Mahdi Zarour

mahdi.zarour.main@gmail.com
514-999-3492 · Canada
Linkedin.com/in/MahdiZarour · Website

SUMMARY

Machine Learning Engineer · Data Scientist · Backend Developer
- **Technically Strong with a focus on Mathematics and Algorithms**.
- **Effective collaborator and Listener** with experience working in cross-functional teams. Proficient in understanding requirements, conducting comprehensive data analysis, and providing insightful recommendations to help make decisions.
- **Proactive and helpful team player** who contributes and shares his ideas, whether when asked or when a need is perceived. Always ready to provide assistance and make meaningful contributions to the team.

EDUCATION

Université de Montréal   Montréal, QC
Master of Science, Computer Science   2020-2023
- **Coursework**: Data Science · Fundamentals of Machine Learning · Probabilistic Graphical Models · Representation Learning · Theory of Deep Learning.
- **GPA**: 3.7.
- **Scholarships**: International Students Scholarship worth $27,000 CAD.

Université Oran 1 Ahmed Ben Bella (IGMO)   Oran, Algeria
Licence (Bachelor of Science), Mathematics & Computer Science   2015-2018
- Top 5% of the program in GPA.

EXPERIENCE

Machine Learning Developer (Self Employed)   Montréal, QC
Skylar   Jan/2023 - Now

Skylar is a startup enterprise with the main mission of pushing the boundaries of AI in Industry. They deliver tailored AI powered models as products. During my time there, I had the chance to work on various projects and manipulate many types of data. I worked on Marketing Campaigns Optimization with Machine Learning, Object Detection with Deep Learning, and Chat-based Question Answering with LLM.

Q&A with LLM (Mars/2023 - Now):
- Built with other team members a tailored chat bot for a dental clinic, utilizing Large Language Models (**LLM**) and **Python**. This involved formulating user questions as prompts for **OpenAI GPT-3** and enriching them with pertinent context coming from a 400-page comprehensive book stored as embeddings in a **vector database**. This process included comparing the user's prompt against the database and extracting the most relevant segments to add to the prompt using **LangChain**. Integrated a real-time speech-to-speech interface via the Microsoft **Azure** API to enable users to verbally pose questions and receive spoken responses.
- Collaborated within a **cross-functional** team as part of an **Agile** Framework to ensure continual client engagement through regular progress demonstrations and feedback integration.
- **Skills: Agile, CI/CD, Big Data, AWS, EMR, SageMaker, Lambda, S3, Python, R, PyTorch, Spark, Hadoop, XGBoost, Supervised Learning, Data Cleaning, Data Preparation, Object Detection, Computer Vision, NLP, Model Training, Deep Learning, LangChain, OpenAi, Vector Database, Microsoft Azure, Embeddings, Jupyter Notebook.**

Data Scientist / Unity Developer Intern   Montréal, QC
Beam Me Up Inc.   Jan/2022 - Dec/2022
- Implemented a python library (with **NumPy, Pandas,** and **Matplotlib**) to process raw Eye Tracking and EEG data by applying the latest feature extraction algorithms to data captured from consumer grade VR and EEG devices, and to visualize the results.
- Developed a virtual environment in **Unity** using **C#** language and **OOP** concepts and successfully conducted experiments on 31 individuals at BMU labs.
- Facilitated **Agile** product development with BMU team, and implemented **CI/CD** practices to increase deployment release frequency and decrease deployment time.
- Developed a state-of-the-art distraction detection system in terms of accuracy (84%, best among those present in the literature) by combining features from EEG and Eye Tracking and applying a machine learning classification algorithm, SVM with **Scikit-learn**.
- Developed the first (in the literature) numerical indicator of the levels of distraction by combining linear dimensionality reduction techniques (Linear Discriminant Analysis) and interpolation for effective Supervised Learning.

