

# ARIJ ZAHRA SOULA

✉ [arijzahra.soula@gmail.com](mailto:arijzahra.soula@gmail.com)  [linkedin.com/in/arijzahasoula](https://linkedin.com/in/arijzahasoula)  [github.com/Arijaa](https://github.com/Arijaa)

## Summary

---

Computer Science Engineering graduate , specialized in Data Science. Experienced in NLP, speech, recommendation systems, healthcare AI, computer vision, and applied ML/DL. Skilled in building scalable pipelines, embeddings, RAG methods, and deep learning models. Passionate about advancing research in AI across multiple domains

## Education

---

### Higher school of Engineering and Technology of Tunisia

Sep. 2021 – June 2025

*Computer science engineering degree with a Data Science specialization ,3.5 GPA*

*Ariana, Tunisia*

Relevant courses : Machine learning , Deep learning , Big Data , Mathematics

### Institute of preparatory engineering studies of Tunis

Sep. 2019 – Aug. 2021

*Preparatory studies engineering degree*

*Tunis, Tunisia*

### Bourguiba Pioneer Highschool of Tunis -LPBT

Sep. 2015 – June 2019

*Baccalauréat in experimental sciences*

*Tunis, Tunisia*

## Experience

---

### Polytechnique Montréal

May 2024 – December 2024

*AI Researcher Intern – Final engineering project*

*Montréal, Canada*

- Conducted research on code clone detection using advanced natural language processing (NLP) techniques, focusing on detecting Type-4 (semantic) clones.
- Implemented and optimized sentence transformer models to generate vector embeddings for code, improving the accuracy of similarity detection in large code datasets.
- Authored a comprehensive literature review on state-of-the-art methods in code clone detection, focusing on NLP-based approaches and embedding models to offer insights and direction for future research in the field.
- **Key technologies :Python, PyTorch, Sentence Transformers (Hugging Face), Qdrant**

### IVA AI Startup

July 2023 – September 2023

*Artificial intelligence Engineer Intern*

*Talinn, Estonia*

- Implemented state-of-the-art text-to-speech (TTS) and speech-to-text (STT) models, resulting in a significant improvement in call transcript conversion accuracy.
- Trained and fine-tuned the TTS model to produce natural-sounding speech from written call transcripts, enhancing overall customer experience.
- Developed and optimized the STT model to accurately transcribe audio call recordings, resulting in improved data accessibility and analysis.
- **Key technologies :Python, NLP , TTS , STT , Transformers , Bert, NLTK ,Microsoft Azure**

### ESPRIT school of engineering

June 2023 – September 2023

*Data engineer Intern*


*Tunis, Tunisia*

- Built ETL pipelines and data workflows to analyze student performance, integrating multiple sources and performing data cleaning, transformation, and normalization using Python (Pandas).
- Designed and implemented a PostgreSQL database for structured storage, querying, and reporting.
- Developed data visualizations and dashboards using Matplotlib and Power BI to identify trends, support data-driven decision making, and enhance analytics capabilities
- **Key technologies :Python (Pandas, NumPy), PostgreSQL, Matplotlib, Power BI**

## Projects

---


### **Risk management Conceptual graph Based Recommendation System** | *NLP, GNN, Spacy, Knowledge graph* **Nov. 2023**

- Extracted actionable insights from textual data using NLTK, resulting in an increase in RM issues solving efficiency.
- Developed a Conceptual Graph-Based Recommendation System for Risk management issues monitoring, reducing decision-making time.
- Utilized advanced Machine Learning models, including Graph Neural Networks (GNNs) with TensorFlow, to achieve an improvement in recommendation accuracy.
-  [Github link](#).


### **Student Job market DevOps project** | *Jenkins,Docker,SonarQube ,Grafana,Prometheus* **Sep.2023 - Nov. 2023**

- Set up Jenkins Environment, executed Maven Clean Compile, and conducted testing with JUnit and Mockito
- Ensured source code quality by performing checks with SonarQube, and created/hosted artifacts in Nexus.
- Published images on DockerHub and implemented monitoring using Prometheus and Grafana.

### **MindMatch: Career center platform NLP Project** | *Python, Pytorch, Sickitlearn, Power BI* **Jan. 2023 – May 2023**

- Retrieving job descriptions and CVs from LinkedIn and job offers.
- Use of data science methods for detailed data analysis.
- Development of a job recommendation system and development of a chatbot.
-  [Github link](#).

### **Chronic Kidney Disease Classification machine learning project** | *KNN, NB, SVM,Random Forest,XGboost* **Nov. 2022**

- Achieved an increase in chronic kidney disease prediction accuracy through data cleaning, visualization, and preparation.
- Optimized model performance by conducting an in-depth evaluation of hyperparameters, resulting in improved diagnostic capabilities and patient care
-  [GitHub link](#)

## Technical Skills

---

- **Programming:** Python, R, SQL
- **AI/ML/DL:** CNN, RNN, LSTM, Transformers, RAG, XGBoost
- **NLP & Speech:** Sentence Transformers, SpaCy, embeddings, retrieval pipelines
- **Frameworks:** TensorFlow, PyTorch, Scikit-learn, Pandas
- **Databases:** PostgreSQL, MongoDB

## Languages

---

**English:**Fluent (written and spoken)

**French:**Fluent (written and spoken)

**Arabic:**Native language

## Certifications and Achievements:

---

**Second place in the hackathon: TICAD 8 Automotive AI challenge**

**Nvidia - Applications of AI for Anomaly Detection certificate**

**Deeplearning.AI- Natural Language Processing specialisation certificate**

## Leadership / Extracurricular

---

### **IEEE Tunisia section**

**January 2023 – January 2024**

*Section Student representative*

*IEEE Tunisia Section SAC*

- Representing the IEEE Tunisia section at national and international conferences, highlighting section activities and achievements and leading meetings with chairs

### **IEEE Computational Intelligence society Student branch chapter**

**May 2022 – May 2023**

*Founder and Chairperson*

*IEEE ESPRIT SB*

- Conducted 4+ Introductions to AI workshops, and organized an AI Hackathon for society members, enhancing their technical proficiency