Yao Chen

514-923-8653 yao.chen5@mail.mcgill.ca linkedin.com/Yao Chen github.com/Yao Chen

Education

McGill University **Jan. 2023 – Present**

Master of Science in Computer Science (Thesis) *Montreal, Quebec*

- Cumulative GPA: 4.00 / 4.00
- Relevant Coursework: Generalized Linear Models, Matrix Computations, Digital Communication, NLP

McGill University **Sep. 2019 – Dec. 2022**

Bachelor of Science in Major Statistics and Computer Science *Montreal, Quebec*

- Cumulative GPA: 3.94 / 4.00
- Relevant Coursework: Data Structures and Algorithm, Applied Machine Learning, Artificial Intelligence, Time Series Analysis, Applied Regression, Linear Algebra, Statistics, Probability
- Course Assistant for Artificial Intelligence Course, Team Mentor for Programming Lang and Paradigms Course

Experience

Business Analytics Intern - Ontario Teachers' Pension Plan **Sep. 2024 – Apr. 2025**

Snowflake, Data Pipelines, Feature Engineering, ML Deployment, Python, SQL *Toronto, Ontario*

- Designed and implemented a modular data pipeline in Snowflake using Python and SQL to automate ingestion and cleaning from heterogeneous sources (IBM DB2, MSSQL).
- Developed XGBoost regression model, reduced manual QA workload by 79% while ensuring $< 5\%$ operational risk, applying end-to-end automation and system monitoring techniques.
- Designed and deployed a user-interactive prediction service via StreamlitApp in Snowflake and to support business use.
- Collaborated with analytics and engineering teams to translate business rules into reliable, maintainable, and reproducible data-driven systems.

Software Engineer Intern - Microsoft (China) Co., Ltd **Apr. 2024 – Jul. 2024**

SCOPE, Redash SQL, Python, Excel, PowerPoint, Data Visualization, User Analytics *Suzhou, China*

- Processed and analyzed 100+ GB telemetry data, identifying key churn patterns and system usage bottlenecks.
- Used GPT4.0 to parse internal docs and automate telemetry metadata extraction and semantic-level aggregation.
- Designed hierarchical activity classification framework to unify and classify raw telemetry events for scalable analysis.
- Built interactive Sankey diagrams to visualize user journeys across interface layers, highlighting frequent navigation paths and high-churn nodes for targeted UX optimization.

Research Assistant - Scientific Computing Lab – McGill University **Apr. 2022 – Present**

MATLAB, Neural Network, Tensorflow, Numpy, Matplotlib *Montreal, Quebec*

- Investigating non-convex optimization methods for sparse signal recovery in MIMO communication systems.
- Developing and analyzing ADMM-based algorithms to solve L_0-regularized least squares problems.
- Implementing deep learning aided sphere decoding(SD) algorithm for an underdetermined system, focusing on channel permutation to accelerate the search for SD.

Projects

Colosseum Survival Gameagent | *Github, AI* **Mar. 2022 – Apr. 2022**

- Designed a relaxation tracking algorithm with an evaluation function that properly define the reward of a position.
- Limited the maximum searching depth of 2 with pruning to achieve low time complexity and space complexity.
- The algorithm had a 100% win rate against random agents, and 85.7% win rate against other agents in the tournament.

BlackJack card games | *C++* **Mar. 2021 – Apr. 2021**

- Developed a BlackJack card game from scratch, utilized object-oriented programming concepts to implement the logic for creating a deck, dealing cards, and shuffling.

Technical Skills

Languages: Python, Java, C, C++, C#, Bash, R, SAS, SQL, MIPS Assembly, Ocaml, HTML, CSS
Frameworks and Tools: VS Code, PyCharm, Jupyter Notebook, Xcode, RStudio, MATALB, AWS EC2, Unity, Cursor
Other: Graphic Design using Procreate, Blender and Adobe Photoshop