Wyett "Huaye" Zeng

wyettzeng@g.harvard.edu | 519-729-8107 | wyett-zeng.com | LinkedIn | GitHub | Google Scholar

Educations

Harvard University — Cambridge, Massachusetts, USA
Master of Science in Data Science — Sep 2025 – Dec 2026 (Expected)

University of Waterloo & Wilfrid Laurier University — Waterloo, Ontario, Canada
Bachelor of Computer Science and Business Administration — Sep 2020 – Apr 2025
Laurier Alumni Gold Medalist (Major Average 94.79 / 100)

Summary of Qualifications

- Fluent in programming languages such as **Python, C++, C, Java, VBA, C#, SQL, Go, Bash, JavaScript, HTML/CSS**
- Strong quantitative analytical skills in **SQL**, **R**, **Tableau**, and **Power BI**. Knowledgeable in database systems such as **Clickhouse, PostgreSQL**, and **MySQL**. Experienced with data science packages such as **Pandas**, **PyTorch, Keras,** and **Numpy**.
- Laurier Student Investment Fund Winter 2025 - fundamental analysis in public equities.
- Actively running a **$200,000 portfolio** in public equity and fixed income with my own money; risk is managed through diversification in asset classes and geographical allocations. Currently, the NA portfolio has a **25.83% IRR**.
- **200+ citations** for machine learning publications in top venues (ACL & TMLR), including first-author papers [Google Scholar].

Work Experiences

Machine Learning Engineer Intern | GPTZero — Sep 2024 – Aug 2025

- Launched the AI-Reviewer product, developing an end-to-end ML server with a **multi-step agent workflow** using LlamaIndex for retrieval, grading, and feedback. The system processed 1,000+ assignments in its first month—surpassing all company target.
- Led the AI-Grader product using transformer-based architectures. On IELTS, a widely used standardized English tests, the model achieves **88% accuracy** within ±1.0 band (out of 12) and **97%** within ±1.5 bands, rivaling human graders.
- Optimized the grammatical error correction model, reducing edit distance by **50%** and improving GLEU score from 0.7 to 0.8.

Quantitative Developer Intern | Gradient Boosted Investments Inc. — Jan 2024 – Apr 2024

- Added features in the **Boosted.ai trading algorithm** with Python to optimize daily stock selection for all the company's clients. The added features expand the algorithm's capabilities to construct portfolios that align closer with the client's needs.
- Developed hedge basket rankings feature, enabling systematic evaluation of multiple hedging strategies through four distinct methods: defensive, emphasizing stability and reduced variance; factor tolerant, targeting maximum performance under market factor exposures; idiosyncratic, maximizing unique portfolio-specific returns; and all-weather, maintaining a balanced allocation.
- Rewrote the factor model algorithm, reducing scheduled inference time for 5,000+ customer models by over 90%. The algorithm uses NumPy, ClickHouse, and PostgreSQL to efficiently compute **economic factor values** for **20,000+ publicly listed securities**.

Data Scientist Intern | Canadian Imperial Bank of Commerce — Jan 2023 – Apr 2023

- Developed the quantitative portfolio builder, which formulates and solves an optimization problem using QSolver to construct a portfolio whose return is within ±2.8% of the benchmark, uncovering insights into "obscure" alternative investment hedge funds.
- Developed a market analysis program that integrates streaming data from Morningstar APIs, generates **300+** interactive market trend graphs for team members in minutes, and delivers short-term forecasts using GRU and LSTM models built with **TensorFlow**.
- Partook in numerous due diligence meetings with portfolio managers from major hedge funds such as TCC, Group RMC, and Hamilton Lane, actively engaging in discussions on fund structures, risk exposures, and investment theses. Following each meeting, produced comprehensive reports that identified and analyzed key areas of concern—including liquidity options, distribution schedules, market correlations, FX risks, and interest rate sensitivities—providing actionable insights for senior decision-makers.

Research Experiences

Transformer Trader

- Completed an **Honours Thesis** on training LLMs to predict market sentiment from economic news by developing a dataset of over 120K entries from Dow Jones News Wires and Wharton Research Data (**Publication under writing**).
- Fine-tuned custom LLMs by combining the Llama base model with custom heads using PyTorch, DeepSpeed, and QLoRA, then transformed the model to a Transformer-based trading algorithm; back tested trade shows **over 15% IRR.**
- Studied topic compositions of financial news over time using methods such as Latent Dirichlet Allocation and Gaussian Mixture Model on the Bert embedding of news text.