

MINERÍA DE DATOS - PRÁCTICA PREDICTIVA 30%

OBJETIVO DE LA PRACTICA:

- Aplicar todos los métodos de minería de datos supervisadas de vistas en clase
 - Métodos de Regresión
 - Árboles de decisión
 - Método Bayesiano
 - Método basado en ejemplos
 - Maquinas de soporte Vectorial
 - Redes neuronales
- Utilizar las herramientas de minería de Datos: R y Weka

OBJETIVO DE LA MINERÍA: Predecir el tipo de fármaco (drug) que se debe administrar a un paciente afectado de rinitis alérgica según distintos parámetros/variables. Se han recogido los datos del medicamento idóneo para muchos pacientes en cuatro hospitales. Se pretende, para nuevos pacientes, determinar el mejor medicamento a probar.

DESCRIPCIÓN DE LOS DATOS: Las variables disponibles en los historiales clínicos de cada paciente son:

- Age: Edad
- Sex: Sexo
- BP (Blood Pressure): Tensión sanguínea.
- Cholesterol: nivel de colesterol.
- Na: Nivel de sodio en la sangre.
- K: Nivel de potasio en la sangre.
- Drug: Hay cinco fármacos posibles:
 {DrugA, DrugB, DrugC, DrugX, DrugY}

****MODELAMIENTO:**

Se requiere:

- 1)** Preparar los datos para el análisis
- 2)** Aplicar todos los métodos vistos en clase con **R** y **WEKA**, finalmente analizar cada resultado obtenido (Observar muy bien los resultados obtenidos con WEKA).
- 3)** Hacer un informe en jupyter notebook indicando los parámetros óptimos utilizados en cada método y los resultados obtenidos para R.
- 4)** Hacer un informe (en el software que) indicando los parámetros óptimos utilizados en cada método y los resultados obtenidos para WEKA.
- 5)** Establecer un criterio para seleccionar el método que mejor se ajuste a las necesidades del problema. Justificar la elección.
- 6)** Crear un archivo de prueba donde se evalúen al menos 2 pacientes nuevos, de los cuales se debe predecir el tipo de medicamento a suministrar con el método seleccionado en el punto anterior, se debe hacer tanto para **R** como **WEKA**.

****ENTREGA:**

La entrega se realizara por medio de AULA DIGITAL

Fecha:

16 de Abril/2018 - 18:00 Horas (6:00 p.m)

Archivos:

Comprimir todos los archivos en un solo archivo .zip

1. Jupyter notebook con todo el trabajo en R (Puntos 1 hasta el 6 de modelamiento)
2. Informe con el trabajo de Weka
3. Archivo de prueba del punto 6 para WEKA

Nombrar el archivo así: Nombre_Apellido_ID.zip

Ejemplo: Peter_Parker_000000616.zip



****EVALUACIÓN**

Implementacion de los Metodos supervisados

No implemento todos los métodos en R o en WEKA

Opuntos

Implemento todos los métodos en R y en WEKA

50puntos

Optimizacion de Parametros

No se realizaron optimizacion de parametros en R o en WEKA

Opuntos

Se realizo la optimizacion de parametros en R y en WEKA

10puntos

Seleccion y Justificacion de Metodo

No se selecciono o no se justifico bien el metodo a utilizar

Opuntos

Se selecciona un método pero la justificación no esta bien.

10puntos

Se selecciona y se justifica bien el método que se utilizo

30puntos

Prueba del método

No se realiza a comprobación del método con nuevos datos en R o en Weka

Opuntos

Se probó el método con datos nuevos en R y WEKA

10puntos