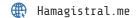
# HAMZA EL BELGHITI

# **Junior Data Engineer**









📈 hamza.lbelghiti@gmail.com



(+212) 660081086



# **WORK EXPERIENCE**

# **FULL STACK DEVELOPER** | July - September 2023

**♀** 3D SMART FACTORY, Mohammedia | **♀** GITHUB **ℰ** DEMO

- Deployed a 3D teeth segmenting deep learning model using a REST API on the AWS Cloud.
- Reduced deployment cost from \$85 per month fixed to \$1.3 per 1000 requests.
- Created a user interface with 3D libraries and deployed the web application with Vercel.

Technologies: Vite, React, Tailwindcss, Three, Vtkjs, Firebase, Pytroch, FastAPI, Docker, Vercel AWS ECR + Lambda + API Gateway

#### JUNIOR DATA SCIENTIST | May - June 2023

♥ OMDENA ÎLE-DE-FRANCE, Remote | • GITHUB Ø DEMO

- Collaborated with a team of 10 international people to develop a conversational chatbot that makes it easier to find alternative public transport during strikes in the Paris region.
- Web scraping of the RATP site for traffic status with an ETL pipeline on AWS.
- Developed an algorithm two times faster than the existing solution using the Navitia API.

Technologies: Streamlit, Folium, Beautifulsoup, AWS S3 + Lambda + Cloudwatch, Navitia API

# **PROJECTS**

# DATA ANALYSIS OF AN ONLINE RETAIL BUSINESS | 2023

FINAL PROJECT OF DE ZOOCAMP | GITHUB & DEMO

- Designed and implemented an ELT data pipeline using Airflow to extract insights from data.
- Ingested data into a storage Bucket, transformed it with dbt, and stored it in BigQuery.
- Created an interactive dashboard that helped identify the total revenue by country, track monthly revenue, and pin down the best-selling products.

Technologies: Python, Terraform, Airflow, dbt, Soda, GCP Storage + BigQuery, Docker, Looker

# MIGRATION OF AN ON-PERMISE DATABASE TO AZURE CLOUD | 2023 DATA ENGINEERING PROJECT | GITHUB

- Connected an on-premise SQL Server database to the Azure Cloud with Microsoft IR.
- Developed a batch pipeline that ingests data with Data Factory, transforms it with Spark and Databricks, then uploads it to a data warehouse, Azure Synapse Analytics.
- Generated an interactive report using PowerBI to help analyze the company's sales.

Technologies: SQL Server, Azure Data Factory + Lake + Synapse, Spark, Databricks, PowerBI

#### ANALYZING GLASSDOOR DATA ENGINEER JOB POSTINGS | 2023

DATA SCIENCE & ENGINEERING PROJECT | GITHUB & DEMO

- Weekly web scraping of data engineer job postings on Glassdoor and storing them into S3.
- Designed a pipeline with Mage that extracts and transforms data from the AWS S3 Bucket.
- Predicting data engineer salary with 86% accuracy using Random Forest Regressor.
- Created a Streamlit app to visualize the most in-demand tools and salaries of data engineers.

Technologies: Python, Selenium, AWS S3, Boto3, Mage, Scikit-learn, Altair Viz, Streamlit

# **EDUCATION**

# MASTERS IN DATA SCIENCE AND ENGINEERING 2022-2024

**♀** Faculty of Sciences of Rabat (FSR), Rabat

Courses: Big Data, Business Intelligence, Datamining, Data Science, Data Base Administration

## **SKILLS**

#### **PROGRAMMING**

Python • Java • C • R

#### **DATABASES**

Oracle • SQL Server • MySQL • Pinecone PostgreSQL • MongoDB • Snowflake

#### **DATA ENGINEERING**

Selenium • Terraform • Soda • Airflow • Mage • Databricks • Spark • Kafka • dbt Dataflow • PowerBI • Looker Studio

#### WEB DEVELOPMENT

HTML/CSS • Java/Typescript • Nextjs • FastAPI • Django • Flask • Streamlit

#### **CLOUD**

AWS • Google Cloud • Azure • Firebase

#### **TOOLS**

Linux • Docker • Git • Gitlab (CI/CD) • Jira

## **LANGUAGES**

Arabic • English (C2) • French (C2)

# **CERTIFICATES**

# **DATA ENGINEER**

Delivered by **DataCamp** Septembre 2023

#### DATA SCIENTIST WITH PYTHON

Delivred by **DataCamp** Décembre 2022

#### **MACHINE LEARNING - PYTHON**

Delivered by FreeCodeCamp Avril 2022

## **CS50 WEB PROGRAMMING**

Delivered by Harvard Février 2022

# **EXTRACURRICULAR**

# STREAMLIT FOR EDUCATION

Student Ambassador Leader | 2023-2024

**INNOVASUS - FSR CLUB** Member | 2023-2024

#### HOBBIES



**I** Chess

Blogging



Hackathons



E-sports