

Título del trabajo: Minería de Datos No Estructurados
Tutor/a: Francisco Javier Arnedo Fernández Correo electrónico: arnedo@ugr.es
Cotutor/a:
Departamento responsable: Estadística e Investigación Operativa
Perfil y número de estudiantes al que va dirigido (máximo 2): 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): Sebastián Joaquín Figueras López
Tipo de trabajo (consultar (*)) 1. Estudio de profundización
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>La Minería de Datos es una rama de la Estadística y de las Ciencias de Computación que intenta descubrir patrones y extraer conocimiento en grandes conjuntos de datos. Dentro de la Minería de Datos se incluyen problemas como la agrupación (clustering) o la clasificación.</p> <p>Los datos disponibles para tales fines pueden ser tanto estructurados como no estructurados. Los datos estructurados siguen una estructura predefinida, de forma tabular, que los hace fácilmente analizables. Por el contrario, los datos no estructurados, los cuales se estiman en un 80% del total de los datos disponibles, pueden tener cualquier tipo de estructura, como por ejemplo textos, imágenes, audios o videos, y su análisis es mucho más complejo.</p> <p>Durante el Grado en Estadística, y en concreto en la asignatura de Minería de Datos se trabaja con datos estructurados, por esta razón se propone este Trabajo Fin de Grado donde se profundice en las técnicas de análisis de los diferentes tipos de datos no estructurados. Tras una recopilación bibliográfica, se definirán los conceptos y técnicas más frecuentemente utilizados y se presentarán casos prácticos utilizando las principales implementaciones de estas técnicas.</p>

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

Actividades presenciales (15-30%)	Planteamiento, orientación y supervisión	35 horas
	Exposición del trabajo	10 horas
	Otras:	
Actividades no presenciales (70-85%)	Preparación del trabajo	150 horas
	Elaboración de la memoria	50 horas
	Otras:	55 horas
Total (12 ECTS)		300 horas

Objetivos que se pretenden alcanzar:

- Realizar un estudio bibliográfico del estado del arte.
- Introducir los principales tipos de datos no estructurados y cómo analizarlos con técnicas de minería de datos.
- Presentar implementaciones de casos de uso.

Bibliografía básica para la puesta en marcha del trabajo:

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013) An introduction to statistical learning. New York: springer.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction. New York: springer.
- Zhang, D. (2019). Fundamentals of image data mining. Springer International Publishing.
- Godec, P., Pančur, M., Ilenič, N. et al. (2019) Democratized image analytics by visual programming through integration of deep models and small-scale machine learning. Nature Communications 10, 4551
- Zong, C., Xia, R., & Zhang, J. (2021). Text data mining. Singapore: Springer.

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.



UNIVERSIDAD
DE GRANADA

GRADO EN ESTADÍSTICA
PROPUESTA DE TEMA PARA TRABAJOS FIN DE GRADO
CURSO ACADÉMICO 2023/2024

- 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.
- 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.