

EXAMEN DE TRATAMIENTO INTELIGENTE DE DATOS ENERO 2018

Nombre y Apellidos: \_\_\_\_\_

PREGUNTA	RESPUESTA	
La Minería de Datos es una parte de la Inteligencia Artificial	<b>V</b>	<b>F</b>
Muchas técnicas estadísticas forman parte de la Minería de Datos	<b>V</b>	<b>F</b>
Extracción del conocimiento y Minería de Datos son una misma cosa	<b>V</b>	<b>F</b>
Los datos a tratar en DM deben ser siempre numéricos	<b>V</b>	<b>F</b>
Para trabajar con Minería de Datos no es necesario partir de un dataset	<b>V</b>	<b>F</b>
Los datos continuos pueden ser demasiado detallados para ser usados por algunas técnicas de DM	<b>V</b>	<b>F</b>
El análisis en componentes principales funciona bien con variables continuas o como mucho enteras	<b>V</b>	<b>F</b>
El cálculo de frecuencias implica discretizar las variables continuas	<b>V</b>	<b>F</b>
El agrupamiento intenta clasificar items en clases ya establecidas	<b>V</b>	<b>F</b>
Los datos binarios tienen medidas de distancia específicas basadas en la concordancia de valores	<b>V</b>	<b>F</b>
El método de las k-medias trabaja mal cuando los grupos son "esféricos"	<b>V</b>	<b>F</b>
El DBSCAN necesita conocer el número de grupos al empezar	<b>V</b>	<b>F</b>
Los métodos jerárquicos son fácilmente escalables	<b>V</b>	<b>F</b>
El método de las k-medias difuso obliga a que cada punto pertenezca al menos a un grupo	<b>V</b>	<b>F</b>
Como es no supervisado la bondad de un agrupamiento no se puede medir bien	<b>V</b>	<b>F</b>
Las medidas basadas en entropía sirven para medir la calidad de un agrupamiento sin información previa	<b>V</b>	<b>F</b>
Las medidas basadas en entropía permiten comparar la coherencia entre dos agrupamientos	<b>V</b>	<b>F</b>
Clasificar equivale predecir una variable categórica	<b>V</b>	<b>F</b>
Ningún método de clasificación da información acerca de cómo dependen las clases de las variables implicadas	<b>V</b>	<b>F</b>
Los métodos de Redes Neuronales no dan información sobre la influencia de las variables en la clasificación	<b>V</b>	<b>F</b>

Para elegir el nodo a particionar hay que tener en cuenta alguna medida de la diversidad del conjunto que representa	V	F
No se pueden aplicar árboles de decisión cuando intervienen variables continuas en la clasificación	V	F
No se pueden generar reglas de clasificación si no es a partir de un árbol de decisión.	V	F
El algoritmo KNN sólo puede actuar con datos continuos y distancia euclídea	V	F
Al aplicar el algoritmo KNN hay que tener en cuenta la naturaleza de las variables	V	F
Las curvas ROC se pueden aplicar a cualquier problema de clasificación	V	F
Predecir una variable significa encontrar una función que la represente	V	F
El coeficiente de correlación mide la relación lineal entre dos variables	V	F
El coeficiente de correlación no se puede extender a varias variables	V	F
Se puede utilizar la regresión logística como un modelo de clasificación binaria con variables independientes continuas	V	F
Para aplicar reglas de asociación a atributos numéricos hay que discretizar previamente	V	F
El algoritmo a priori calcula itemsets frecuentes usando estrategias de poda	V	F
Cuando el consecuente de una regla es muy frecuente la confianza es muy baja	V	F
Como se obtienen muchas reglas de asociación no es fácil seleccionar las mejores de forma sencilla	V	F
El OLAP Mining consiste en hacer OLAP con datos discretizados	V	F