



1. DATOS BÁSICOS DEL TFG:

Título: Ciencia de Datos para análisis de comunidades y temas de conversación en redes sociales

Descripción general (resumen y metodología):

El objetivo del trabajo es aplicar técnicas estadísticas y de ciencia de datos al análisis de datos obtenidos de redes sociales (por ejemplo, Telegram o TikTok). El foco principal es la identificación de comunidades de usuarios en función de sus interacciones, así como la detección y análisis de temas de conversación relevantes en dichas comunidades.

El estudiante trabajará con datos reales recolectados mediante APIs públicas o conjuntos de datos disponibles, y empleará herramientas de análisis de redes, procesamiento del lenguaje natural (PLN) y visualización de datos. El trabajo permitirá aplicar conceptos clave del Grado como el análisis exploratorio, el modelado estadístico y la minería de datos.

Metodología

1. Obtención de datos: Recogida de datos de redes sociales mediante APIs o conjuntos de datos abiertos. En caso necesario, se proporcionarán datos preprocesados para asegurar la viabilidad del trabajo.
2. Preprocesamiento de datos: Limpieza y transformación de los datos, detección de fuentes y mensajes relevantes, extracción de elementos clave (por ejemplo, hashtags, fechas, contenido textual, etc.).
3. Análisis de redes: Construcción de grafos de interacción entre usuarios (por ejemplo, seguimiento, reenvíos, etc.) e identificación de comunidades mediante software especializado (por ejemplo, Louvain en Gephi).
4. Análisis temático: Extracción de temas usando modelos de PLN como LDA o BERTopic, clasificación de mensajes por temática y estudio de la evolución temporal de los temas dentro de las comunidades detectadas.
5. Visualización: Desarrollo de gráficos explicativos (por ejemplo, series temporales, nubes de palabras, grafos de relaciones) mediante bibliotecas avanzadas de Python o R.
6. Conclusiones: Interpretación de los resultados, comparación entre comunidades, análisis de temas dominantes y discusión de posibles aplicaciones (monitorización de discursos, tendencias, campañas informativas, etc.).

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

Objetivos planteados:

Los objetivos de este trabajo se centran en aprender los fundamentos teóricos y prácticos de las siguientes tareas de la Ciencia de Datos:

- Obtener y procesar datos reales de redes sociales.
- Analizar redes para identificar comunidades.

- Aplicar modelos avanzados para la detección de temas de conversación.
- Visualizar datos complejos.
- Extraer conclusiones del análisis computacional de fenómenos sociales.

Bibliografía básica:

- Barabási, A.-L. (2022). Network Science. Web: <http://networksciencebook.com>
- Wickham, H., Cetinkaya-Rundel, M., & Grolemund, G. (2023). R for Data Science (2.ª ed.). O'Reilly Media. ISBN: 1492097403
- Van Atteveldt, W., Trilling, D., & Calderón, C. A. (2022). Computational Analysis of Communication. Wiley-Blackwell. ISBN: 1119680239
- Ramalho, L. (2022). Fluent Python (2.ª ed.). O'Reilly Media. ISBN: 1492056359
- Alammari, J., & Grootendorst, M. (2024). Hands-on Large Language Models. O'Reilly Media. ISBN: 9781098150969

Recomendaciones y orientaciones para el estudiante:

Conocimientos intermedios de Python o R

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: JUAN GÓMEZ ROMERO

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

Correo electrónico: jgomez@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: