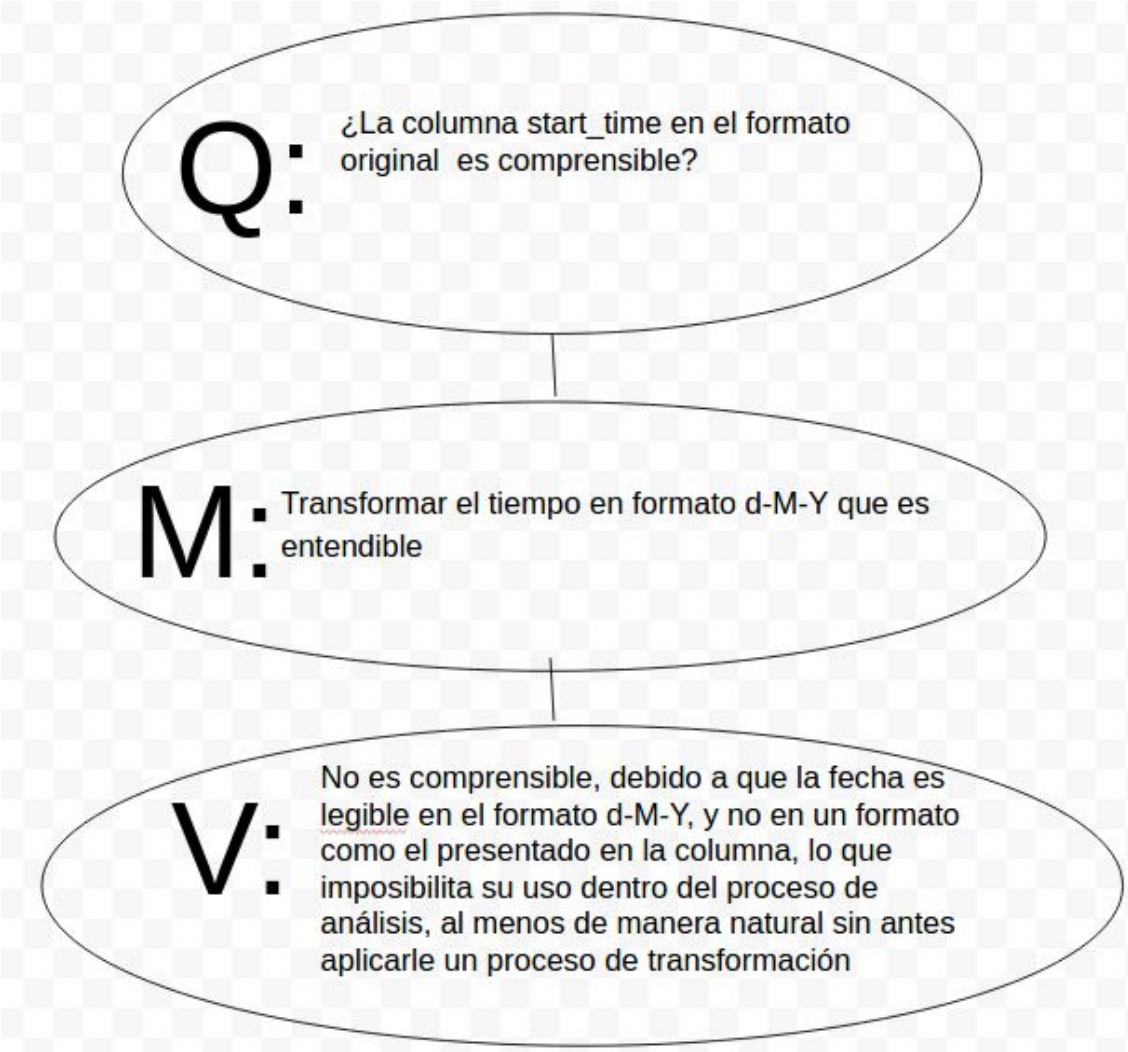

Q:
¿Cuáles columnas son determinantes en la prueba para lograr determinar si hubo ataque?
M:
source_ip,destination_ip,start_time(mediante la cual hallamos la columna intertime) , #bytes
V:
Elección de las características necesarias del dataset.


Q:
¿Qué métodos se pueden utilizar para calcular la probabilidad de ataque?
M:
Reglas de Asociación
V:
La confianza devuelve un valor entre 0 y 1 que sirve como la probabilidad de que existió un ataque


Q:
¿El número de bytes de los paquetes es importante para calcular la probabilidad de un ataque?
M:
Discretizar el #bytes por paquetes entre bajos (low) y altos (high) , bajos entre 0 y la mediana y ,altos mayor que la mediana
V:
Sí es importante,, tomando en cuenta lo que se logra investigar y lo que se dice en *"databook.txt"* por lo general manda paquetes con poco número de bytes. Además se necesita los valores discretizados para usar reglas de asociación


Q:
¿La columna start_time en el formato original es comprensible?
M:
Transformar el tiempo en formato d-M-Y que es entendible
V:
No es comprensible, debido a que la fecha es legible en el formato d-M-Y, y no en un formato como el presentado en la columna, lo que imposibilita su uso dentro del proceso de análisis, al menos de manera natural sin antes aplicarle un proceso de transformación