



1. DATOS DEL TFG OFERTADO

Título del trabajo: Análisis del Conocimiento obtenido sobre Conjuntos de Datos Relacionados con la Biología mediante el Uso de Métodos de Extracción de Reglas de Asociación

En la actualidad han aumentado los estudios de Minería de Datos (MD) debido a las dificultades que existen para analizar grandes volúmenes de datos usando sólo los sistemas de información tradicionales. En general la MD consiste en extraer conocimiento interesante a partir de grandes conjuntos de datos. Una de las técnicas de MD más utilizada para extraer conocimiento interesante a partir de Bases de Datos (BD) ha sido el descubrimiento de reglas de asociación .

Las reglas de asociación son utilizadas para representar e identificar dependencias entre las variables de una BD. Estas reglas son expresiones del tipo $A \rightarrow B$, donde A y B son conjuntos de ítems (parejas variable-valor) y cumplen que $A \cap B = \emptyset$. Esto significa que si todos los ítems de A están en un ejemplo de la BD entonces todos los ítems de B están también en el ejemplo con una alta probabilidad, y A y B no tienen ningún ítem en común. El manejo y comprensión de éstas herramientas de la MD es de especial interés en el área de la biología, la química, la medicina, etc...

El objetivo del presente TFG es familiarizar al alumno con dicho tipo de técnicas mediante el uso de las mismas en algunos problemas relacionados con la biología. Para ello se propone el siguiente plan de trabajo:

- El alumno debe estudiar, comprender y resumir el funcionamiento de los algoritmos básicos más conocidos para la extracción de reglas. Se proponen los siguientes 5 algoritmos a estudiar: Apriori, FPGrowth, GENAR, MODENAR y Alatasetal..
- El alumno debe familiarizarse con la herramienta KEEL (<http://www.keel.es/>), para aplicar los 7 algoritmos nombrados en los siguientes 8 conjuntos de datos: Ecoli, Zoo, Abalone, Yeast, Breast, Bupa, newThyroid e Iris.
- El alumno debe realizar un análisis comparativo de los distintos algoritmos para determinar cuál es el más adecuado para cada problema. Adicionalmente, deberá analizar desde el punto de vista de un usuario final los conjuntos de reglas obtenidos por dicho algoritmo, mostrando que tiene capacidad para interpretar las reglas obtenidas.

Número de alumnos por trabajo ofertado: 1

Ofertado por:

1. Profesor del Departamento
2. Profesor del Departamento junto con Empresa ó Institución
3. Propuesto por alumno ()

| |
|---|
| X |
| |
| |

Palabras clave: Minería de datos, Bases de Datos, Extracción de reglas.

2. MODALIDAD

| | |
|---|-------------------------------------|
| 1. Trabajo bibliográfico | <input type="checkbox"/> |
| 2. Trabajo de investigación (*) | <input checked="" type="checkbox"/> |
| 3. Informe o proyecto de naturaleza profesional (*) | <input type="checkbox"/> |

3. DATOS DEL TUTOR Y COTUTOR (en su caso) DE LOS TFG OFERTADOS

| | | |
|--|---------------------|-------------------------------|
| Nombre y apellidos del tutor: JESÚS ALCALÁ FERNÁNDEZ | | |
| Teléfono: 958240429 | Fax: +34 958 243317 | e-mail: jalcala@decsai.ugr.es |

Granada, 24 de Junio 2014

Fdo: Director/a del Departamento de Ciencias de la Computación