Mohamed Bilel Hasni

Senior Data Engineer

Country of residence : Luxembourg
Mobile : +352-661702111
Email : hasnimedbilel@gmail.com

Portfolio   https://hasnimedbilel.github.io/

Professional Summary

Cloud-savvy Data Engineer with 5+ years of experience designing scalable data platforms, ETL frameworks, and real-time data integration solutions. Expert in Microsoft Azure, Databricks, and Apache Spark. Currently supporting KPMG through SFEIR Luxembourg on high-impact cloud data projects. Strong foundation in data modeling, pipeline optimization, and automation.

Core Competencies

Cloud Platforms	○ Azure (Data Factory, Synapse, Data Lake, Azure SQL, Logic Apps), Databricks, Azure DevOps
Data Engineering	○ ETL/ELT design and orchestration using Azure Synapse Pipelines, Dataflows and Notebooks ○ Delta Lake, Lakehouse Architecture, Bronze/Silver/Gold data modeling ○ Stream and batch processing with Spark, PySpark ○ API-based ingestion (e.g., SAP SuccessFactors) and large dataset handling
Languages & Tools	○ Python, SQL, Java, Apache Hive, Docker, Git, CI/CD
Visualization	○ Power BI, Tableau
ML & NLP (Supporting Skillset)	○ Scikit-learn, PyTorch, HuggingFace

Professional Experience

Jan 2023– Ongoing   **Data Engineer**, *SFEIR, LUXEMBOURG*.
- ○ Big Data analytics/processing, design and architecture of ETL pipelines in Azure to support data ingestion&integration of operational data sources.
- ○ Big Data Engineering, Azure Synapse Analytics, Databricks, ETL, Spark.

Jan 2020– December 2022   **Data Scientist**, *PUTNAM PHMR*.
- ○ Application of AI in the healthcare industry: Clustering, unsupervised learning, Natural Language processing, Image processing ...
- ○ Data Engineering, Data Warehousing, ETL, PySpark, SQL, Dask, Pandas.
- ○ Data visualization & Dashboarding.

Jan 2019–Oct 2019   **Data Science Intern**, *SMARTTEK SERVICES, DUBAI*.
- ○ Developed and deployed a machine learning solution to detect Simbox fraud using PySpark, Hive, and Rapidminer on Hortonworks Data Platform (HDP).
- ○ Processed large telco datasets, including CDR extraction, data transformation, and warehouse creation.
- ○ Monitored and managed HDP clusters using Apache Ambari for performance optimization.

Education

2016–2019   **National Engineering School of Sousse, Tunisia**,
Computer-Science Engineering degree.

2014–2016   **Preparatory Institute of engineering studies el Manar, Tunisia**,
Mathematics & Physics.

Certifications

Microsoft   o   Microsoft DP-203 : Azure Data Engineer Associate [Link]
Azure   o   Microsoft AZ-900 : Azure Fundamentals [Link]

Databricks   o   Databricks : Certified Data Engineer Associate [Link]

Key Projects

CDH
Platform
Data Engineer, *KPMG.*
- Developed enterprise-grade, automated data pipelines in Azure Synapse Analytics and Databricks for reporting and analytics.
- Built CI/CD for automated deployment of data workflows via Azure DevOps.
- Designed lakehouse architecture to support structured and unstructured data.
- Collaborated on governance, lineage, and access management using Azure Purview.
- Leveraged SAP SuccessFactors (SF) and SAP HANA for handling large-scale data to support internal operations, using Synapse Pipelines, Dataflows, and Spark for efficient data processing and integration.

TAK
Treatment
Clustering
(ISPOR)
PySpark, Flask, PostgreSQL, *ISPOR 2022 Conference.*
- Developed an ML solution for clustering treatment sequences.
- Used unsupervised learning techniques with image processing for treatment sequence analysis.

InstaPaper
PyTorch, CRAFT/CRNN, Flask, *World Bank Project.*
- Developed an OCR-based pipeline for extracting data from audit report images.
- Implemented deep learning models like CRAFT for text detection and CRNN for text recognition.

Publications

2022   TAK (Treatment Sequences Analysis Through K-Clustering), [ispor]

2022   Could Artificial Intelligence Support Prediction of Reimbursement Decisions in Scotland? A Pilot Project, [ispor]

2022   Promises of AI-Assisted Patient Monitoring Methods, [ispor]