



Línea de Trabajo fin de Máster

(Fecha última actualización: ...02/03/2022.....)

Máster en Estadística. CURSO ACADÉMICO 2021-2022	
Título	Análisis de datos de proximidad y su tratamiento computacional.
Tipo	INVESTIGACIÓN <input checked="" type="checkbox"/> ORIENTACIÓN PRÁCTICA <input type="checkbox"/>
Número de alumnos	2
Profesor(es)/ email	José Fernando Vera
Descripción	<p>Uno de los problemas más interesantes en muchas disciplinas y que forma parte de lo que hoy día se conoce como “aprendizaje estadístico no supervisado”, se plantea cuando necesitamos medir y entender las relaciones entre objetos o variables, siendo desconocidas las dimensiones subyacentes de los mismos, especialmente cuando la información disponible se refiere exclusivamente a alguna medida de similaridad o de disimilaridad entre ellos. Tanto “Multidimensional Scaling” (MDS) para datos de proximidad como Unfolding para datos de preferencias, permiten el análisis de esos datos, así como su combinación con otras técnicas tradicionales del Análisis Multivariante como el análisis clúster para facilitar la interpretación de los datos o abordar problemas relacionados con “big data”. Estos modelos también se han combinado con técnicas estadísticas clásicas como el análisis de ecuaciones estructurales, los modelos log-lineales o los modelos logit.</p> <p>Uno de los aspectos importantes consiste en el uso de información auxiliar para la interpretación de los resultados puesto que existe un problema de indeterminación e las soluciones obtenidas. Así, la aplicación y el desarrollo de técnicas específicas para cada problema en cuestión y su tratamiento computacional constituye una área de investigación muy atractivo y de gran actualidad, en la que diversas técnicas estadísticas y de optimización heurística juegan actualmente un papel importante.</p>
Objetivos particulares	Estudio y tratamiento computacional de técnicas estadísticas multivariantes para la representación e interpretación multidimensional de datos de proximidad.
Prerrequisitos y recomendaciones	Haber cursado o estar cursando la asignatura: Análisis de datos. Técnicas aplicadas a datos de proximidad
Plan de trabajo	<p>Estudio de las técnicas adecuadas para la resolución de problemas y el análisis de datos de proximidad.</p> <p>Estudio de software estadístico para la aplicación de las técnicas: SPSS, R, MatLab</p> <p>Determinación de un modelo aplicado para el análisis real de datos de proximidad.</p> <p>Elaboración de la memoria y presentación de resultados.</p>
Competencias generales y específicas	<p>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.</p> <p>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.</p> <p>G03. Saber reunir e interpretar datos relevantes para emitir juicios que incluyan una</p>



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.

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

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.

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.

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.

Bibliografía

1. BORG, I. & GROENEN, P.J.F. (2005).- Modern Multidimensional Scaling. Theory and Applications. Second Edition. Springer Series in Statistics. Springer.
2. COX, T.F. & COX, M.A.A. (2001).-Multidimensional Scaling. Second Edition. Monographs on statistics and applied probability, 59. London: Chapman Hall.
3. HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2010). The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Second Edition. Springer Series in Statistics. Springer.
4. KRUSKAL, B. & WISH, M (1981). Multidimensional Scaling. Sage.
5. MARDIA K.V. & BIBBY, K.J.M. (1997). Multivariate Analysis. London: Academic Press.
6. HOAGLIN, D., MOSTELLER, F. & TUKEY, J.W.(1983).-Understanding robust and Exploratory Data Analysis. New York. John Wiley & Sons.
7. SEBER, G.A.F. (1984). Multivariate Observations. New York. John Willey.
8. ARABIE, P. & CARROLL, J.D. & DeSARBO, W. (1987). Three-Way scaling and clustering. Sage.
9. VERA, J.F. and HEISER, W.J., and MURILLO, A. (2007). Global Optimization in Any Minkowski Metric: A Permutation-Translation Simulated Annealing Algorithm for Multidimensional Scaling. Journal of Classification, 24, 277-301