

Ethan Thoma

Vancouver, BC | | [Linkedin](#) | [GitHub](#) | [Personal Website](#)

Graduate researcher specializing in neural-symbolic AI and large language models (LLMs), with expertise in developing novel ML architectures. Focused in natural language processing (NLP), neuro-symbolic AI, and interpretability/explainability in machine learning. **Expected graduation: January 2025.**

RESEARCH 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**
- Developed the Dynamic Computation Tree Model (DCTM), a neural-symbolic framework leveraging dynamic computation trees to enhance LLM numerical reasoning, achieving a significant improvement (**0.52 vs. 0.01** F1-score) on the NumGLUE benchmark compared to baselines

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 for sentiment analysis, improving prediction accuracy on limited compute resources

Graduate Research Volunteer

May 2023 - September 2023

University of British Columbia, Vancouver, BC

- Pioneered a novel approach for automatic evaluation metrics generation for natural language generation tasks 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 532V), Intelligent Systems (CPSC 422), Advanced Statistical Modelling (STAT 538), Computer Vision (COSC 544)
- Received 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)

RESEARCH CONTRIBUTIONS

Ethan Thoma, Gema Rodríguez-Pérez. (2025) Dynamic Compute Tree Model. *In preparation.*

Ethan Thoma, Gema Rodríguez-Pérez. (2024) Knowledge Graph Representations of GitHub Issues: An Exploratory Study. *Unpublished.*

Jingqian Liu, Ethan Thoma, Vered Shwartz. (2023) An Automatic Plausibility Evaluation Metric for Generative Commonsense Reasoning. *Unpublished.*