

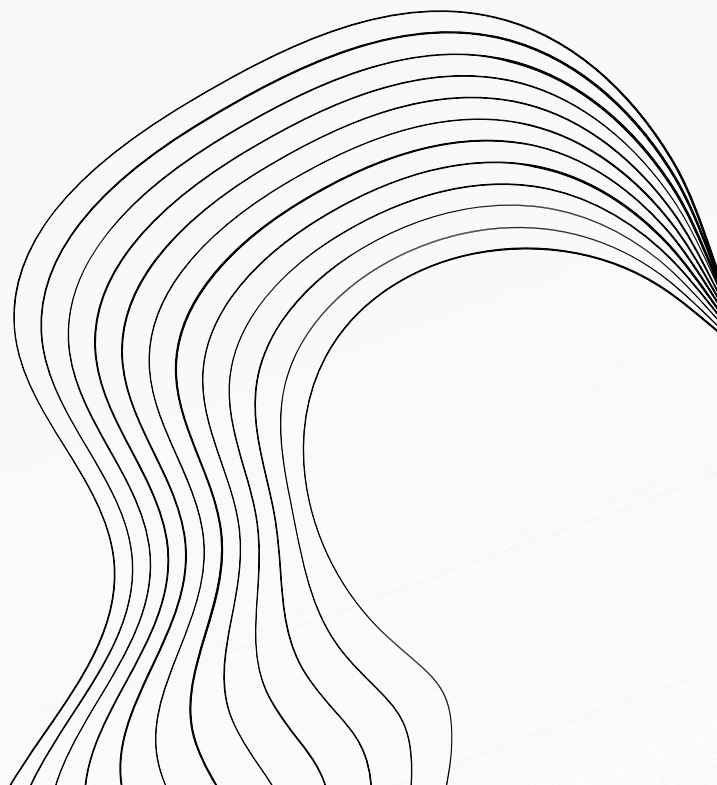
APRENDIZAJE AUTOMÁTICO PARA LA GESTIÓN DE DATOS MASIVOS

BY DR. PAULO LOPEZ MEYER

6/08/24

EVALUACION 3

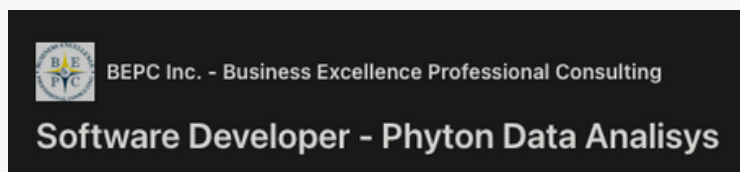
**ISAAC
MENCHACA**



1.¿Porqué Hadoop ya no es la herramienta más popular? ¿Cuáles son las herramientas modernas alternativas a Hadoop?

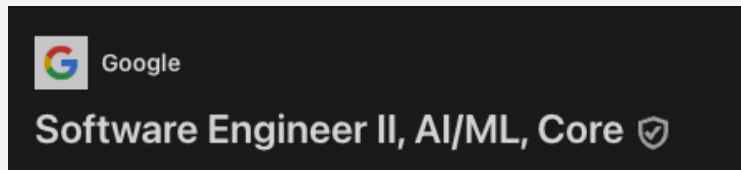
- Apache Hadoop, lanzado en 2006, prometía almacenar grandes volúmenes de datos con hardware básico y usar el esquema de lectura (Schema-on-Read) para simplificar el análisis. Sin embargo, esta ventaja resultó en problemas como la dificultad para gestionar y rastrear la calidad de los datos, y la acumulación de "pantanos de datos". Además, la complejidad de integrar múltiples motores de computación como Hive y Spark, y el enfoque erróneo en almacenar datos en lugar de utilizarlos efectivamente en aplicaciones empresariales, llevó a que Hadoop y los Data Lakes fueran menos efectivos y más difíciles de manejar.
- Apache Spark es la alternativa, y en muchos aspectos el sucesor, de Apache Hadoop. Fue creado para abordar las deficiencias de Hadoop y lo hace increíblemente bien. Por ejemplo, puede procesar datos por lotes y datos en tiempo real, y funciona 100 veces más rápido que MapReduce.

3.Buscar 2 trabajos en México donde se mencionen al menos 2 temas de los vistos en clase (LinkedIn, Glassdoor, etc.).



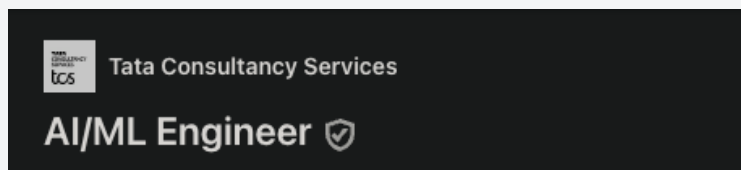
Required Skills

- Data processing Python libraries (Pandas, NumPy)
- SQL Databases (MSSQL, MySQL or PostgreSQL)
- Machine Learning libraries (TensorFlow, PyTorch, Scikit-learn)
-



Required Skills

- 1 year of experience with machine learning algorithms and tools (e.g., TensorFlow), artificial intelligence, deep learning, or natural language processing.
- Experience in ML model coding languages (e.g., Python).
- Master's degree or PhD in Computer Science or related technical field.



Required Skills

- Proficiency in programming languages such as Python, Java, C++, C, R, and JavaScript.
- Strong understanding of machine learning frameworks (e.g., TensorFlow, PyTorch) and libraries (e.g., scikit-learn, Keras).