

Kharanshu Naghera
Machine Learning Engineer

(647) 926-4850 | realknaghera1@gmail.com | [LinkedIn](#) | Toronto, ON

PROFESSIONAL SUMMARY

Machine Learning Engineer with 3 years of hands-on experience in Generative AI, NLP, and Computer Vision. Skilled in developing and deploying end-to-end LLM-driven RAG pipelines, multimodal AI applications, and model safety frameworks. Demonstrated success in improving engagement, ensuring compliance, reducing bias, and scaling ML APIs to deliver impactful and production-ready AI solutions for enterprise environments.

EXPERIENCE

RBC (Royal Bank of Canada) Software Engineer, AI/ML	Jan 2025 - Present Toronto, ON
- Design, develop, and deploy scalable machine learning models for financial risk assessment and credit scoring, improving predictive accuracy by 23% and minimizing model drift across quarterly evaluations. - Engineer data ingestion pipelines using Python, SQL, and Apache Airflow to automate high-volume transactional data processing and reduce manual reporting time by 40%. - Optimize deep learning architectures with TensorFlow and PyTorch, increasing model inference speed and accuracy for customer behavior and fraud detection analytics by 30%. - Collaborate with product managers and data scientists to translate business objectives into AI-driven solutions, ensuring model explainability, compliance, and ethical AI alignment. - Develop and maintain APIs and microservices for model serving and real-time scoring, integrating AI insights into digital banking platforms and customer dashboards. - Conduct feature engineering and A/B testing on complex financial datasets to improve predictive clarity and model interpretability using SHAP and LIME. - Apply natural language processing (NLP) to automate document classification and anomaly detection, improving compliance and audit efficiency by 18%.	

Apex infolab Software Engineer, Applied Machine Learning	Aug 2021 – Jul 2023 Bengaluru, INDIA
- Led design of personalization models for e-commerce pricing (Logistic Regression, XGBoost, Neural Nets) increasing customer engagement 20% and increased new sign-ups 15%. - Built distributed pipelines with Spark, Airflow, Databricks on AWS EC2, processing over 100M+ customer records. - Automated data preprocessing in Python and SQL to label 500K+ text records, supporting ETL for NLP training. - Developed an NLP pipeline with BERT for support ticket classification, cutting manual triage workload 30%. - Integrated Computer vision CNN embeddings from TensorFlow, improving recommendation accuracy 12%. - Deployed ML models on AWS SageMaker; delivered real-time APIs serving 2 million monthly users with 200 ms latency. - Reduced compute costs 25% by automating optimized ML services with Git-based workflows. - Created Power BI dashboards to track metrics, improving ticket resolution visibility.	

PROJECTS

- From-Data-to-Justice:** Offender sex/age prediction on 573K FBI homicide records; 45-combo leak-free oversampling benchmark, SMOTEENN + XGBoost at 0.635 macro-F1 / 0.75 ROC-AUC; FastAPI + Docker on AWS App Runner.
- eBay-iPhone-Price-Prediction:** 30K+ scraped eBay sold listings, 56-fit GPU model sweep, LightGBM at R² 0.94 / MAE C\$62 with conformal bands; FastAPI + Go/ONNX serving on Docker, AWS App Runner, and Kubernetes.
- Global Power-Plant Analysis:** Conducted risk assessment of global power plants using SQL and Power BI, identifying energy production inefficiencies.

TECHNICAL SKILLS

Programming	Python, R, Java, C++, SQL, JavaScript
Software Engineering	SDLC, Agile, Scrum, Code Reviews, Design Patterns, System Design, TDD, SaaS, Git Control
APIs & Services	RESTful APIs, FASTAPI, Django, API Integration, gRPC, Microservices
ML/AI Framework	PyTorch, TensorFlow, LangChain, Hugging Face, Scikit-learn, OpenCV, XGBoost, LightGBM
Deep Learning	LLM fine-tuning, LoRA, Embeddings, Multi-Agent systems, CNNs, Transformers, BERT
MLOps & Cloud	Docker, Kubernetes, Kafka, CI/CD, AWS, GCP, Azure, DevOps, MLflow, MCP Server
Search/GenAI	RAG, Vector Databases (FAISS, Pinecone), Knowledge Graphs

CERTIFICATIONS

[Google Data Analytics](#), [AWS Cloud Management Console](#).

EDUCATION

Trent University, Peterborough, ON Masters in Big Data Analytics	Sep 2023 – Jan 2025 GPA: 3.7 / 4.0
Gujarat Technological University, India Bachelors in Computer Engineering	Jun 2018 – Jun 2022 GPA: 3.2 / 4.0