- **Skills: Agile, CI/CD, Data Analysis, Data Visualization, Data Acquisition, Data Cleaning, Data Preparation, C#, Unity, OOP, Python, R, EEG, Eye Tracking, Scikit-learn, PyTorch, Numpy, Pandas, Matplotlib, Virtual Reality, Supervised Learning, Unsupervised Learning, Signal Processing, Jupyter Notebook.**

Data Scientist — Oran, Algeria
Betobex — Jun/2018 - Jan/2020

- Managed and optimized part of the database effectively using **PostgreSQL**, to ensure data security.
- Leveraged **Business Intelligence** and **Dashboarding** tool **Tableau** to perform data visualization and data analysis, extracting insights to inform business strategies at the company.
- Assisted the senior members in optimizing inventory management by employing **NumPy**, **Pandas**, **Scikit-learn** and **XGBoost** for Data Processing and Supervised Learning.
- Collaborated in bi-weekly sprints and daily stand-ups as part of an **Agile** framework, to enhance team motivation and communication.
- Effectively communicated complex data insights to non-technical management teams, ensuring clear interpretation and understanding of mathematical results, while backing my analysis with **Matplotlib** plots.
- **Skills: Python, R, Scikit-learn, Matplotlib, Pandas, XGBoost, Tableau, Business Intelligence, Agile, PostgreSQL, Data Analysis, Data Visualization, Data Acquisition, Data Cleaning, Data Preparation, Machine Learning, Supervised Learning, Jupyter Notebook.**

Software Developer Intern — Oran, ALgeria
Betobex — Jan/2018 - Jun/2018

- Continuously identified and resolved **bugs** in the Inventory Management System, to improve system **stability, security** and **performance**.
- Developed complex **SQL cascading** queries to automate data management processes and reduce manual workload.
- Transformed methods of the Inventory Management System into a **SOAP web service** using **Spring** Framework, to enable efficient **XML** based data exchange with external applications and code reusability.
- Utilized **Maven** to handle packages easily and efficiently.
- **Skills: Spring, Maven, Java, Web Services, SOAP, Java, XML, OOP, SQL, MySQL, phpMyAdmin.**

Software Developer Intern — Oran, Algeria
Betobex — May/2017 - Sep/2017

- **Collaborated** within a **team**, and leveraged the **Model-View-Controller** (MVC) architecture to build an Inventory Management System in **Java**, using **OOP** concepts including **Inheritance** and **Polymorphism** with **Overloading** and **Overriding**.
- Implemented **exception handling** in Java, to manage runtime errors and ensure robustness.
- Employed **MySQL** and **PhpMyAdmin** for database management, to store, retrieve, and update data.
- Developed the frontend using **JavaFX**, and focused on intuitivity and simplicity.
- Utilized **Maven** to handle packages easily and efficiently.
- **Skills: Java, Maven, OOP, SQL, MySQL, phpMyAdmin, JavaFX, Full-Stack, Exceptions, Collaboration.**

SKILLS

- Proficient in recent Machine Learning algorithms by implementing from scratch SVM, Logistic Regression, Naive Bayes, XGBoost, and Decision Trees using only Numpy and Scikit-learn.
- Proficient in recent Deep Learning architectures by implementing from scratch MLP, CNN, RNN, and LSTM using only Numpy.
- Proficient in Pytorch library after implementing in Pytorch DL architectures including CNN, Transformers, Vision Transformers (ViT), Variational Autoencoders (VAE), and Generative Adversarial Networks (GAN).
- Experience with other Deep Learning Libraries Tensorflow and Keras.
- Experience in applying machine learning to Computer Vision for Image Classification and Object Detection.
- Experience in applying machine learning to Natural Language Processing (NLP) for Text Classification with NLTK and SpaCy, Language Modeling, Sentiment Analysis, and Topic Modeling..
- Competent with Microsoft Azure and Amazon Web Services (AWS) SageMaker, Lambda, and S3.
- Familiar with Google Cloud Platform (GCP).
- Competent with Big Data tools Apache Spark (Pyspark), Hadoop, and Hive.
- Proficient with Data Manipulation, Visualization, and Analysis tools Pandas, Matplotlib, Seaborn, and Tableau..
- Familiar with Git for code collaboration and version control.
- Strong Foundation in Statistics, Probability, Calculus and Linear Algebra.
- Proficient with OOP, Algorithms, and Data Structures.
- Proficient in Programming with Java and Python.
- Experienced in Programming with C, C++, and R.
- Experience in Generative AI with ChatGPT, Llama, Stable Diffusion and GANs.

SOFT SKILLS

- Strong problem-solving and analytical skills.

- Excellent communication and collaboration skills.
- Proficient in English, fluent in French and Arabic.