Gagandeep Singh Thind

Phone: (825) 962-9131 | Email: 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 about 2 years of experience in **predictive modelling, statistical analysis, data engineering, and datadriven decision-making**. Skilled in **Python, R, SQL, and cloud platforms (AWS, Databricks)** to design and implement end-to-end analytical and data pipeline solutions. Experienced in building scalable ETL workflows, managing large datasets, and developing machine learning models that deliver actionable insights. Proficient in creating interactive dashboards, translating business needs into data strategies, and collaborating with cross-functional teams. Passionate about retail analytics and applying data science to solve realworld business challenges.

Technical Skills

- **Languages and Tools**: Python, R, SQL, SAS, R, PySpark, PyTorch, TensorFlow, Keras, REST APIs, Tableau, Power BI
- **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, Retrieval-Augmented Generation (RAG).
- **Big Data & Cloud Platforms**: Azure (Databricks, Azure ML, ADLS Gen2, Event Hubs, Docker), AWS (SageMaker, Lambda, Glue, S3), GCP (BigQuery), Alteryx Designer, Apache Spark, Kafka, Airflow, GitHub Actions, NoSQL

Soft Skills: Stakeholder Management, Analytical problem solving, Critical Thinking, Project Management, Attention to Detail

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
Projects:

- **Customer Churn Prediction:** Developed an end-to-end ML pipeline for a telecom dataset (70K+ records) achieving 0.92 AUC, with engineered features boosting performance by 8% and uncovering key churn drivers.
- **Real-Time Fraud Detection System:** Deployed an LSTM autoencoder via Flask API (<50 ms latency) to flag anomalous credit-card transactions in real time, enhancing fraud detection accuracy and financial data security.

Bachelor of Technology (Computer Science) | Gyan Ganga College of Technology (GPA: 8.2/10)   **Jul 2019 - Jun 2023**
Relevant Coursework: Machine Learning; Cloud Computing; Database Management Systems; Analysis and Design of Algorithms

EXPERIENCE

University of Calgary
Data Scientist Research Assistant   **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.
- Developed an interactive Streamlit dashboard integrated with a fine-tuned LLM Gemini API chatbot for real-time forecast visualization and conversational explainability.

Bureau Veritas
Graduate Engineer Trainee (GET) – Risk Analyst   **Jun 2023 – Jun 2024**

- Collaborated with stakeholders to understand data requirements and deliver insights that align with business objectives, leading to a 100% client satisfaction.
- Applied Python and SAS for predictive modeling, anomaly detection, and data analysis to optimize client data-driven decisions.
- Digitalized risk assessments and audits for Fortune 500 companies, increasing scope and reducing manual effort by 50%.
- Developed and implemented statistical analysis, hypothesis testing, and forecasting models to enhance targeting and performance measurement, aligning with business goals.

Alteryx
Data Analyst Intern   **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.

ADDITIONAL ACTIVITIES

Volunteered and organized Campus Drive and Events in College | State and University Football Captain | Travel and Fitness Enthusiast