Omid Vaheb

ML Engineer Toronto, Canada ovaheb@gmail.com 437-4226-606 ovaheb ovaheb

- **3+ years** of experience in **LLMs** & **audio/vision processing**, model development & optimization, and advanced algorithm design through roles at **REDspace**, **Vector**, **NRCC**, **UofT**.
- **4+ years** of hands-on experience in **PyTorch**, **Tensorflow**, **C++**, and **Cloud Computing**.

Professional Experience

Machine Learning Engineer | *REDspace* — Remote, Canada | Jan. 2025 – Present

- Reduced **Automatic Speech Recognition (ASR) word error rate & hallucination up to 70%**, leveraging contextual data and GPT-4o for **generative error correction (EC)** in Mockingbird product.
- Deployed **multimodal** EC solution on top of OpenAI Whisper, achieving SOTA performance as a PoC.

Machine Learning Associate | *Vector Institute* — Toronto, Canada | Sep. 2024 – Present

- Developed and deployed deep learning solutions for SMEs, collaborating with cross-functional teams to integrate AI into real-world applications while ensuring scalability and efficiency.
- Gained hands-on expertise in **RAG**, **LLM fine-tuning** & **alignment** for enterprise AI optimization.

Research Assistant | *University of Toronto* — Toronto, Canada | Sep. 2022 – Dec. 2024

- Improved astronomical **image denoising** by **15dB PSNR**, boosting **object detection by 7%**, and cutting observation costs **by 66%**.
- Optimized data loaders, reducing memory usage **by 80%**, execution time **by 60%**, and limiting disk access, **accelerating model training**.
- Developed an **unsupervised** denoising framework combining N2N and SURE loss, translating findings to non-technical stakeholders.

Applied Scientist | *National Research Council Canada* — Victoria, Canada | Jun. 2023 – Sep. 2023

- Reduced **MAE by 15%** in astronomical **image reconstruction** by optimizing Restormer **Transformer** and U-Net architecture. Implemented in **PyTorch** and deployed on a **multi-GPU** cluster.
- Generated **large-scale** synthetic data from real galaxy morphologies, reducing CPU **synthesis time by 20%**. Developed preprocessing pipelines for raw telescope data.

Research Assistant | *Computational Audio-Vision Lab* — Remote | Apr. 2021 – Aug. 2022

- Achieved SOTA **92% accuracy** in autism detection from infants' crying voice by fine-tuning Google YAMNet and using an **ensemble** classifier in **TensorFlow**.
- Attained **99.5%** similarity & consistency with human annotations by fine-tuning YAMNet and using energy-thresholding for cleaning voice recordings.

Data Scientist | *Virasad Startup* — Remote | May 2020 – Apr. 2021

- Reduced maintenance costs **by 10%** with real-time **anomaly detection** using LSTM and ARIMA for production line time series forecasting.

Related Skills

- **Programming:** Python, C/C++, Linux Shell Scripting, MATLAB, SQL, R, Verilog
- **AI Libraries:** PyTorch, Tensorflow, Langchain, LlamaIndex, Keras, NLTK, OpenCV, AutoML, Fastai, JAX
- **Development Tools:** Git, Slurm, Docker, MySQL, RAG, Cloud Computing, AWS (Sagemaker, Bedrock), GCP, Vertex AI, BigQuery, Wandb, HDF5

Education

Master of Applied Science in Electrical and Computer Engineering — Sep. 2022 – Sep. 2024
University of Toronto; GPA: 3.86/4 — *Toronto, Canada*

B.Sc. in Electrical Engineering, Minor in Computer Engineering — Sep. 2017 – Feb. 2022
University of Tehran; GPA: 3.92/4 — *Tehran, Iran*