Hooman **Ramezani**

Machine Learning & ML Ops

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

EDUCATION

University of Toronto, *MASc, Operations Research (Machine Learning)* — **EXPECTED JAN 2025**

- **Thesis:** Tuning medical-focused LLM and ViT models with **multimodal** clinical data for lung cancer treatment planning
- **Relevant Coursework**: Cloud Data Eng. (Hadoop, AWS), Visual Computing, Deep Learning Theory, Graph Neural Nets.

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

- **Relevant Coursework**: Intro. Deep Learning, Computational Neuroscience, Intro. Machine Learning, Pattern Recognition,

EXPERIENCE

Machine Learning Engineer (ML-Ops) | *Advanced Micro Devices* — **MAY. 2022 - AUG. 2022**

- Deployed **LLMs** (BERT / GPT) for inference on AMD CPUs, applying model compression with **pruning** and **quantization**
- Architected a scalable, distributed data pipeline with **Hadoop + Azure**, handling multi-terabyte **cloud** datasets
- Trained and deployed **BERT-based LLM** for sentiment analysis with Few-Shot learning, with **80% latency** improvement

Deep Learning Developer | *DarwinAI* — **SEPT. 2021 - DEC. 2021**

- Delivered vision model for diagnosis of Liber Fibrosis to **Pfizer** team which reduced examination time by **40%** for medical practitioners, utilizing **PyTorch CNN / ViT** + custom data augmentations to produce sensitive and accurate model
- Utilized **CUDA GPU optimization** (prefetch, **parallel** processing, cache) to train models in a **distributed** setting
- Collaborated with customers to translate requirements into scalable deep learning solutions

Deep Learning Developer | *Applied Brain Research* — **JAN. 2021 - APRIL 2021**

- Engineered entire lifecycle for a **drone object-detection** model to identify surface defects on complex structures
- Built **Unreal Engine 4** simulation for synthetic data generation to overcome a lack of real-world data, achieved **94%** **YOLOv3** accuracy using **CNN-LSTM** model to tracking **temporal data** in video frames

Backend Data Engineering | Baron Biosystems — **MAY. 2020 - AUG. 2020**

- Manipulated large **data models** for data-driven cycling analytics platform, building functionality with PHP and MongoDB
- Constructed scalable and flexible messaging service with **10k MAU**, designing custom **API** using best practices

RESEARCH & PROJECTS

Deep Learning Lung Cancer Treatment Planning | *morLAB University of Toronto* — **JULY 2023 - PRESENT**

- Developing a novel approach for lung cancer treatment planning, integrating **Vision Transformer (ViT)** tokens with medical **Large Language Models (LLMs)** for **multi-modal** analysis of lung-CT imagery and clinical notes.
- Developed lung-CT segmentation system for unlabelled data, demonstrating significant potential achieving **90%** DiceC

Information Bottleneck Attribution for CNNs and Transformers | *University Of Toronto* — **2024**

- Developed an **attribution** method to enhance ML model **interpretability**, demystifying decision-making in CNN and Transformer models using Information Bottleneck Theory, which identifies critical information for transparent decisions

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

- Introduced a **novel time-series** model for the early detection of **Freezing of Gait (FOG)** in Parkinson's patients, analyzing biometric data (EMG, ECG) for prediction, enhancing fall prevention strategies 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 models to develop a **3D computer vision** system capable of identifying optimal grasp points from LiDAR camera, demonstrating an **89.5%** grasp success rate