

SUMMARY

Machine Learning Engineer · Data Scientist · Data Engineer · Software Engineer

- **With 5 years practicing in Data Science and Machine Learning**, dedicated to continuous learning and staying up to date with the latest advancements in machine learning, data science, and MLOps. Enjoy tackling new problems without being afraid to explore solutions directly from research papers to deliver the most effective results.
- **Technically Strong with a focus on Mathematics, Statistics, and Probability.**
- **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

- **Grade:** A (Top 5% in the program).

EXPERIENCE

Machine Learning Engineer / Software Engineer

Montréal, QC

Skylar

Jan/2023 - Now

Q&A with LLMs and RAG:

- Built a tailored chat bot app for a startup, with the large language model **ChatGPT** and **Flask**.
- Used OpenAI **Embeddings** and **Vector Databases** to augment ChatGPT prompts with relevant context (retrieval-augmented generation).
- 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.

Discord Digital Store:

- Designed an online store on the Discord Platform using UML and implemented it with **Python** and **Django**.
- Employed **OOP** concepts including **Inheritance** and **Polymorphism** with **Overloading** and **Overriding**.
- Managed data effectively using **PostgreSQL (CRUD)** and **SQLAlchemy**, including users, transactions, orders, and other information.
- Successfully deployed the store on **Heroku**, then actively monitored the application, to identify and rectify flaws and **bugs**.
- **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.

Beam Me Up Inc.

Montréal, QC

Data Scientist / Unity Developer Intern

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 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

Betobex

Oran, Algeria

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, Data Analysis, Data Acquisition, Data Cleaning, Data Preparation, Machine Learning, Supervised Learning, Jupyter Notebook.**

PROJECTS

- **Object Detection:** Implementation in **Pytorch** of the famous "You Only Look Once" CNN based object detection algorithm used in real time computer vision systems.
- **Fraud Detection:** Detection of frauds using Machine Learning with Apache Spark for distributed processing and XGBoost.
- **Tutorial System Management:** A Spring Boot, Angular, and PostgreSQL project to implement CRUD operations.
- **DentalGPT :** Utilizing Retrieval Augmented Generation (**RAG**) to answer questions about dental aligners with ChatGPT, everything with a beautiful Streamlit UI.
- **PicAI:** Django based web application for basic image processing using Pandas, Sklearn, and Numpy.
- **Semi-Supervised Image Classification:** Combining Unsupervised Representation Learning and only 100 labels to maximize classification accuracy on MNIST Dataset.
- **River Flow and Water Level Forecasting for HydroQC:** Using Time Series models such as Prophet and Arima, ranked second at HackQC 2022.
- **Plant Disease Detection:** Using Deep Learning CNN model, won first place at CodeML 2022 at Polytechnique Montreal.
- **Movie recommendation system:** with NLP Sequence BERT based on the user's movie history.

SOFT SKILLS

- Strong problem-solving and analytical skills.
- Excellent communication and collaboration skills.
- Proficient in English, fluent in French and Arabic.