



1. DATOS BÁSICOS DEL TFG:

Título: Análisis de datos asociados a problemas de clasificación mediante técnicas de minería de datos

Descripción general (resumen y metodología):

Actualmente hay gran cantidad de problemas de clasificación cuyos conjuntos de datos están disponibles por Internet para su estudio, por ejemplo, clasificación de enfermedades, clasificación de plantas, clasificación de platos de comida etc. Por otro lado, en los últimos años se ha mejorado el proceso de análisis y extracción de conocimiento a partir de grandes cantidades de datos por el incremento en la velocidad de cómputo de los ordenadores y por el desarrollo de nuevos algoritmos de minería de datos.

De estos dos hechos podemos deducir un interesante estudio consistente en aprovechar las capacidades de los algoritmos de minería de datos para extraer conocimiento de un conjunto de datos asociado a un problema de clasificación. No se trataría de obtener una caja negra que me clasifique correctamente sino más bien el objetivo consistiría en obtener conocimiento comprensible que me explique el por qué una instancia del conjunto de datos tiene asignado un valor de clase determinado.

Tipología: Estudio de casos, teóricos o prácticos, relacionados con la temática del Grado.

Objetivos planteados:

-) Buscar y seleccionar un conjunto de datos asociado a un problema de clasificación.
-) Limpiar estos datos y prepararlos para su análisis mediante técnicas de minería de datos.
-) Aplicar algoritmos de minería de datos sobre los datos anteriores.
-) Obtener información comprensible referente a los valores de clasificación asociados a los datos.
-) Redactar una memoria detallando todos los pasos desarrollados y objetivos logrados.

Bibliografía básica:

- Scott V. Burger.(2018). Introduction to Machine Learning with R. Rigorous Mathematical Modeling. O'Reilly Media.
- Witten, Ian H., Eibe Frank, and Mark A Hall. (2011). Data Mining: Practical Machine Learning Tools and Techniques . 3a. ed., Elsevier.
- B. Lantz. (2015). Machine Learning with R: Expert techniques for predictive modeling to solve all your data analysis problems, 2nd Edition.
- Han, J., & Kamber, M. (2001). Data mining: Concepts and techniques. Morgan Kaufmann.

Recomendaciones y orientaciones para el estudiante:

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: CARLOS JAVIER MANTAS RUIZ

Ámbito de conocimiento/Departamento: CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL

Correo electrónico: cmantas@ugr.es

3. COTUTOR/A DE LA UGR (en su caso):

Nombre y apellidos:

Ámbito de conocimiento/Departamento:

Correo electrónico:

4. COTUTOR/A EXTERNO/A (en su caso):

Nombre y apellidos:

Correo electrónico:

Nombre de la empresa o institución:

Dirección postal:

Puesto del tutor en la empresa o institución:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos:

Correo electrónico:



Universidad de Granada

GRADO EN ESTADÍSTICA
PROPUESTA DE TEMA PARA TRABAJOS FIN DE GRADO
CURSO ACADÉMICO 2024/2025

Título del trabajo: Análisis de datos asociados a problemas de clasificación mediante técnicas de minería de datos
Tutor/a: Carlos J. Mantas Ruiz Correo electrónico: cmantas@decsai.ugr.es
Cotutor/a:
Departamento responsable: Ciencias de la Computación e I.A.
Perfil y número de estudiantes al que va dirigido (máximo 2): 2 (Dos TFGs de 1 estudiante) <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> 8,9
Competencias <i>(estas son las 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>Actualmente hay gran cantidad de problemas de clasificación cuyos conjuntos de datos están disponibles por Internet para su estudio, por ejemplo, clasificación de enfermedades, clasificación de plantas, clasificación de platos de comida etc. Por otro lado, en los últimos años se ha mejorado el proceso de análisis y extracción de conocimiento a partir de grandes cantidades de datos por el incremento en la velocidad de cómputo de los ordenadores y por el desarrollo de nuevos algoritmos de minería de datos.</p> <p>De estos dos hechos podemos deducir un interesante estudio consistente en aprovechar las capacidades de los algoritmos de minería de datos para extraer conocimiento de un conjunto de datos asociado a un problema de clasificación. No se trataría de obtener una caja negra que me clasifique correctamente sino más bien el objetivo consistiría en obtener conocimiento comprensible que me explique el por qué una instancia del conjunto de datos tiene asignado un valor de clase determinado.</p>



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

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

Objetivos que se pretenden alcanzar:

-) Buscar y seleccionar un conjunto de datos asociado a un problema de clasificación.
-) Limpiar estos datos y prepararlos para su análisis mediante técnicas de minería de datos.
-) Aplicar algoritmos de minería de datos sobre los datos anteriores.
-) Obtener información comprensible referente a los valores de clasificación asociados a los datos.
-) Redactar una memoria detallando todos los pasos desarrollados y objetivos logrados.

