# Dennis Darko

**Machine Learning & AI Engineer | MLOps Specialist**
North Vancouver, BC, Canada · +1604 977 2276 · dennisdarko0909@gmail.com · LinkedIn · GitHub

## Professional Summary

Certified Google Cloud Professional Machine Learning Engineer with over seven years of experience designing, deploying and maintaining scalable AI and machine-learning solutions across cloud environments. Specializes in generative AI, natural language processing and large language model operations (LLMOps), with hands-on expertise in building retrieval-augmented generation applications using LangChain, vector databases and Hugging Face transformers. Proven track record of translating business requirements into data-driven solutions that increase revenue, reduce costs and accelerate delivery. Skilled in MLOps practices (CI/CD, MLflow, DVC), data engineering and cross-cloud deployments (Azure ML Studio, AWS SageMaker, Vertex AI), and adept at collaborating with cross-functional teams to deliver ethical and reliable AI systems.

## Technical Skills

- **Programming & Data:** Python, SQL, Bash, LookML, basic JavaScript
- **Machine Learning & Deep Learning:** scikit-learn, XGBoost, TensorFlow, PyTorch, CatBoost, OR Tools, DistilBERT, Hugging Face, Keras, LangChain
- **Generative AI & NLP:** Large Language Models (LLMs), transformer fine-tuning, prompt engineering, retrieval-augmented generation (RAG), LangChain, Hugging Face Transformers
- **MLOps & Data Engineering:** MLflow, DVC, Airflow/Cloud Composer, Vertex AI, Azure ML, AWS SageMaker, Docker, Kubernetes, Bitbucket Pipelines, Azure DevOps, Git
- **Data Engineering & Warehousing:** ETL, data pipelines, data modeling, BigQuery, Redash, MySQL, PostgreSQL, Dataflow
- **Monitoring & Logging:** Grafana, Prometheus, Azure Monitor, Cloud Logging
- **Cloud Platforms:** Google Cloud Platform (BigQuery, Dataflow, Cloud Functions, Cloud Run, Vertex AI), AWS, Azure
- **Tools & Collaboration:** Jira, Confluence, Slack, Looker, Tableau, Matplotlib, Plotly, Notion, Bitbucket

- **Soft Skills:** Cross-functional collaboration, project management, mentorship, problem solving, communication

## Experience

### Machine Learning Engineer – GoMaterials, Vancouver, BC (Feb 2025 – Dec 2025)

- Developed **TransCostML**, a CLI pipeline that extracted delivery data, performed preprocessing and trained ensemble and linear regression models (Random Forest, XGBoost, stacking), improving transportation cost prediction accuracy by ≈18 % versus baseline.
- Built robust ETL pipelines using Python, Pandas, scikit-learn and SQL; integrated with Redash and MySQL to pull production data and ensure data quality.
- Introduced MLflow for experiment tracking and model registry and adopted DVC for data version control; configured Azure Blob Storage as the remote store, enabling reproducibility and collaboration.
- Implemented CI/CD pipelines in Azure DevOps and Bitbucket to automate testing, model training and Docker deployment; orchestrated releases to Azure ML Studio, reducing deployment time and errors.
- Co-developed a markup optimization system: trained conversion and markup optimization models using CatBoost and Random Forest; built an inference API with FastAPI; increased conversion rates by recommending optimal e-commerce markups.
- Collaborated with product managers, data scientists and engineers to align models with business objectives; documented pipelines in Confluence and managed tasks in Jira.

### Machine Learning Operations (MLOps) Engineer – Vosyn Inc., Etobicoke, ON (Sept 2024 – Present)

- Designed and automated end-to-end data pipelines on Google Cloud Platform for large-scale text-to-speech models, orchestrating training, deployment and monitoring at scale and ensuring high availability and reliability.
- Deployed models using Kubernetes and Vertex AI, reducing infrastructure costs by 30 % and improving deployment efficiency by 50 %.
- Integrated BigQuery and Looker to develop real-time reporting pipelines, providing stakeholders with actionable insights.
- Improved CI/CD workflows using Google Cloud Build; accelerated deployment cycles by 40 %.

### Machine Learning Engineer – Sinewy Technologies, Kumasi, Ghana (Dec 2021 – Aug 2023)

- Built scalable data-processing pipelines and models on Google Cloud Platform to power real-time ad targeting and recommendation systems; boosted click-through rates by 25 %.
- Developed predictive-maintenance models using Google Cloud IoT Core and BigQuery ML, saving US$500 K annually in operational downtime.
- Automated ETL pipelines using Dataflow and SQL; reduced time to insight by 40 % and improved reliability across departments.

