Areel Khan

LinkedIn   areel.khan.career@gmail.com   GitHub

EDUCATION

University of Waterloo | Bachelor of Mathematics – Statistics | 3.8/4.0 GPA   **Aug 2025**

WORK EXPERIENCE

Software Engineer Intern | Trayd   New York | May 2025 - Aug 2025
• Building an internal admin portal customer success to manage all customer processes independently from engineering.

Software Engineer Intern | Fable Security   San Francisco | Jan 2025 - Apr 2025
• Migrated ETL from **Pandas** to **PySpark** + **Avro**, improving data quality and execution speed + memory usage by ~**80x**.
• Worked with customers to define requirements for exposing raw security event data. Built 1) pipelines in **PySpark** to get the data, 2) **GraphQL** API in **Django** to serve the data, and 3) frontend in **React** to present it.
• Built a **LangGraph** ReAct agent that integrates into internal APIs to generate security training content. Shipped in 2 weeks.
• Built a test env that mirrors prod on a schedule while anonymizing PII, cutting testing time and accelerating releases.

Lead Machine Learning Engineer | Health Canada   Toronto | Sep 2023 - Jan 2025
• Led the development of a full-stack RAG chatbot in LangGraph, Postgres and Milvus. Implemented LLM red-teaming and guardrails with Promptfoo, monitoring with Prometheus and Grafana, and deployment on Azure with K8s and Nginx.

Software Engineer Intern | Replicant   Remote | May 2024 - Aug 2024
• CTO and founder: "You're the **most impactful intern we've ever hired**". Hooray! Nominated for intern of the year award.
• Geo-sharded a critical **Postgres** DB and added data archiving, reducing query latency by **300%** and storage cost by **70%**.
• **Won Hackathon.** Implemented an enterprise search tool now used by **90%** employees daily. Deployed on **GCP** with **Helm**.
• Built a CI bot with **Flask** and **webhooks** to enforce security checks on **100%** of PRs, improving QA and ISO compliance.

Software Engineer Intern | Hopper   Remote | May 2023 - Sep 2023
• Introduced data quality testing in **Airflow** and alerting in **Slack** & **PagerDuty**, improving error detection rates by **5%**.
• Built a **SQL** model to predict the home country of users, enabling targeted marketing and preventing promotion scams.

Product Engineer | Bamboo Housing   Remote | Jun 2022 - Sep 2023
• One of five founding members running Bamboo Housing, a fast-growing student housing start up with **over 10k users**.
• Generated **$15,000** revenue by developing web features in **Node** and **React** enabling users to boost postings for money.

Analytics Engineer Intern | Super   Remote | Sep 2022 - Dec 2022
• Built an ETL pipeline in **Airflow** and **dbt** to integrate payments into NetSuite, saving the finance team **100+ hrs** per week.
• Enhanced **SQL** testing by introducing CSV data fixtures, boosting coverage and enabling non-technical users to write tests.

Data Science Intern | TD Bank   Remote | Jan 2022 - Apr 2022
• Reduced manual annotation load by **60%** by training ML models on partially labeled data using **Sklearn** and **Pandas**.
• Developed a text augmentation pipeline that **doubles** data size, resulting in up to **17%** increases in model accuracy scores.

TECHNICAL SKILLS

Programming Languages: Python, TypeScript, SQL, C/C++, R.
Cloud: AWS (S3, EC2, CloudWatch), GCP (CloudSQL, Spanner), Docker, K8s, Terraform, Helm, Prometheus, Grafana.
Data Engineering: Airflow, dbt, Spark, BigQuery, Snowflake, Spanner, Dask, PyTorch Dataloaders.
AI & ML: PyTorch, TensorFlow, DVC, Weights & Biases (WandB), LangChain, LangGraph, MilvusDB, ChromaDB.
Web/Other: Django, FastAPI, Node, Next, React, ChakraUI, GraphQL, PostgreSQL, MongoDB, Celery, Redis.

LEADERSHIP & HACKATHONS & PROJECTS

Director of Technology & Lead Engineer | WAT.ai   Waterloo | Aug 2022 – Aug 2024
• Slashed cloud bill **in half** by caching with **Redis** and multiprocessing to **20x** data pipeline speed and reduced ingress by **90%**.
• Building ML infra on a GPU cluster with **Slurm**, **Docker**, **WandB** and **DVC** enabling **500+** students to train models easily.

Hackathons Wins:
• ClimateHack.ai - **1st Place.** Trained the most accurate PV prediction model **in the world** using 600 GB of data. News article.
• Wildfire Hackathon - **1st Place.** Trained a CNN to predict wildfire spread on a very messy dataset in 4 hours. News article.
• Cohere RAG Hackathon - **1st Place.** Built a medical RAG pipeline with tool calling to medical vision models. News article.