

Ethan Thoma

ethanthoma@gmail.com · linkedin.com/in/ethanthoma

PROFESSIONAL SUMMARY

Graduate researcher specializing in neuro-symbolic AI and large language models, with extensive experience in machine learning, data pipeline development, and distributed computing. Demonstrated expertise in implementing novel ML models and conducting large-scale experiments using cloud infrastructure.

EXPERIENCE

Graduate Research Assistant

May 2024 - Present

University of British Columbia

Kelowna, BC

- Leading research on knowledge graph representations for conversational data, improving model F1 score from 0.3564 to 0.4762
- Developing novel approaches to enhance LLM performance on GitHub issue data through structured knowledge representation
- Collaborating with Dr. Gema Rodríguez-Pérez to advance research in AI reasoning and knowledge representation

Graduate Research Assistant

September 2023 - December 2023

UBC Sauder School of Business

Vancouver, BC

- Designed and implemented a multi-process data pipeline that reduced embedding time for 9.5M Yelp reviews from months to one week
- Engineered robust pipeline with exponential backoff, rate limiting, and fault tolerance for OpenAI API integration
- Developed a novel Gaussian-distance ordinal classification model, improving prediction accuracy on limited compute resources

Research Volunteer

May 2023 - September 2023

University of British Columbia

Vancouver, BC

- Pioneered a novel approach for evaluation metrics generation using discriminative datasets
- Leveraged SLURM to orchestrate distributed testing of 12 Hugging Face models on Compute Canada, generating 11k sentences
- Coordinated and analyzed 600 human evaluation tasks on AWS Mechanical Turk, validating the effectiveness of the new metrics

EDUCATION

Masters of Computer Science*January 2024 - Present**University of British Columbia**Vancouver, BC*

- Relevant Coursework: NLP Commonsense (CPSC 532), Intelligent Systems (CPSC 422), Advanced Statistical Modelling (STAT 538)
- Received academic bursaries for academic performance

Bachelor of Science in Computer Science*September 2019 - May 2023**University of British Columbia**Vancouver, BC*

- Relevant Coursework: Computer Vision (CPSC 425), Methods for Statistical Learning (STAT 406), Natural Language Processing (CPSC 436N), Advanced Machine Learning (CPSC 440)

SKILLS

- Programming Languages: Python (PyTorch, Lightning), TypeScript, Odin, Zig, Gleam
- ML/AI Tools: Hugging Face, OpenAI, Google Cloud AI
- Infrastructure: Git, SLURM, Nix, AWS, Google Cloud
- Core Competencies: Machine Learning, Natural Language Processing, Distributed Computing
- Research & Development: Problem Analysis, Experimental Design, Technical Writing
- Collaboration: Cross-functional Team Leadership, Project Management, Scientific Communication