

Khush Jani

Toronto, ON | khush-jani.work | khushjani971@gmail.com | +1 (438)-464-0015 | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL SUMMARY

Data and AI Engineer with a **Master's in Applied Computer Science** and **3+ years** across data engineering, machine learning, and GenAI systems, with production pipelines processing billions of records, AI agents saving 10+ hours weekly, and ML models trusted for business decisions.

SKILLS

- **Programming & Analytics:** Python, SQL, Java, Scala, TypeScript, C++, Power BI, Tableau
- **Data Engineering:** Spark, PySpark, Kafka, Airflow, ETL/ELT Pipelines, Data Modeling, Query Optimization
- **Cloud Platforms:** AWS (S3, Glue, Redshift, SageMaker), GCP (BigQuery, Vertex AI), Azure, Snowflake, Databricks
- **AI/ML & GenAI:** LLMs, Agentic AI, RAG Pipelines, LangGraph, XGBoost, TensorFlow, PyTorch, SHAP
- **MLOps & DevOps:** Docker, Kubernetes, MLflow, GitHub Actions, Terraform, CI/CD, Grafana

PROFESSIONAL EXPERIENCE

Associate, PwC

September 2025 – Present

- Led the deployment of a **GenAI-powered reporting automation system on Azure**, integrating LLM-based content generation across data pipelines to eliminate manual inputs for 17+ Project Managers and save **10+ hours/week**.
- Automated legacy Informatica mapping analysis and SQL artefact generation for a **Tier-1 Canadian bank** using Hypertool AI, cutting manual effort by **98%** with production-ready, fully traceable data pipelines at scale.

Achievements: Boosted data throughput **35%** and reporting accuracy **25%** through modernizing legacy MRP infrastructure with AI-driven automation across AWS and Azure.

Machine Learning Engineer, BRAINYBEAM TECHNOLOGIES PVT. LTD.

January 2022 – December 2022

- Engineered scalable ML pipelines for recommendation and predictive models using **Python, Spark, and XGBoost**, implementing CI/CD workflows and A/B experiment tracking for dependable model deployment and evaluation.
- Streamlined model retraining, drift detection, and performance monitoring using **MLflow and Dockerized** microservices, strengthening data preparation and inference performance across live environments.

Achievements: Released **8+ production-grade ML pipelines**, reducing model deployment lifecycle time by **40%** and enabling reliable low-latency predictions that drove analytics-driven decisions across enterprise clients.

Data Analyst, ECUBIX

June 2021 – November 2021

- Directed SQL-based ETL pipelines and conducted analytics on high-volume datasets, applying advanced statistical methods to uncover trends, detect anomalies, and support evidence-based strategic planning.

Achievements: Delivered actionable insights through robust data pipelines and visualization solutions, driving a 20% increase in customer acquisition and a 15% revenue boost by enabling timely interventions.

Python Developer Intern, BASCOM BRIDGE EDUCATION PVT. LTD.

January 2020 – December 2020

- Developed automation scripts for high-volume data processing tasks, reducing operational overhead by 32%.

Achievements: Optimized Python-based solutions for complex workflows, accelerating project delivery by 10%.

EDUCATION

Master of Applied Computer Science, CONCORDIA UNIVERSITY

Montreal, Canada

Bachelor of Computer Engineering, GUJARAT TECHNOLOGICAL UNIVERSITY

Gujarat, India

PROJECTS

Real-Time Stock Market Data Pipeline | Python, Kafka, AWS (EC2, S3, Glue, Athena)

- Built a real-time streaming pipeline on AWS, boosting analysis efficiency by 50% on high-throughput market data.

GenAI Agentic Customer Support System | Python, LangGraph, Vertex AI, Docker

- Orchestrated a multi-agent RAG system on GCP with supervisor coordination and full observability.

CERTIFICATIONS

- AWS Certified Data Engineer | GCP Professional ML Engineer | Databricks Certified Data Engineer