

Mirza Ahmadi

Guelph, ON, N1L 0N6 | (647) 390-8697 | mirza.am.ahmadi@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SKILLS SUMMARY

- **Languages & Tools:** Python, R, Bash, Linux, High-Performance Computing (HPC), SQL, Git, HTML & CSS
- **Certifications:** [Machine Learning \(ML\) + Deep Learning \(DL\) - Compute Ontario](#) | [AI + Python - Harvard CS50](#) | [DL + Reinforcement Learning - CIFAR](#) | [Linear Algebra + Calculus - Coursera](#)

PERSONAL PROJECTS

Spero: A Chatbot for Mental Health Support | [Research Proposal](#) | *Python* Jan 2025 - Present

- Designing an **NLP** chatbot to give personalized emotional support for users seeking mental health assistance
- Utilizing **RAG** to retrieve therapist-client transcript data to ensure contextually-accurate responses
- Evaluating quality of chatbot-generated responses against real transcript data using **BLEU** and **ROUGE** scores
- Bridging mental health care gaps by offering 24/7 support for individuals facing accessibility barriers
- Leveraging **Pandas** and **NumPy** for data preprocessing and **OpenAI's API** for **LLM**-based text generation

[Evaluating Viability of Synthetic EEG Data \(BrainHack 2024, SickKids\)](#) | [Demo](#) | *Python* Dec 2024

- Evaluated **GAN**-generated synthetic EEG data as a proxy for limited real-world clinical data in model training
- Collected and cleaned EEG data from individuals performing two cognitive exercises, creating a labeled dataset
- Trained a **Random Forest** model achieving **95%** accuracy using **Pandas**, **NumPy**, and **Scikit-learn**
- Generated 200 000 synthetic datapoints with a **GAN** and compared them to collected data in model evaluation
- Visualized results using Confusion matrices, ROC-AUC curves, loss plots, and learning curves with **Matplotlib**

[Evaluating Algorithms for Gene Sequence Classification](#) | *R (Tidyverse)* Nov 2024

- Compared the performance of **Random Forest** and **Linear Regression** models in gene sequence classification
- Performed **exploratory data analysis** to summarize sequence statistics and identify patterns in the data
- Visualized and compared sequence length distribution, k-mer frequency proportions, and feature importance

[Venom Type Classifier](#) | [Demo](#) | *Python* Mar 2024

- Developed a **Support Vector Machine** model to predict venom type from protein data, achieving **94%** accuracy
- Utilized **Scikit-learn**, **Matplotlib**, and **Pandas** for model building and visualization of PCA data

EXPERIENCE

M. Sc. Research Project - *University of Guelph* Guelph, ON | Jan 2024 - Present

- Developing a **data science pipeline** to derive genome insights, informing disease and genetic variability
- Integrating **Python**, **R**, **SQL**, bioinformatics tools, and **ML** for data processing, storage, and model development
- Managed **HPC** jobs for genomic data analysis using **SLURM** on Compute Canada **Linux** systems
- Deploying the pipeline on a cloud platform to enhance efficiency and data accessibility for genomic researchers

Evolutionary Biology Researcher - *University of Guelph* Guelph, ON | May 2022 - Present

- Designed and executed an independent research report on snake venom evolution across 127 species
- Analyzed proteomic and phylogenetic data in **Excel** and **R** to compile phylogenies and create graphical analyses
- Authoring a research paper for publication to enhance our understanding of venom and complex trait evolution

EDUCATION

M. Sc., Bioinformatics & AI (Research-Based) - *University of Guelph* Guelph, ON | Jan 2024 - Present

- Courses: ML, Mathematics for ML (Linear Algebra, Statistics, Calculus), Bioinformatics Software Tools

B. Sc., Zoology Major with Minor in French - *University of Guelph* Guelph, ON | Sept 2019 - April 2023