Bibliografía básica para la puesta en marcha del trabajo:

- Scott V. Burger.(2018). Introduction to Machine Learning with R. Rigorous Mathematical Modeling. O'Reilly Media.
- Witten, Ian H., Eibe Frank, and Mark A Hall. (2011). Data Mining: Practical Machine Learning Tools and Techniques . 3a. ed., Elsevier.
- B. Lantz. (2015). Machine Learning with R: Expert techniques for predictive modeling to solve all your data analysis problems, 2nd Edition.
- Han, J., & Kamber, M. (2001). Data mining: Concepts and techniques. Morgan Kaufmann.

Tipo de trabajo (*):

1. Estudio de profundización en algún tema concreto de Estadística, o como proyecto de aplicación de la misma a estudios o problemas de otros ámbitos científicos o sociales.
2. Realización completa de todas las fases de un proyecto estadístico, bien con auxilio de prácticas en empresas o con prácticas propuestas y dirigidas por el tutor.
3. Estudio de casos, teóricos o prácticos, relacionados con la Estadística.
4. Elaboración de un informe o un proyecto de naturaleza profesional.
5. Elaboración de un plan de empresa.
6. Simulación de encargos profesionales.
7. Trabajos bibliográficos sobre el estado actual de una temática relacionada con la Estadística.
8. Creación y/o empleo de herramientas informáticas para su uso en Estadística.



9. Trabajos de inicio a la investigación.
10. Trabajos cuya finalidad sea la divulgación de la Estadística en diversos contextos.
11. Trabajos sobre Historia de la Estadística.
12. Trabajos relacionados con la docencia de la Estadística.

Competencias ()**

Competencias generales:

G01. Poseer los conocimientos básicos de los distintos módulos que, partiendo de la base de la educación secundaria general, y apoyándose en libros de texto avanzados, se desarrollan en la propuesta de título de Grado en Estadística que se presenta.

G02. Saber aplicar los conocimientos básicos de cada módulo a su trabajo o vocación de una forma profesional y poseer las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de la Estadística y ámbitos en que esta se aplica directamente.

G03. Saber reunir e interpretar datos relevantes para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.

G04. Poder transmitir información, ideas, problemas y sus soluciones, de forma escrita u oral, a un público tanto especializado como no especializado.

G05. Haber desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

G06. Saber utilizar herramientas de búsqueda de recursos bibliográficos.

G07. Poder comunicarse en otra lengua de relevancia en el ámbito científico.

G08. Poseer habilidades y aptitudes que favorezcan el espíritu emprendedor en el ámbito de aplicación y desarrollo de su formación académica.

G09. Fomentar y garantizar el respeto a los Derechos Humanos, a los principios de accesibilidad universal, igualdad, y no discriminación; y los valores democráticos, de la cultura de la paz y de igualdad de género.

Competencias específicas:

E01. Conocer los fundamentos básicos del razonamiento estadístico, en el diseño de estudios, en la recogida de información, en el análisis de datos y en la extracción de conclusiones.

E02. Conocer, saber seleccionar y saber aplicar, técnicas de adquisición de datos para su tratamiento estadístico.

E03. Conocer los fundamentos teóricos y saber aplicar modelos y técnicas estadísticas en estudios y problemas reales en diversos ámbitos científicos y sociales.

E04. Saber seleccionar los modelos o técnicas estadísticas para su aplicación en estudios y problemas reales en diversos ámbitos científicos y sociales, así como conocer herramientas de validación de los mismos.



Universidad de Granada

**GRADO EN ESTADÍSTICA
PROPUESTA DE TEMA PARA TRABAJOS FIN DE GRADO
CURSO ACADÉMICO 2024/2025**

- E05.** Comprender la importancia de la Investigación Operativa como metodología de optimización, toma de decisiones y diseño de modelos particulares para la resolución de problemas en situaciones específicas.
- E06.** Comprender y utilizar básicamente el lenguaje matemático.
- E07.** Conocer los conceptos y herramientas matemáticas necesarias para el estudio de los aspectos teóricos y prácticos de la Probabilidad, la Estadística y la Investigación Operativa.
- E08.** Conocer y saber utilizar aplicaciones informáticas de análisis estadístico, cálculo numérico y simbólico, bases de datos, visualización gráfica y optimización, que sean útiles para la aplicación y desarrollo de las técnicas estadísticas.
- E09.** Conocer los conceptos básicos y habilidades propias de un ámbito científico o social en el que la Estadística o la Investigación operativa sean una herramienta fundamental.
- E10.** Tomar conciencia de la necesidad de asumir las normas de ética profesional y las relativas a la protección de datos y del secreto estadístico, como premisas que deben guiar la actividad profesional como profesionales de la Estadística.