

HAOCHONG (ROGERS) YANG

☎ (437) 981-3043 ✉ haochong.yang@mail.utoronto.ca 🌐 HaochongYANG 📄 haochongyang.github.io 🌐 LinkedIn

EDUCATION

BSc (Hons). in Statistics and Computer Science

University of Toronto (UofT) | Toronto, ON, Canada

Sep 2021 – May 2025

cGPA: 3.96/4.00

– Thesis Supervisor: Prof. Elena Tuzhilina

PUBLICATIONS

[M]: MANUSCRIPT, [P]: PUBLICATION, [R]: UNDER REVIEW

[M] **Yang, H.**; Tuzhilina, E. Dimension Reduction Methods for Single-Cell DNA Structural Heterogeneity Analysis. *Manuscript in preparation*

[P] **Yang, H.**; Huang, M.; Chen, X.; He, Z.; Pu, S. Enhanced Real-Life Data Modeling with the Modified Burr III Odds Ratio–G Distribution. *Axioms* (SCIE Q1 journal) 2024, 13, 401. <https://doi.org/10.3390/axioms13060401>

[R] **Yang, H.**; Sun, Y. H.; Lee, K. A Novel Approach for Efficient Multi-class Anxiety Level Prediction Approximation for Long Assessments. *Submitted: Journal of Anxiety Disorders*

[P] Sun, Y.; **Yang, H.** Influence of Known Genetic Risk Factors for Atherosclerosis on People with Different Racial Backgrounds. *URN CST Journal* 2024, Volume 8, Issue 11. <https://doi.org/10.26685/urncst.712>

SCHOLARSHIPS & AWARDS

[A]: ACADEMIC, [R]: RESEARCH

[R] University of Toronto Excellence Award (UTEA, \$7500)

University of Toronto | 2025

[A] Louis Savlov Scholarships in Sciences and Humanities (\$500)

University College | 2025

[R] Summer Undergraduate Data Science Scholarship (SUDDS, \$7500)

Data Science Institute | 2024

[A] C.L. Burton Scholarship for Mathematics and Physical Sciences (\$500)

University College | 2023

[A] Louis Savlov Scholarships in Sciences and Humanities (\$500)

University College | 2022

[A] University of Toronto Scholar Award (\$1500)

University of Toronto | 2022

[A] Dean's List Scholar Awards

Faculty of Arts & Science | 2022, 2023, 2024

INDUSTRY EXPERIENCE

Data Analyst (Co-op)

May 2023 – Aug 2023

Toromont CAT | Concord, ON

- Trained a LLM with documentation and system data using LangChain and OpenAI API to develop a company chatbot “CatGPT” which will benefit over 2000 people as users to look for company business solution
- Optimized the Strategic Asset Management platform with Power BI and Python, enabling real-time monitoring and efficient tracking of heavy mining machinery status and operational activities
- Created model to keep track of machine usage patterns and forecast component replacement using Python
- Collaborated with 16 Canadian mine owners to update machine component status, resulting in a significant reduction of \$57 million in overdue value

Software Developer (Internship)

May 2022 – Aug 2022

BL Innovare | Markham, ON

- Spearheaded daily data analysis, software development, and support operations for a leading company specializing in vehicle inspection and maintenance products and services
- Enhanced the performance of Bodyguard 2.0, an advanced machine learning application designed to simulate human perceptual and decision-making processes for vehicle inspection and damage analysis
- Contributed to the back-end development of a language translating service integrated to the vehicle analysis platform using Python and Django
- Demonstrated proficiency in crafting intricate SQL queries to process customer data and conduct in-depth data analysis

UNDERGRADUATE RESEARCH EXPERIENCE

Heterogeneity of Single Cell DNA Structure Analysis

Department of Statistical Sciences | U of T

Supervised by Prof. Elena Tuzhilina

- Conducted individual research under professor's guidance, with regular progress updates via weekly meetings
- Explored the quantitative measurements such as local contacts difference and Euclidean distance for classifying single cells based on DNA Hi-C data and contact matrices
- Applied dimension reduction techniques such as PCA, t-SNE, and Uniform Manifold Approximation & Projections to find representations of cells in low dimensions for clustering
- Designing low-rank embeddings for cell matrices to achieve effective cell approximation and further provide efficient classification procedure via proof of concepts

Efficient Anxiety Level Prediction using Machine Learning

Department of Applied Psychology and Human Development | U of T

Supervised by Prof. Kang Lee

- Enhanced machine learning model for DASS-42 mental health survey result prediction, reducing the length of the original assessment while maintaining high prediction accuracy.
- Developed an ensemble model that achieved an accuracy of 85% in classifying anxiety levels across five categories with high reliability, while the baseline chance level is 20% in five-level classification tasks.
- Implemented web app to enable anxiety assessments based on the model using React for deployment.
- First-authored a manuscript detailing method and findings, submitted to the *Journal of Affective Disorders*

Data Modeling with the Modified Burr III Odds Ratio-G Distribution

Department of Mathematics and Statistics | University of West Florida

Supervised by Prof. Shusen Pu

- Developed a novel statistical model integrating odds ratio with Burr III distribution, focusing on the Burr III Scaled Inverse Odds Ratio-G subclass
- Explored mathematical properties of the distribution like hazard rates and quantiles through rigorous proofs
- Conducted simulations to test the model's robustness and applied it to medical datasets, demonstrating improved flexibility and predictive accuracy compared to established distributions
- Presented findings in a comprehensive research paper, showcasing the model's efficacy and potential applications in statistical analysis

Design LLM Systems in Mental Health Assessments

Department of Applied Psychology and Human Development | U of T

Supervised by Prof. Kang Lee

- Developed a system of LLMs to approximate the role of psychologists to conduct mental health assessments
- Led prompt engineering work for LLMs and coded a duo-agents psychological analysis robust system
- Worked on front and back-end to enable the multimodality detection of facial expression, tone, and actions
- Collaborated with team members to create research proposals, draft ethics reports, and plan clinical studies

Social Theories Validation using Reinforcement Learning (RL)

Department of Psychology | U of T

Supervised by Prof. William Cunningham

- Led Leaky Emotions project to research on how agent emotions affect others' behavior in a social setting
- Fine-tuned RL environment to validate various social cognition theories with support of neural networks
- Designed visuals for training process, facilitating rigorous game logic execution and components modification
- Leveraged ML techniques to train CPU agents for optimal game results, and performed statistical analysis on game data to show the agents gaining social impacts on others to receive higher mutual rewards

TECHNICAL COMPETENCIES

Coding Python, Java, JavaScript, Shell, Assembly
Data R, SQL, Tableau, PowerBI, Kafka, Docker

Frameworks TensorFlow, PyTorch, Langchain, Django
Packages Pandas, NumPy, Matplotlib, Scikit-Learn

PRESENTATIONS

Poster Session

- Efficient Multi-class Anxiety Level Prediction Approximation for Long Assessments, *Data Science Institute, Toronto, Canada* *DSI Summer Research Day Aug 2024*

Presentation

- Dimension Reduction Methods for Single-Cell DNA Structural Heterogeneity Analysis, Department of Statistical Sciences, UofT, Toronto, Canada *DoSS Readings Showcase Apr 2025*
- Application of Statistical Methods in Psychological Assessments, Seattle, WA, United States *Invited: Joint Mathematics Meetings, Jan 2025*