



1. DATOS BÁSICOS DEL TFG:

Título: Análisis estadístico de datos de disimilaridad

Descripción general (resumen y metodología):

Antecedentes y resumen del tema propuesto:

En general, resulta cada vez más frecuente en muchas disciplinas el análisis estadístico de datos basado en datos de proximidad entre los elementos de estudio (objetos, individuos, series, funciones, modelos, etc.), bien porque se desconocen cuáles pueden ser las variables involucradas, porque éstas no se pueden medir o son latentes, por la presencia de datos faltantes o simplemente porque la naturaleza de las variables no se adecúa a las técnicas estadísticas usuales para su análisis y requieren de un análisis basado en datos de proximidad. El análisis de datos basado en distancias es por tanto una disciplina fundamental en muchas áreas y sobre todo en los problemas que involucran una cantidad ingente de datos (medicina, genética, marketing, ciencias sociales, etc.) relacionados con la reducción de la dimensión, el análisis cluster y el análisis discriminante, entre otras.

La temática general de esta propuesta de TFG consiste en el estudio y aplicación de técnicas estadísticas y computacionales para el análisis de datos de disimilaridad y de preferencia mediante Multidimensional Scaling (MDS) y Unfolding. La propuesta no solo se limita al estudio y aplicación de las técnicas sino también al aprendizaje de la metodología de aplicación en el área particular. Esta propuesta de trabajo permite la realización de hasta un máximo de 2 TFG por 2 alumnos en función de tres módulos de actuación: Así, cada trabajo podría centrarse en uno de los objetivos del análisis de datos anteriormente descritos.

Cada trabajo contemplará una parte teórica asociada a uno de los objetivos descritos anteriormente, una parte computacional que consistirá en la recopilación bibliográfica y descripción general de la metodología para el desarrollo del objetivo correspondiente, y finalmente la aplicación mediante el análisis de datos reales o simulados, principalmente con R. Finalmente, el TFG también puede contemplar la metodología correspondiente para el análisis de datos en un problema real.

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

Objetivos planteados:

Para cada módulo, esta propuesta de trabajo fin de grado contempla los siguientes objetivos, contenidos y actividades concretas a desarrollar para cada módulo:

- Descripción de la problemática abordada por los métodos correspondiente y/o del encargo profesional o problemática real abordada, si es el caso.
- Recopilación y breve descripción teórica de las técnicas más utilizadas para la problemática concreta elegida.
- Recopilación y descripción del software estadístico para el análisis de los datos relacionados con el área.
- Análisis de los datos y comparación de resultados entre las distintas técnicas.
- Conclusiones y presentación de resultados.

Cabe la posibilidad de que el análisis real de datos esté relacionado con prácticas en una empresa.

Objetivos que se pretenden alcanzar:

- Conocimiento de la metodología para abordar una problemática real para el análisis de datos de proximidad.
- Conocimiento de las técnicas estadísticas relacionadas con el problema.
- Conocimiento del software adecuado para el análisis con MDS y/o Unfolding y presentación de resultados.

Bibliografía básica:

1. BORG, I, GROENEN, P. (2005). Modern Multidimensional Scaling. Theory and Applications. Springer.
2. BORG, I, GROENEN, P., and MAIR, P. (2018). Applied Multidimensional Scaling and Unfolding, 2nd ed. New York, NY: Springer Science & Business Media.
3. COX, T.F. & COX, M.A.A. (2001). Multidimensional Scaling. Second Edition. Monographs on statistics and applied probability, 59. London: Chapman Hall.
4. HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2010). The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Second Edition. Springer Series in Statistics. Springer.
5. MARDIA K.V. & BIBBY, K.J.M. (1997). Multivariate Analysis. London: Academic Press.

Recomendaciones y orientaciones para el estudiante:

Breve descripción de las actividades presenciales y no presenciales a realizar: A título orientativo, puede sintetizarse una breve descripción de las actividades a realizar correspondientes a 12 créditos ECTS que equivalen según la normativa actual a 300 horas de trabajo total (1 ECTS = 25 horas): Actividades presenciales (15-30%) Planteamiento, orientación y supervisión: 90 horas Exposición del trabajo: 1 horas Otras: 9 horas. Actividades no presenciales (70-85%). Preparación del trabajo: 250 horas. Elaboración de la memoria: 50 horas: Total 300 horas (12 ECTS).

Plazas: 2

2. DATOS DEL TUTOR/A:

Nombre y apellidos: JOSÉ FERNANDO VERA VERA

Ámbito de conocimiento/Departamento: ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

Correo electrónico: jfvera@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:

Centro de convenio Externo:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos:

Correo electrónico: