Gagandeep Singh Thind

Phone: (825) 962-9131 | Email: [gagandeeps.thind@gmail.com](mailto:gagandeeps.thind@gmail.com) | LinkedIn: <https://www.linkedin.com/in/gagandeeps-thind> | Portfolio: <https://gagandeeps-thind.github.io>

**Professional Summary**

Result Driven Data Scientist with hands-on **machine learning & statistics**, **model governance** (MRM, documentation, audit-readiness), and **model monitoring/drift** experience. Built production analytics/ML systems on **Azure Databricks + Spark**, and delivered **explainable, stakeholder-ready insights** through dashboards and clear narratives. Comfortable partnering across business, control, and engineering teams to **independently challenge** approaches, align to **enterprise standards**, and support **ethical, compliant** model use across the lifecycle.

**Technical Skills**

* Languages and Tools: Python (Pandas, NumPy, Scikit-learn), SQL, SAS, Scala, R, PySpark, PyTorch, TensorFlow, Keras, REST APIs
* Machine Learning & AI: Predictive Modeling, Statistical Modeling, Deep Learning (CNNs, RNNs/LSTMs, Transformers), Prompt Engineering, NLP (LLMs, BERT, FinBERT, LangChain, HuggingFace), Copilot Studio/Azure AI Studio, Dataverse, SharePoint, ServiceNow, Retrieval-Augmented Generation (RAG), Hyperparameter Tuning, Model Monitoring, Drift Detection
* Big Data & Cloud Platforms: Azure (Databricks, Azure ML, ADLS Gen2, Event Hubs, Docker), AWS (SageMaker, Lambda, Glue, S3), GCP (BigQuery), Alteryx Designer, Apache Spark, Apache Kafka, Apache Airflow, Github Actions, Delta Lake, PostgreSQL, MySQL
* Data Engineering & MLOps: ETL/ELT, CI/CD pipelines, MLflow, Data Wrangling, Feature Engineering, Model Monitoring & Validation
* Reporting & Visualization: Tableau, Power BI, Matplotlib, Spreadsheets(Excel), Executive Reporting, Presentation, Model Documentation

**Education**

* **Master of Data Science and Analytics**, University of Calgary - Sep 2024 - Aug 2025 (GPA: 3.9/4.0)  
  *Relevant Coursework:* Developing Big Data Applications; Statistical Modelling with Data; Working with Data at Scale
* **Bachelor of Technology (Computer Science)**, Gyan Ganga College of Technology (Jabalpur, India) - Jul 2019 - Jun 2023 (GPA: 8.2/10)  
  *Relevant Coursework:* Machine Learning; Cloud Computing; Database Management Systems; Analysis and Design of Algorithms

**Experience**

**Data Scientist Research Assistant**, University of Calgary (Calgary, AB) - May 2025 - Aug 2025

* Architected QuantEdgeX, a real-time stock prediction platform on Azure, achieving 70-85% directional accuracy for stock price forecasts with millisecond latency to support high-frequency trading decisions.
* Productionized a hybrid SARIMAX + XGBoost model via an Azure ML model registry and CI/CD pipeline, boosting trading signal accuracy to 85% by optimizing feature engineering and hyperparameter tuning.
* Implemented an automated MLOps pipeline on Databricks to ingest over 1 GB of daily streaming market data for continuous model training and inference.
* Documented model development and validation processes according to industry Model Risk Management guidelines, ensuring transparency, compliance, and readiness for audits or regulatory review.
* Developed an interactive Streamlit dashboard integrated with a fine-tuned LLM Gemini API chatbot for real-time forecast visualization and conversational explainability.
* Enabled traders and analysts to query model insights in plain English via the chatbot interface, powered by a reinforcement learning agent generating trading signals.

**Data Analyst Intern**, Alteryx (India) - Dec 2022 - Mar 2023

* Optimized ETL workflows in Alteryx Designer, cutting data processing time by 30% and improving reporting frequency from weekly to daily for faster business intelligence in a global team setting.
* Automated reporting pipelines for 5+ cross-functional teams, reducing manual effort by over 20 hours per month and enabling quicker access to key insights for non-technical stakeholders across departments.
* Developed and validated complex SQL queries to improve data quality and consistency, achieving over 99% accuracy in dashboards and ensuring reliable data for executive decision-making and audit compliance.

**Projects & Contributions**

* Canadian Housing Price Predictor & Market Analysis: Developed a tool to forecast housing prices across major Canadian markets. Scraped 50,000+ property listings using BeautifulSoup and Selenium, and trained an XGBoost regression model that achieved an RMSE 12% lower than the baseline. This project provided more accurate pricing insights to inform real estate investment decisions.
* Customer Churn Prediction: Built an end-to-end machine learning pipeline for a Canadian telecom to predict customer churn, achieving an AUC of 0.92 (92%). Engineered new features (e.g., contract duration trends, support call frequency) that improved model performance by 8%, and performed extensive exploratory data analysis on 70,000+ customer records to identify key churn drivers.
* Real-Time Fraud Detection System: Designed and deployed a deep learning solution to detect fraudulent credit card transactions in real time. Trained an LSTM autoencoder in TensorFlow on a highly imbalanced transaction dataset to identify anomalous spending patterns. Deployed the model as a Flask REST API, capable of returning fraud probability scores in under 50 ms, enabling instant transaction flagging. Demonstrated robust handling of sensitive financial data and integration of AI models into fraud prevention workflows.