### Software Engineer (LLM & Backend Infrastructure) – Sinewy Technologies, Kumasi, Ghana (Oct 2020 – Dec 2021)

- Architected and deployed scalable cloud-native solutions using Kubernetes and Looker, increasing business-process automation by 50 %.
- Implemented ETL processes and data-modelling frameworks to streamline data extraction from multiple sources and improve integration across business functions.
- Enhanced system reliability through automated monitoring and error detection, reducing downtime by 20 %.

### Programmer (Backend Developer) – Ahafo Regional Coordinating Council, Goaso, Ghana (Sept 2019 – Oct 2020)

- Developed event-driven integration pipelines with Google Cloud Functions and Eventarc, reducing data-synchronization latency.
- Implemented workflow orchestration using Cloud Composer to automate complex tasks and minimize operational errors.
- Built Slack bots integrated with Dialogflow to automate task notifications, improving team productivity by 50 %.

## Projects

| Project | Summary |
|---|---|
| **TransCostML (GoMaterials Transport Price Estimation)** | Python, scikit-learn, XGBoost, MLflow, Azure. Developed a CLI pipeline that extracts delivery data, preprocesses it and trains global cost ensemble and linear regression models using Random Forest and XGBoost with stacking; organized the codebase into modular ETL, preprocessing, training and MLflow tracking components and deployed the model to Azure ML Studio's production environment. |
| **Gomat Markup Optimization** | Python, CatBoost, FastAPI, MLflow. Built conversion probability and markup optimization models to recommend optimal selling markups; applied data balancing, feature engineering and synthetic data generation; exposed an inference API with FastAPI; containerized the application and integrated CI/CD and MLflow tracking; deployed via Azure ML Studio. |
| **GoSource Routing Optimization** | Python, OR Tools, Flask. Developed a proof-of-concept route optimization model to determine optimal vehicle routes among suppliers; built CLI tools to parse JSON inputs and compute optimal routes; wrote QA scripts and integrated Slack notifications. |

| Project | Summary |
|---|---|
| **Ads Recommendation System** | GCP Vertex AI, BigQuery, Docker. Designed a real-time ads recommendation system using Vertex AI and BigQuery, increasing click-through rates by 35 %; deployed personalized models via Docker and Cloud Run. |
| **Participedia Capstone Project** | DistilBERT, Kubernetes, DVC, Vertex AI. Developed a multi-task learning pipeline to analyze participatory democracy data, generating embeddings and classifications; automated deployment with Vertex AI, Kubernetes and DVC. |
| **Loan Approval Prediction** | Python, scikit-learn. Developed classification models (Logistic Regression, SVM, Decision Tree, Random Forest, Gradient Boosting) for loan approvals; handled missing data and outliers; achieved an F1-score of 0.947. |

## Education & Certifications

- **Master of Professional Studies in Analytics**, Northeastern University – *Graduated Dec 2024*
  Concentration: Applied Machine Intelligence
- **Bachelor of Science in Information Technology**, University of Education, Winneba, Ghana – *2017*
- **Certifications:** Google Cloud Professional Machine Learning Engineer; Google Cloud Skill Badge – Building and Deploying Machine Learning Solutions on Vertex AI

## Professional Development & Additional Skills

- **Large language models & generative AI:** Experience with GPT, DistilBERT and fine-tuning transformer models; building retrieval-augmented generation pipelines using LangChain and vector databases (e.g., FAISS) for context-aware applications.

- **LLMOps & evaluation:** Design RLHF pipelines, synthetic data generation and evaluation frameworks to monitor model hallucinations, factuality and bias, ensuring ethical AI practices.

- **Cloud deployments & monitoring:** Deploy models across Azure ML Studio, AWS SageMaker, Google Vertex AI and Cloud Run; implement CI/CD workflows and monitoring tools like Grafana and Prometheus to ensure reliability and compliance.

- **SQL & data engineering:** Advanced SQL, ETL and data warehousing using BigQuery, Redash, MySQL, PostgreSQL; design and optimize data pipelines for scalable

machine-learning applications.

- **Project management & collaboration:** Proficient in Jira, Confluence, Notion, Slack and Git; experienced in Agile methodologies and cross-functional team leadership.

## References

Available upon request.