

ARSHDEEP SINGH BAMRAH, M.Eng, P.Eng

Toronto, ON | arsh.bamrah27@gmail.com | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL SUMMARY

Senior Data Scientist and AI Engineer with 5+ years of experience delivering machine learning and Generative AI solutions across regulated, enterprise environments. Strong background in GenAI product design, **Agentic AI frameworks**, and Retrieval-Augmented Generation (RAG), with proven ability to translate business problems into production-ready AI systems. Experienced in leading external vendors, shaping AI architectures, and driving adoption of AI products including conversational systems and AI-generated podcasts.

CORE SKILLS

Generative AI & Agentic Systems

- LLMs, Agentic AI frameworks, orchestration, RAG, embeddings, semantic search, Text-to-Speech (TTS).

Machine Learning & Data Science

- Time-series modeling, anomaly detection, forecasting, feature engineering, supervised & unsupervised learning.

Cloud, Data & MLOps

- Azure Machine Learning, Databricks, AWS SageMaker, Docker, Kubernetes, CI/CD pipelines, Snowflake, SQL, PySpark.

Business & Product

- AI product design, solution architecture, vendor leadership, stakeholder engagement, and requirements translation

PROFESSIONAL EXPERIENCE

Sanofi — Data Scientist (Commercial GenAI) - Contract Aug 2025 – Present | Toronto, ON

Focus: Applied Generative AI, agentic workflows, and business-driven data science

Tech: Langchain, LLMs (GPT), RAG, embeddings, Retrieval, TTS, Snowflake, Python

- Defined Agentic AI framework standards, including orchestration logic, **multi-step reasoning**, **RAG** and feedback loops to align technical solutions with business intent.
- Led the solution design and architecture for **Generative AI products**, integrating LLMs, RAG pipelines, and TTS systems to support enterprise decision-making
- Delivered a **GenAI-powered audio insight platform** (LLM-generated scripts + multi-voice TTS), enabling non-dashboard consumption of commercial insights across global stakeholders.
- Shaped **Commercial GenAI strategy** by contributing to high-value use cases (talk-to-data, conversational analytics) and aligning solutions with business priorities and responsible AI guidelines.
- Developed data quality standards, grounding strategies, and validation logic to improve the reliability and traceability of AI-generated outputs.
- Defined **Agentic AI architectural decision records (ADRs)** covering feedback loops, grounding, validation, data quality, and governance to ensure reliable AI outputs in regulated settings.
- Acted as the technical lead bridging business strategy and execution, translating commercial needs into AI-ready problem statements and success metrics.

Sanofi — Reliability AI / Predictive Maintenance Engineer (Manufacturing) - Contract
Aug 2024 – Aug 2025 | Toronto, ON

***Focus:** Enterprise PdM AI architecture, regulated manufacturing environments*

***Tech:** Python, PowerBI, SQL, Snowflake, BERT, Time-Series ML, IIoT, SAP PM*

- Led vendor collaboration and architecture decisions to design and deploy **predictive maintenance** ML workflows using time-series sensor data for anomaly detection across critical manufacturing assets.
- Owned solution design for **Snowflake** SQL feature pipelines integrating **IIoT signals**, alarm histories, and SAP PM data to enable downtime risk scoring and maintenance prioritization.
- Built a **BERT-based NLP pipeline** to extract FMEA-relevant entities from maintenance logs, **reducing manual analysis by ~80%** and accelerating failure mode identification.
- Translated ML outputs into actionable maintenance strategies by partnering with reliability and operations teams to define interventions, thresholds, and scale PdM across assets.
- Introduced **Power BI reliability dashboards** for FFIP (MTBF, MTTR, alarm rates, risk scores) and delivered weekly insights to **senior leadership** to drive data-driven decisions.

Apotex — Reliability Data Scientist
May 2021 – Aug 2024 | Toronto, ON

***Focus:** Predictive maintenance, equipment reliability modeling, industrial data science*

***Tech:** Python, Azure ML, Azure WebApp, Docker, XGBoost, BERT, GPT/BERT embeddings, SQL, IIoT*

- Designed full-lifecycle Azure ML pipelines for data ingestion, feature engineering, model training, and controlled production deployment.
- Built a transformer-based semantic search engine using GPT/BERT embeddings, significantly reducing engineering research time.
- Developed and deployed an XGBoost-based anomaly detection model with 95% accuracy for real-time process deviation alerts.
- Implemented Docker and CI/CD retraining pipelines to ensure the reliability and maintainability of deployed AI systems.

TECHNICAL PROJECTS

Landlord-Bot: Agentic AI Property Management System

- **Agentic Framework Design:** Engineered an autonomous AI agent capable of multi-step reasoning and planning to assist landlords with operational tasks and tenant interactions.
- **RAG & Contextual Embedding:** Implemented advanced Retrieval-Augmented Generation (RAG) using custom embedding techniques to ensure high-accuracy responses based on specific lease agreements.
- **Production Readiness:** Developed the system with a focus on reliability and safety, utilizing GitHub for version control and structured change management.

EDUCATION

Master of Engineering (M.Eng), Machine Learning — University of Toronto

Bachelor of Engineering (B.Eng), Mechanical Engineering — Ryerson University