

Osebom Joshua Onyeagwu

oj.osebom@gmail.com • 905-324-2825 • [LinkedIn](#) • [GitHub](#)

TECHNICAL SKILLS

Languages: Python, SQL, JavaScript, PHP, Bash

Tools: Pandas, Spark, Flask, FastAPI, Azure, Docker, scikit-learn, TensorFlow, spaCy, LangChain, HuggingFace

EXPERIENCE

Consultant, Business Effectiveness

August 2024 – Present

Canadian Imperial Bank of Commerce (CIBC)

Toronto, Ontario

- Designed **data pipelines** with **Databricks**, **SQL**, **SAS**, and **Azure Data Studio** to monitor user journeys
- Utilized **PySpark** to transform large clickstream datasets into **schemas** for **actionable business insights**
- Built **SQL** datasets from **backend API** logs in **Splunk**, tracking lending product adoption for **100K+** users
- Leveraged **PySpark** to uncover **70%** user drop-off in meeting application forms, due to poor UI design

Backend Engineering Intern

May 2024 – August 2024

Vector Institute

Toronto, Ontario

- Deployed **Flask API** on **GCP** with **user recommendation** & **semantic search** endpoints serving **+450 users**
- Integrated **Qdrant** vector database, boosting **search relevance** by **30%** and reducing processing time by **40%**
- Built **data pipeline** using **cron jobs**, **Cohere** & **WordPress API** to update database with **+200 resources** weekly
- Deployed **WordPress** plugin with **PHP** to easily push AI enchantments to company's main website

Data Scientist

January 2023 – December 2023

McMaster Manufacturing Research Institute (MMRI)

Hamilton, Ontario

- Built **ETL data pipelines** with **Pandas** to **automate** healthy analytics of **40+** machines at Honda
- Deployed **web app** with **PostgreSQL** & **Flask** to **automate SQL queries** and **insights** for **20+ ad-hoc projects**
- Designed schema & API to serve business requirements and improve inventory tracking by **over 50%**
- Productized **scikit-learn** machine learning models for predictive maintenance, achieving **30% cost savings**

Data Analyst

May 2022 – December 2022

McMaster Manufacturing Research Institute (MMRI)

Hamilton, Ontario

- Deployed **Power BI** dashboards displaying KPIs on **40+** machines at **Honda plant**, focusing on **scalability**
- Streamlined data processing** and **transformation** for **30GB+** unstructured **large datasets** with **Python**
- Leveraged **scikit-learn**, **Pandas**, **NumPy** for **statistical modelling**, used to identify **100+ machine defects**
- Automated data cleaning, visualization** and **plotting** with **Plotly**, facilitating exploratory data analysis

PROJECTS & AWARDS

RAG Chatbot for Customer Reviews | CIBC

November 2024 – February 2025

- Built a **RAG-based chatbot** to **provide insights** on **600+** client reviews using **Weaviate**, **Pandas** and **Llama3**
- Attempted an **agentic workflow** with **LlamaIndex** for query intent detection and automatic data filtering
- Implemented **hybrid search** combining dense and sparse retrieval, **reducing query response time** by **15%**

CIFAR Inclusive AI Scholarship | Vector Institute

October 2024 – December 2024

- Awarded **CIFAR AI Scholarship** as part of Vector Institute's Black & Indigenous Applied Internship Program

Backend Engineer | Owambe Technologies

November 2024 – Present

- Designed **relational database tables** and **schemas** in **Supabase** to facilitate a group event planning app
- Connected the calendar page to **backend event data** to provide scheduling updates with **React Native** and **Expo**
- Building **robust deployment pipelines** with **Expo**, ensuring reliable cross-platform functionality and releases

EDUCATION

McMaster University

Bachelor of Materials Engineering

Hamilton, Ontario

September 2018 – December 2023