Bill (Yuan Hong) Sun

billyhsun@gatech.edu | GitHub | LinkedIn | Portfolio | (647) 987-3896 | Toronto, ON, Canada

EXPERIENCE

KINAXIS | Machine Learning Engineer II
Feb 2022 - Present | Toronto, ON / Ottawa, ON
- Working in an Agile environment with cross-functional teams to integrate ML and LLM solutions into the **Maestro** platform.
- Implementing and productionalizing machine learning models in retail and supply chain demand forecasting using Python, Argo, & Kubernetes.
- Building Maestro Chat - An OpenAI-powered chatbot and Retrieval Augmented Generation (RAG) system for answering user questions.
- Improved RAG retrieval and generation accuracy by over 30% through agent improvements, using Mosaic AI and LangChain tools.
- Developing, designing, and scaling machine learning pipelines and infrastructure to onboard new customers in retail, manufacturing, and supply chain. Helped reduce cloud costs by over 90%.
- Driving impact in ML in supply chain through blog posts and patents.

NURALOGIX | Data Science Software Developer
May 2019 – Aug 2020; Jul 2021 - Jan 2022 | Toronto, ON
- Worked in an Agile environment involving both development and research work with multiple interdisciplinary teams.
- Contributed to the development of the **Anura** health monitoring app.
- Developed internal tools and ETL pipelines in Python and Jenkins to automate data collection, processing, and cleaning.
- Simplified access to the DeepAffex cloud by developing an open-source Python API package, as well as writing public documentation.
- Developed a full-stack web application for testing survey-based machine learning models. Includes a microservice back-end (using AWS Lambda), a Flask front-end, and PostgreSQL database.
- Helped prepare multiple customer demos of proof of concept products.

UNIVERSITY OF TORONTO | Machine Learning Team Lead
Sep 2020 – Dec 2024 | Toronto, ON
- Working with Dr. **Kang Lee** of the Centre of Smart Learning to develop ML and LLM methods for assessing mental & physical health conditions.
- Leading the machine learning research team. Supervising undergrad research students on machine learning and data analysis projects.
- Led and developed a web application (in Django and React) for hosting machine-learning-based health assessments and for data collection.
- Leading a cross-functional R&D team in designing and developing an LLM video chatbot web application powered by OpenAI (implemented using LangChain, React, and TypeScript) for testing mental health therapy and assessment using Large Language Models.
- Written and first-authored several publications in academic journals.

PROJECTS

VACCINE SENTIMENT DASHBOARD | Co-developer
Sep 2020 – Dec 2020 | Public Health Ontario
- Applied Natural Language Processing techniques to develop a sentiment model that detects Tweets containing misinformation on vaccines.
- Developed an interactive dashboard and a data pipeline in Python Flask, Dash, PostgreSQL, and Heroku that scrapes and analyzes new Tweets daily and displays vaccination misinformation statistics.

EDUCATION

GEORGIA TECH
MSc in Computer Science
major in Machine Learning
2023 – 2025 (Expected)

UNIVERSITY OF TORONTO
BASc in Engineering Science
major in Machine Intelligence
minor in Engineering Business
2016 – 2021

MA in Applied Psychology
major in Developmental Psychology and Education
2021 – 2023

SKILLS

PROGRAMMING
Python • Java • SQL • C/C++
HTML/CSS/JavaScript • R

FRAMEWORKS / LIBRARIES
Pandas • Spark • Argo • Airflow
TensorFlow • Sklearn • PyTorch • Keras
NLTK • Dash • Flask • Django • React

OTHER TECHNOLOGIES
Git/GitHub • AWS • Azure • Docker
Heroku • Atlassian toolkit • Databricks
Linux / Unix • Kubernetes • REST / gRPC
API / Microservices • Machine learning
Deep learning / Neural networks • NLP
LLMs • OpenAI • LangChain • Terraform

PUBLICATIONS

[1] S. Liu, Y. H. Sun, and et al. Symptom based detection models of covid-19 infection using ai. In Y. Eldar and N. Lidströmer, editors, *Artificial Intelligence in COVID-19*. Springer Nature, 2022.

[2] Y. H. Sun and et al. A novel machine learning approach to shorten depression risk assessment for convenient uses. *Journal of Affective Disorders*, 2022.

[3] Y. H. Sun, H. Luo, and K. Lee. A novel approach for developing efficient and convenient short assessments to approximate a long assessment. *Behavior Research Methods*, 2022.