



Universidad de Granada

**GRADO EN ESTADÍSTICA
PROPUESTA DE TEMA PARA TRABAJOS FIN DE GRADO
CURSO ACADÉMICO 2014/2015**

Título del trabajo: Selección de variables para clasificadores basados en probabilidades imprecisas
Tutor/a: Joaquín Abellán Mulero
Cotutor/a:
Departamento responsable: Ciencias de la Computación e Inteligencia Artificial
Perfil y número de estudiantes al que va dirigido (máximo 3): Estudiantes que cursen la asignatura de Minería de Datos. Número de alumnos : 1 <i>(Rellenar sólo en caso de que la propuesta esté realizada a través de estudiante)</i>
Estudiante que propone el trabajo (Nombre, Apellidos, DNI):
Tipo de trabajo <i>(consultar (*))</i> Trabajos de inicio a la investigación
Competencias <i>(estas son la mínimas; consultar (**) si se considera añadir otras)</i> <i>Competencias generales:</i> G02, G03, G04, G05, G06, G07, G08, G09. <i>Competencias específicas:</i> E01, E02, E03, E04, E05, E08, E09, E10.
Resultados de aprendizaje <i>(estos son los mínimos; añadir otros si se considera)</i> <ul style="list-style-type: none">• Adquirir competencias globales ligadas al desarrollo y aplicación de los conocimientos del Grado.• Adquirir competencias ligadas a la búsqueda y organización de información y documentación relevante sobre el tema objeto de estudio.• Aplicar el “pensamiento estadístico” y tener capacidad para enfrentarse a las distintas etapas de un estudio estadístico (desde el planteamiento del problema hasta la exposición de resultados).• Saber presentar, de forma escrita y oral, la memoria, los resultados y las conclusiones del trabajo realizado.
Antecedentes y resumen del tema propuesto: <p>La clasificación es un importante problema de la Minería de Datos en el que se pueden incorporar representaciones de la información basadas en probabilidades imprecisas (generalizaciones de la teoría de la probabilidad).</p> <p>Un método de clasificación puede ser precedidos de otros procedimientos que ayudan a mejorar el resultado obtenido, tales como selección/eliminación de variables, determinación y eliminación de datos erróneos, etc</p> <p>El software Weka permite, de forma eficiente, el manejo de clasificadores conocidos, así como la incorporación de nuevos.</p> <p>Se propone analizar distintos métodos de selección de variables para mejorar el funcionamiento de nuevos clasificadores basados en probabilidades imprecisas y medidas de información/incertidumbre, utilizando la herramienta Weka. Se compararan los resultados obtenidos a través de un conjunto de test de hipótesis. Para todo esto será necesario el manejo del módulo Experimenter en Weka, así como de otros paquetes estadísticos y/o software para Minería de Datos.</p> <p>Como paso final al trabajo anterior, se propone incorporar nuevos métodos de selección de variables al entorno Weka, realizando pequeñas modificaciones de los existentes.</p>



**GRADO EN ESTADÍSTICA
PROPUESTA DE TEMA PARA TRABAJOS FIN DE GRADO
CURSO ACADÉMICO 2014/2015**

Universidad de Granada

Breve descripción de las actividades presenciales y no presenciales a realizar:

Actividades presenciales (15-30%)	Planteamiento, orientación y supervisión	45 horas
	Exposición del trabajo	15 horas
	Otras:	
Actividades no presenciales (70-85%)	Preparación del trabajo	200 horas
	Elaboración de la memoria	40 horas
	Otras:	
Total (12 ECTS)		300 horas

Objetivos que se pretenden alcanzar:

- Extender los conocimientos obtenidos en las asignaturas de Minería de Datos e Informática I, pertenecientes al Grado en Estadística.
- Conocer conceptos básicos sobre probabilidades imprecisas y su uso en Minería de Datos, con enfoque principal en clasificación.
- Uso del Experimenter en Weka y manejo a nivel programación de dicho software.
- Inicio a la investigación

Bibliografía básica para la puesta en marcha del trabajo:

- J. Abellán y S. Moral, Building classification trees using the total uncertainty criterion, International Journal of Intelligent Systems, 18(12) (2003), 1215-1225.
- I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.