



Zineb EL BACHA

+212657850067 | zinebbacha2004@gmail.com | linkedin.com/in/zineb-el-bacha
github.com/zinebelbacha

Profile

Fourth-year data science student with strong expertise in AI and Machine Learning. Passionate about emerging technologies and artificial intelligence, I am highly motivated, eager to learn, and driven to make meaningful contributions. Currently seeking an End-of-Year Internship in data science, where I can apply my skills, gain hands-on experience, and contribute to innovative and impactful projects.

Education

National Institute of Posts and Telecommunications - INPT	Sept 2023 – Present
Data Engineering	
LM6E (formerly Lydex)	Sept 2021 – June 2023
Preparatory Classes, Technology and Industrial Sciences (TSI)	
Jaafar Elfassi Elfihri Highschool	Sept 2018 – June 2021
Baccalaureate in Electrical Sciences and Technologies (STE), with High Honors	

Projects

AskINPT	See project
<ul style="list-style-type: none">Intelligent and interactive assistant designed for the INPT website, providing visitors with precise and instant answers.Utilized RAG architecture to extract and generate relevant responses from documents such as PDFs or text files. Tools: Retrieval-Augmented Generation, LLM, LangChain, Pinecone.	
Reddit Data Streaming Pipeline	See project
<ul style="list-style-type: none">Streamed real-time posts from the Reddit API and published them to a kafka topic.Processed and transformed the incoming data using spark streaming before storing it in a database.Orchestrated the pipeline with Airflow and containerized the services using Docker and Docker Compose. Tools: Reddit API, Apache Kafka, Apache Spark, Docker, Apache Airflow, PostgreSQL.	
AI Medical Assistant	See project
<ul style="list-style-type: none">AI model capable of classifying pulmonary X-rays.Collected and preprocessed pulmonary X-ray images from multiple sources for model training.Used a pre-trained convolutional neural network, DenseNet121, to classify pulmonary X-rays. Tools: DenseNet121, Python, NumPy, Seaborn	
Car Data Scraper and Analyzer	See project
<ul style="list-style-type: none">Used web scraping techniques to extract car-related data from Edmunds.com, a dynamically generated website.Handled JavaScript-rendered content and implemented anti-detection measures using user agents and request header.Preprocessed and cleaned the extracted data, and visualized key insights using Power BI. Tools: Selenium, Pandas, Power BI	
CNE Reader	See project
<ul style="list-style-type: none">Tool for automatic extraction of student information from their student cards.Developed a model combining object detection and OCR. Tools: OCR, PyTesseract, TensorFlow 2, SSD-MobileNet	

Certifications:

Coursera Supervised Machine Learning: Regression and Classification

Languages

English: Read, spoken, written
French: Read, spoken, written

Arabic: Native language
Spanish: Beginner level

Technologies

Languages: Python, Java, Scala, C, SQL, Bash

Databases and Big Data: MySQL, PostgreSQL, MongoDB, HBase, Hadoop, Spark, Kafka

Skills: Machine Learning, Deep Learning, Data Engineering, Cloud Computing, NoSQL, Data Analysis, Data visualization

Interests and Volunteering

Interests: Art, Painting, Drawing, Reading, Personal Development

Volunteering: Member at Social Affairs Club INPT, Member at FGE Club INPT