

PROPUESTA DE TRABAJO FIN DE MÁSTER (Modalidad B)
Curso 2023-2024
MÁSTER EN TÉCNICAS ESTADÍSTICAS

Título	Desarrollo e implantación de un sistema de ML enteramente basado en tecnologías SQL de la nube
Nombre de la Empresa	SDG CONSULTING ESPAÑA, S.A.
Tutor/a en la empresa	ANDRES PADRONES
Tutor/a académico/a	ROSA M. CRUJEIRAS CASAIS
Descripción del trabajo a realizar	<p>Cada vez es más habitual que servicios de <i>data warehousing</i> en la nube incorporen capacidades de ML con el objetivo de permitir la creación y/o explotación de modelos algorítmicos mediante interfaces de fácil uso y conocidas por el analista de datos —i.e., SQL—, al mismo tiempo que se logre realizar todo el proceso allí donde estén los datos y, gracias a ello, conseguir grandes capacidades de procesamiento dada la alta escalabilidad de estas arquitecturas. En base a eso, este TFM propone la evaluación de todos estos aspectos en lo relativo a BigQuery ML, el componente con capacidades de ciencia de datos que Google Cloud ha habilitado en su producto de <i>data warehousing</i> en la nube. Así pues, el foco de este trabajo estará en dar cobertura al ciclo de vida completo de la ciencia de datos (análisis exploratorio, <i>feature engineering</i>, experimentación algorítmica, evaluación, <i>fine tuning</i>, operativización, explotación y monitorización) usando como componente central BigQuery ML, y enriqueciendo todo ello con un enfoque MLOps para garantizar la ágil iteración sobre la solución. Este trabajo se realizará sobre un caso de uso vehicular, con el fin último de demostrar la viabilidad de este tipo de enfoques a la hora de aportar valor a negocio desde una perspectiva de analítica avanzada en un corto tiempo, y utilizando tecnologías que posiblemente ya estén aplicando las compañías en relación a su estrategia de datos global.</p> <p>Beneficio esperado: Establecer el grado de madurez de las tecnologías de <i>data warehousing</i> en la nube con respecto a</p>

	sus capacidades de ML, validando la viabilidad de su aplicación para el desarrollo e implantación <i>end-to-end</i> de soluciones de ciencia de datos.
Recomendaciones	
Fechas de las practicas	A CONVENIENCIA DEL MASTER Y DEL ALUMNO.
Lugar de trabajo y horario	HORARIO SE ADAPTARÁ CONVENIENTEMENTE AL ALUMNO. LUGAR FLEXIBLE: PODRÁ SER REMOTO O EN NUESTRA OFICINA EN SANTIAGO DE COMPESTELA, A CONVENIENCIA DEL ALUMNO
Convenio	
Otras observaciones	