

PARSA MORSAL

parsa.morsal@gmail.com | +1 (365) 888-0774 | Toronto, Ontario | [in](#) | [📧](#) | [🌐](#)

PROFESSIONAL EXPERIENCE

Bitstrapped

Machine Learning Engineer

Aug 2023 - Present

Toronto, Ontario

- Designing, implementing, & deploying tailored Generative AI solutions for companies in various sectors over GCP Vertex AI and other cloud providers, using Retrieval-Augmented Generations, Large Language Models (*LLMs*) & Recommendation Systems based on Reinforcement Learning (*RL*) & matrix decomposition.
- Fine-tuning, re-training & maintaining state-of-the-art *LLMs*, *ML*, *MLOps* & human-in-the-loop pipelines based on client's PII data & business requirements to achieve over 96% accuracy on various Natural Language Processing (*NLP*) tasks.
- Taking ownership of projects involving 5-10 engineers and communicating with multiple clients regarding their business requirements, ongoing development & deployment processes, as well as product delivery, documentation & knowledge transfer.

BlancLabs

Machine Learning Engineer

Sep 2021 - Aug 2023

Toronto, Ontario

- Investigated, designed & delivered custom Generative AI and Large Language Model (*LLM*) solutions based on enterprise data using various models with 1 to 65 billion parameters *e.g.* GPT-2, GPT-3, Bloom-3B, Falcon-40B, LLaMA-65B, GPT-NeoX-40B, GPT-J-6B.
- Fine-tuned, optimized & deployed Large Language Models (*LLM*) on AWS EC2 (*P3*, *P4*, *G3*, *G4*, and other *Nvidia A100* and *Tesla T4 GPU instances*) using model retraining & prompt engineering (few-shot learning).
- Designed, implemented & deployed custom Generative AI products *e.g.* Q&A chatbots, customer service assistant chatbots, and automation pipelines on custom datasets using various technologies *e.g.* Retrieval-Augmented Generation (*RAG*), Reinforcement Learning from Human Feedback (*RLHF*), LangChain, and Amazon SageMaker JumpStart.
- Designed, implemented & delivered end-to-end customized machine learning & deep learning products, proof of concepts (*POC*), and data analysis pipelines based on structured & unstructured enterprise data using various methods & technologies, *e.g.* Sentiment Analysis (*BERT*, *NLTK*), document classification (*RoBERTa*), data extraction (*LayoutLM*, *Textract*), and image classification (*ResNet-50*, *ViT*).

McMaster Center for Software Certification (McSCert)

Research Assistant

Sep 2020 - Aug 2022

Hamilton, Ontario

- Investigated, designed & implemented a library deprecation detection tool by applying sentiment analysis (*BERT*, *Word2vec*) on GitHub & Stack Overflow Q&As.
- Investigated, implemented & validated vulnerability discovery methods for source code security analysis using Bidirectional Long-Short Term Memory (*Bi-LSTM*) classifiers based on serialized Abstract Syntax Tree (*AST*) data.

Sharif University of Technology

Research Assistant

Jan 2018 - Aug 2020

Tehran, Iran

- Participated in detecting *Spectre* attacks by classifying cache calls based on hardware performance counter & kernel event handler using a Multilayer Perceptron (*MLP*) models & Support Vector Machines (*SVM*) with 96% accuracy over the *Spectre-V1* dataset.

EDUCATION

McMaster University

M.A.Sc. Software Engineering

Sharif University of Technology

B.Sc. Computer Engineering

SKILLS & QUALITIES

Programming Languages, Frameworks & Libraries

- Python, Java, C/C++, Bash, Linux, TensorFlow, PyTorch, AWS SageMaker, Azure ML, GCP Vertex AI, NLTK, NumPy, pandas, SciPy, OpenCv, XGBoost, SQL, Apache Spark, Pinecone, LangChain, Git, Docker, Firebase.