# Mirza Ahmadi

Guelph, ON, N1L 0N6 | (647) 390-8697 | mahmad15@uoguelph.ca | LinkedIn | GitHub | Portfolio

## SKILLS SUMMARY

- **Languages & Tools:** Python, R, Bash, Linux, High-Performance Computing (HPC), 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

**Enhancing EEG Classification with GANs (BrainHack 2024, SickKids)** | Demo | ***Python*** **Dec 2024**

- Collected and cleaned EEG data from individuals performing two cognitive exercises, creating a labeled dataset
- Trained a **Random Forest** model with **95%** accuracy using **Pandas**, **NumPy**, and **Scikit-learn**
- Generated 200 000 synthetic datapoints with a **GAN** to test if synthetic data can serve as a proxy for real data
- 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
- Visualized sequence length distribution, k-mer frequency proportions, and feature importance

**Road Sign Categorizer** | ***Python*** **May 2024**

- Implemented a **convolutional neural network** to classify road signs into 43 categories with **90%** accuracy
- Leveraged the libraries **TensorFlow**, **Scikit-learn** and **OS** for model development and data preprocessing

## EXPERIENCE

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

- Developing an **ETL pipeline** to derive genome annotation insights, informing disease, variability, and evolution
- Integrating **Python, R**, bioinformatics tools, and **ML** for data processing, sequence classification, and modeling
- Using Compute Canada systems and **SLURM** to manage **HPC** jobs for genomic data processing and storage
- 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: Machine Learning , Math for Machine Learning, Genomic Methods, Bioinformatics Software Tools
- Graduate student representative on the Bioinformatics Steering Committee and Graduate Student Council

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