

PETER (HAO) YU

Montreal, Quebec, Canada

✉ hao.yu2@mail.mcgill.ca  haoy  YHPeter  yupeter (**Competition Expert**)

[Homepage \(yhpeter.github.io\)](http://yhpeter.github.io) [Google Scholar](#)

Education

McGill University

Software Engineering B.S. 3.82/4.00

Computer Science M.S. (Thesis) 4.00/4.00

Montreal, Quebec, Canada

Sep 2021 – May 2024

Sep 2024 – Exp. Jan 2026

Skills

Languages: Python/JavaScript/SQL/C++/C#/Java/C/OCaml

ML/Data Science: Pandas, Numpy, PyTorch, TF2, HuggingFace, LangChain, LlamaIndex, vLLM

Tools: PostgreSQL, Elastic Search, MongoDB, Docker, Git, AWS, Prometheus, Grafana, SLURM

Work Experience

▷ **McGill University** Teaching Assistant

Montreal Contract **Sep 2024 – Currently**

Teaching Assistant for COMP 202 (Foundations of Programming) and COMP 206 (Intro. to Software Systems, Linux/Bash)

▷ **SmartPrep.ai** ML Engineer

Remote Part-time **Sep 2023 – Currently**

Develop **ThetaWave.ai**, the all-in-one note-taking and study platform powered with AI and RAG. Designed and implemented RAG and quality evaluation systems based on user-uploaded content, history dialogue, and web search results with NextJS and AWS Lambda. Build the releasing CI/CD pipeline and the monitoring system through Prometheus and Grafana.

▷ **Tencent** Research Intern

Shanghai Full-Time **May 2024 – Aug 2024**

Research in the Retrieval-Augmented Generation (RAG), focusing on data analysis of players' in-game interaction.

▷ **Morgan Stanley** Summer Analyst

Montreal Full-time Contract **May 2023 – Aug 2023**

▷ **Drakkar Capital** Hedge Fund Programmer

Montreal Full-time Contract **May 2022 – Aug 2022**

Publications

[1]Hao Yu, et al. [INJONGO: A Multicultural Intent Detection and Slot-filling Dataset for African Languages](#), Under Review.

[2]Hao Yu, et al. [How to Unlock Time Series Editing? A Diffusion-Driven Multi-Grained Control](#), Under Review.

[3]Hao Yu, et al. [Evaluation of Retrieval-Augmented Generation: A Survey](#), CCF BigData 2024, Springer Published.

[4]Jacob-Junqi Tian, Hao Yu, et al. [Web Retrieval Agents for Evidence-Based Misinformation Detection](#), COLM 2024.

[5]Hao Yu*, et al. [Weakly Supervised Person Name Extraction for Fighting Human Trafficking](#), Findings, EMNLP 2023.

[6]Hao Yu*, Zachary Yang*, et al. [Open, Closed, or Small Language Models for Text Classification?](#) In: arXiv:2308.10092 (2023).

Awards and Competitions

\$15,000, Research Funding from Canada First Research Excellence Fund

Dec 2024

Tuition plus \$29,000, Full Scholarship, McGill Master of Computer Science (Thesis)

Mar 2024

\$7,000 CAD, Science Undergraduate Research Awards (2023) by McGill University

Apr 2023

4th place, NeurIPS 2022, OGB-LSC PCQM4Mv2 Predict Quantum Property of Molecular Graphs

Oct 2022

Top 2%, rank 71/3204, Wechat Big Data Challenge, Multimodal Video Classification

Jul 2022

Top 1.5% in Kaggle competitions, highest rank 2660/185,584, Kaggle Expert

Apr 2022

Bronze, rank 86/1588, Kaggle Happywhale - Whale and Dolphin Identification

Apr 2022

Bronze, rank 159/2058, Team Leader, Kaggle Feedback Prize - Evaluating Student Writing

Mar 2022

Silver, rank 46/1505, Team Leader, Kaggle Sartorius Cell Instance Segmentation

Dec 2021

First prize, top 0.7%, rank 20/2849, Team Leader, AI-Earth AI Innovation Challenge

Feb 2021

Global AI Technology Innovation Competition Track I, Top 5%, rank 198/4337

Jan 2021

First place, rank 1/1619, Global AI Technology Innovation Competition Warmup Track

Jan 2021