



1. DATOS BÁSICOS DEL TFG:

Título: Desarrollo de un caso de estudio en ingeniería y ciencia de datos

Descripción general (resumen y metodología):

El trabajo de grado propuesto busca realizar un estudio exhaustivo sobre la manipulación y análisis de grandes volúmenes de datos, destacando el papel fundamental de los científicos e ingenieros de datos en este proceso. Se emplearán las últimas herramientas de la nube proporcionadas por Google Cloud Platform (GCP), como Dataflow, Pub/Sub y BigQuery, para llevar a cabo la ingeniería de datos. Además, se aplicarán técnicas de machine learning en este contexto para optimizar el procesamiento y la extracción de información útil de los datos. El objetivo final es desarrollar un dashboard interactivo con gráficos y estadísticas relevantes que faciliten la comprensión del problema y la toma de decisiones informadas.

La metodología seguida será la siguiente

1. Se recopilarán grandes volúmenes de datos en tiempo real de diversas fuentes relevantes para el caso de estudio.
2. Se realizará una limpieza y transformación de los datos utilizando herramientas de ingeniería de datos de GCP, como Dataflow, para prepararlos para su análisis.
3. Se realizará un análisis exploratorio de los datos para comprender su estructura y características, identificar patrones y anomalías, y obtener información preliminar sobre el problema en cuestión.
4. Se aplicarán algoritmos de machine learning para extraer conocimientos útiles de los datos y mejorar la eficiencia y precisión del proceso de análisis.
5. Se desarrollará un dashboard interactivo utilizando herramientas como Google Data Studio, que presentará gráficos y estadísticas relevantes derivadas del análisis de datos, proporcionando una visión clara y comprensible del problema.
6. Se evaluará la eficacia y la validez de los resultados obtenidos mediante el uso de métricas apropiadas y pruebas estadísticas.
7. Se presentarán conclusiones detalladas sobre el caso de estudio, destacando los hallazgos clave, las lecciones aprendidas y las posibles recomendaciones para futuras investigaciones o aplicaciones prácticas.

Este enfoque metodológico permitirá abordar de manera integral y rigurosa el análisis de datos en tiempo real y proporcionará una base sólida para resolver problemas complejos en el campo de la ingeniería y ciencia de datos.

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

Objetivos planteados:

1. Describir la tarea común de los científicos e ingenieros de datos.
2. Conocer y familiarizarse con las últimas herramientas cloud para ingeniería de datos ofrecidas por Google Cloud Platform (Dataflow, Pub/Sub, Bigquery...)
3. Identificar y plantear un caso práctico de interés para aplicar la metodología completa de trabajo
4. Realizar un dashboard, con gráficos y estadísticas relevantes que ayuden a entender el problema y a la toma de decisiones.
5. Aplicar técnicas machine learning para la generación de modelos que representen la información y permitan cierta accionabilidad sobre los datos.

Bibliografía básica:

1. Marz, N., & Warren, J. (2015). Big Data: Principles and Best Practices of Scalable Realtime Data Systems. Manning Publications.
2. Ng, A. (2018). Machine Learning Yearning. Deeplearning.ai.
3. Geewax, J. J. (2018). Google Cloud Platform in Action. Manning Publications.

Recomendaciones y orientaciones para el estudiante:

Realizar cursos de formación complementarios en Cloud Skill Boost:
<https://www.cloudskillsboost.google/>

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: ALBERTO LUIS FERNÁNDEZ HILARIO

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

Correo electrónico: alfh@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: Airam Ariza

Correo electrónico: airamariza@correo.ugr.es