

Souheil Bichiou

☎ (216) 52720115 | ✉ souheil.bichiou@ensi-uma.tn | 🔗 linkedin.com/in/souheil | 🐙 github.com/souheil

Education

National School of Computer Science

Sep 2022 – Jun 2025

Master in Computer Science

Tunisia

Preparatory Institute for Engineering Studies of Nabeul

Sep 2020 – Jun 2022

Engineering Preparatory Program — Mathematics, Physics, and Engineering Sciences

Tunisia

Experience

Cloudsquid

Feb 2025 – Jun 2025

Data Software Engineer Intern

Germany, Berlin

- Crafted an evaluation platform using **ReactJS** and **FastAPI**, achieving a precision score of **0.937** on data processed through the Cloudsquid pipeline, leveraging both collected and custom-created datasets
- Leveraged **Large Language Models (LLMs)** and **Optical Character Recognition (OCR)** to develop a solution to visualize **bounding boxes** of all extracted data, thus handling documents with **200+** pages
- Integrated the solution into the Cloudsquid platform using **Golang**, **GoProfiles**, and **ReactJS**, boosting data traceability by **95%**

SkillMapper

Sep. 2024 – Nov 2024

Data Scientist

France, Remote

- Engineered a **sentiment analysis system** for the SkillMapper platform to score courses and analyze reviews, helping users make informed enrollment decisions and increasing enrollment efficiency by **30%**
- Benchmarked **ML models** and **LLMs** for course review scoring, achieving **0.92** accuracy with **DistilBERT**
- Provided users with the ability to assess a course's score before enrolling, reducing course **dropouts** by **25%**
- Utilized **BERTopic** to extract dominant themes from **5,000+** course reviews, offering insights into the most discussed topics and user concerns

Elyadata

Jun 2024 – Sep 2024

Data Scientist Intern

Tunisia, Tunis

- Fine-tuned **T5** model for Arabic summarization and question answering, reducing manual review time by **65%**
- Applied **LoRA (PEFT)** and **model quantization** to fine-tune the model, updating only **0.47%** of parameters and reducing memory and computational requirements
- Leveraged **Weights & Biases** to track and validate experiments, reducing result interpretation time by **70%**
- Achieved an evaluation **loss** of **0.16** after **3 epochs** and a **ROUGE-1** score of **0.72**

Projects

Resume Skill Extractor And Scorer | [Source Code](#)

NER | Scikit-learn | Flask | Keras

- Enhanced resume screening process by using **BERT-base-cased** for **Named Entity Recognition (NER)** to extract skills, experiences and other key information achieving a **92%** accuracy in identifying relevant entities
- Scored CVs based on their skills and experiences using a **Sequential Model** achieving a training **loss** of **0.14**
- Enhanced candidate selection, boosting accuracy by **35%** and cutting manual screening time by **75%**.

VideoCode Extractor | [Source Code](#)

FastAPI | ReactJS | OpenCV | LangChain

- Designed a user-friendly interface with **FastAPI** backend and **ReactJS** frontend, enabling users to extract and retrieve code from any frame of local or YouTube videos, reducing manual code retrieval time by **80%**
- Implemented a frame-by-frame analysis system with **OpenCV**, boosting video processing speed by **50%**
- Integrated a chatbot using **LangChain**, enabling video code queries and reducing search time by **60%**.

Technical Skills

Languages: Golang, Python, C/C++, Java, SQL (Postgres), JavaScript, HTML/CSS

Frameworks: ReactJS, Node.js, Flask, FastAPI

Developer Tools: Git, Docker, AWS, Postman, BitBucket, Kubernetes, Apache Spark, Apache Airflow

Languages

French (B2) | **English (B2)** | **Arabic (Native)**