Hooman **Ramezani**

Machine Learning Engineer

hooman125@gmail.com
linkedin.com/in/hoomanramezani
github.com/HoomanRamezani
hoomanramezani.com

EDUCATION

University of Toronto, ***MASc***, *Operations Research & Artificial Intelligence* **2023 - 2024**

- **Thesis:** Tuning medical-focused LLM and ViT models with **multimodal** clinical data for lung cancer treatment planning
- **GPA 4.0 / 4.0** | **Coursework**: Large Language Models, Cloud Data Engineering (Spark, AWS), Deep Learning Theory

University of Waterloo, ***BASc***, *Systems Design Engineering* **2018 - 2023**

- **GPA 3.7 / 4.0** | **Coursework**: Intro. Deep Learning, Intro. Machine Learning, Pattern Recognition, Neuroscience

EXPERIENCE

Machine Learning Researcher | *University Health Network* **Full Time - MAY 2023 - PRESENT**

- **Authored 'Lung-DETR'** architecture for lung nodule detection, achieving F1 score of **94.2%** (SOTA), utilizing Detection-Transformer, custom Focal Loss, custom data preprocessing to overcome severe class imbalance and **sparsity**
- Building **medical multi-modal LLM** for lung cancer treatment via cross-modal analysis of CT images and physician clinical notes to prescribe personalized plans, utilizing **LoRA** fine-tuning and cross-attention for **token alignment**

Machine Learning Engineer | *Advanced Micro Devices* **Internship - SEPT. 2022 - APRIL. 2023**

- Accelerated **LLM inference** on AMD CPUs with ZenDNN, applying model compression with **pruning** and **quantization**
- Built content **recommendation** model with candidate retrieval and reranking stages, with **80% latency** reduction
- Managed workflows within a distributed training pipeline with **Spark + Azure**, handling multi-terabyte **cloud** datasets

Machine Learning Engineer | *Apple (formerly DarwinAI)* **Internship - JAN. 2022 - APRIL. 2022**

- Delivered a **CNN** based Liver Fibrosis diagnosis model to **Pfizer** team, reducing examination time for clinicians by **40%**
- Produced a highly **sensitive** and **generalized** model using custom data **augmentations** and multi-GPU training (DDP)
- Collaborated with customers to translate requirements into scalable deep learning solutions deployed on edge devices

Deep Learning Developer | *Applied Brain Research* **Internship - MAY. 2021 - AUG. 2021**

- Engineered entire lifecycle for a **drone object-detection** model to identify surface defects on complex structures
- Built **Unreal Engine 4** simulation to synthetically generate **50,000** annotated images to overcome lack of real-world data
- Achieved **94% YOLOv3** accuracy using CNN-LSTM model to track **temporal** data in video frames

Other Internships (12 months) Data Engineer. (2020) Android Developer. (2020) QA Developer. (2019)

RESEARCH // PROJECTS

Machine Learning Lead | *Livy Education Chatbot Team* **JAN 2024 - PRESENT**

- Leading AI for assignment research chatbot built for Canvas students, integrating **GPT-4**, **RAG**, **Web** using **LangChain**
- Optimizes content relevancy through **RAG reranking**, utilizing a **Chroma** vector database and **finetuned embeddings**
- User queries returned with highly relevant course materials, **inline citations**, providing informed ideas for assignments.

Advanced Detection of Parkinson's Freezing of Gait | *University Of Waterloo* **2023**

- Introduced a **novel time-series InceptionTime** model (time-series CNN) for the early detection of **Freezing of Gait (FOG)** in Parkinson's patients, analyzing biometric data (EMG, ECG) for fall forecasting with **94%** testing accuracy

Enhanced Robotic Grasping with Adapted PointNet | *UW Vision and Image Processing Lab* **2022**

- Enhancing **robotic arm grasp** capabilities, adapted custom PointNet deep learning models to develop a **3D computer vision** system capable of identifying optimal grasp points from **LiDAR** camera, demonstrating an **87%** grasp success rate