

RAJ KRISHNAN VIJAYARAJ

+1 (249) 876-0601 ✧ Toronto, Canada

raj.vijayaraj@mail.utoronto.ca ✧ [linkedin.com/in/raj-krishnan-vijayaraj](https://www.linkedin.com/in/raj-krishnan-vijayaraj) ✧ unni12345.github.io

EDUCATION

University of Toronto

MEng, Electrical and Computer Engineering

2022–2023

GPA: 3.94/4.00

NIT Tiruchirappalli

B.Tech, Electrical and Electronics Engineering

2015–2019

GPA: 8.00/10.00

SKILLS

Python, Scala, C++, SQL, Spark, LangChain, Transformers, PyTorch, Hugging Face, Kubernetes, Docker, Jenkins, KubeRay, GraphQL, REST, AWS, FSDP, LoRA, PCA, RAG, Prompt Engineering

WORK EXPERIENCE

Machine Learning Engineer

Jul 2024 – Present

LexisNexis / Appfabs, Toronto, Canada

- Deployed production-grade SQL generation and RAG agents using Spark, LangChain, Hugging Face
- Automated deployment using Docker, Jenkins, KubeRay; integrated GraphQL and REST APIs with credentialing
- Optimized ETL pipelines, reducing latency by 50% via schema tuning and query design

Applied Research Intern

May – Jul 2024

Vector Institute, Toronto, Canada

- Fine-tuned biomedical LLMs (e.g., BioGPT) and served inference using FSDP across multi-GPU clusters
- Designed a hallucination-resistant diagnostic RAG pipeline achieving sub-1% factual error rate

Machine Learning Engineer (Mitacs Fellow)

Sep 2022 – May 2024

Scribble Data, Toronto, Canada

- Built privacy-aware LLM agents for legal QA and contract redaction with domain-specific masking
- Led low-rank finetuning (LoRA) and quantization benchmarks for secure, cost-efficient inference

Backend Developer

Jul 2019 – Nov 2021

G2, Bengaluru, India

- Engineered GraphQL APIs and real-time ETL flows using Ruby, PostgreSQL, and Sidekiq
- Automated reporting jobs, reducing export latency from 3 days to under 20 hours

SELECTED PROJECTS

- **Embeddings to Diagnosis (KDD 2025)** — Independent research introducing LDFR, a latent fragility metric exposing diagnostic instability in LLMs under structured perturbation
- **Clinical Tiger** — Independently developed a multimodal RAG agent combining GPT-4-Vision and ClinicalBERT with chain-of-verification for clinical QA
- **LLM Infra @ LexisNexis** — Collaborated with infra and product teams to deploy Kubernetes-based RAG/SQL agents with AWS Secrets and automated CI/CD
- **PIIguardLLM @ Scribble** — Co-led design of a quantization-aware legal redaction pipeline using LoRA-finetuned LLMs and rule-based domain masking

PUBLICATIONS

Raj Krishnan Vijayaraj. *Embeddings to Diagnosis: Robustness Evaluation of Clinical Reasoning with Agentic Synthetic Notes.* Agentic & GenAI Eval @KDD, 2025.

Raj Krishnan Vijayaraj, Anand Asokan, Allan J. Pothen. *ARMatron: A Wearable Gesture Glove for Robotic Control.* IEEE RAHA, 2016. [DOI:10.1109/RAHA.2016.7931882](https://doi.org/10.1109/RAHA.2016.7931882).