# Yongan(Michael) Yu

**Montréal, Quebec**

yya040327@gmail.com | (438)-821-1823 | Linkedin | Github | Portfolio

## EDUCATION:

**McGill University**

**AUGUST 2022 - APRIL 2026**
**MONTRÉAL, QUEBEC**

Bachelor in **Computer Science** Major & **Geographical Information System** Minor
- CIHR, Canadian Inst of Health Research Award 2023

## SKILLS:

- **Programming:** Python, Java, C/C++, R, CSS, HTML, Javascript, Bash, XML
- **Tools**: LangChain, Scikit-learn, Git, Matplotlib, Pandas, NumPy, Selenium, Plotly
- **Frameworks**: TensorFlow, PyTorch, ReactJS, MongoDB, Django, SQL, NodeJS, FastAPI, AWS
- **Skills:** Supervised/Unsupervised Learning, Deep Learning (CNNs, RNNs), Transformers, Natural Language Processing, API implementation, Object-Oriented Programming, Test-Driven Development
- **Spoken Languages:** Mandarin, English, French

## PROJECTS:

**Circular Economy Business Idea Evaluator Tool**

**HARVARD BUSINESS SCHOOL**

- Architected an advanced GenAI framework to evaluate over **2000 startup products** and optimize circular economy enterprises based on key factors like **embedded product value**, **material access**, and **processing levels**, contributing to informed investment decisions
- Incorporated advanced **Tree-of-thought** reasoning for nuanced statistical data analysis, utilized **K-means clustering** to enhance semantic similarity-based retrieval, and achieved **79% accuracy** from validation data
- Accomplished a **Streamlit-powered** chatbot that stored in-chat contextual memory and information to provide grounded responses
- https://github.com/techandy42/GreenTechGuardians

**Web Crawler Powered by Google Gemini-Pro**

**GOOGLE GEMINI GENAI HACKATHON**

- Developed and **fine-tuned** a web crawling system using **Selenium** and **Beautiful Soup** to boost data extraction efficiency, achieving a **30% improvement in processing speed**. Streamlined the ingestion of large-scale data
- Implemented the **Gemini API** to classify main topics from URLs, outputting the processed data in JSON format. Enhanced data handling capabilities, facilitating the development of data-driven AI solutions
- Engineered a web service leveraging **FastAPI** and **Django** on runtime calculations, optimizing the performance
- https://github.com/Marvin-Deng/Gemini-Scraper

**Weather Vulnerability Analysis Tool**

**DATA RESCUE ARCHIVES & WEATHER MCGILL**

- Optimized **OpenAI API** performance with **SQL** to clean and organize around **500 datasets** from Quebec newspaper archives, focusing on extracting sections about extreme weather events, facilitated in-depth historical weather pattern analysis
- Achieved an **85% accuracy** in predicting weather vulnerabilities and retrieving weather-related data using advanced **OpenAI embeddings**, significantly enhancing the predictive analytics capabilities of the system
- Applied **Retrieval-Augmented Generation** via leveraging **Pinecone** and **LangChain** to identify historical weather vulnerabilities
- https://github.com/Michaelyya/vulnerability-Prediction-research

## WORK EXPERIENCES:

**Undergraduate Project Researcher**

**JANUARY 2024 - APRIL 2024**
**HARVARD UNIVERSITY**

Harvard Business School

- **Published** "Case study on how Artificial Intelligence Can Be Used to Enhance Venture Capital (VCs) Decision-Making" with D^3 at Harvard
- Implemented an app that VCs can use to make investments and improve startup ideas, which companies can enhance **decision-making** with predictive analytics and real-time data integration
- Conducted research on successful startups based on the **Harvard Circular Economy Database**

**Undergraduate Research Assistant**

**JANUARY 2023 - PRESENT**
**MCGILL UNIVERSITY**

Rosalind & Morris Goodman Cancer Institute

- Solved **ODEs** and **PDEs** for computing the deformation of a nonlinearly elastic body and generated Chaste source from **Oxford University** through **Cmake** and **Javac** in a **Linux** environment
- Contributed to modeling cell simulation to assist in investigating how macrophages destroy the membrane and affect cell mitosis in the epithelial layer
- Compiled a parallel center mass graph window to track cell simulation and analyze them by the **CPM (GGH) algorithm**