

Brian Denis Castelino

Dallas, TX (Open to Relocate) | briancastelino07@gmail.com | +1(571)899-1663 | linkedin.com/in/cas7elino
github.com/bcastelino | Portfolio: brianc.framer.website

Summary

AI & Data Engineer with 5+ years across data and analytics roles, now building production ML platforms, LLM-powered analytics pipelines, and real-time inference systems on Databricks, AWS, and Snowflake. Specialized in MLOps, NLQ-to-BI automation, and applied AI for enterprise-scale analytics. Databricks (Data Engineer + GenAI), SnowPro Advanced, and Google Cloud certified.

Education

George Mason University | MS in Data Modeling/Warehousing and Database Administration | Aug 2023 - May 2025

Professional Experience

AI Data Engineer, WorldLink US

Sept 2025 – Present

- Architected and deployed an end-to-end commodity price forecasting platform on Databricks using **MLflow** and **XGBoost** with full **CI/CD**, cutting deployment cycle time by **40%** across three production model releases.
- Built a scalable ML inference application (**React + Databricks Apps**) supporting **1,000 concurrent users** with less than **2 seconds** latency, enabling real-time enterprise-wide access to predictive insights.
- Implemented automated **MLOps** workflows (model validation, versioning, Dev to Prod promotion), reducing production rollbacks by **30%** and release errors by **45%**.
- Improved forecasting performance through advanced feature engineering and hyperparameter tuning, achieving an **18% reduction** in **MAPE** relative to baseline time-series models.
- Engineered an LLM-powered **NLQ-to-BI** visualization pipeline integrating **Databricks Genie** with **Power BI**, reducing manual report creation time by **60%** and increasing self-service analytics adoption by **35%**.
- Automated **SQL-to-visual** translation and **YAML to TMDL** metadata conversion using **Agent Skills** and **LLM** orchestration, achieving **92%** query-to-visual accuracy and reducing BI configuration effort by **50%**.

Graduate Teaching Assistant, George Mason University – SEOR Dept

Aug 2024 – Jan 2025

- Led weekly sessions for 40+ students on **Excel modeling** using Frontline Solver simulations to increase student accuracy in solving linear and mixed-integer optimization problems by 30%.
- Facilitated interactive labs on **Gurobi in Python**, resulting in a 20% increase in student engagement scores.

Senior Product Analyst, Media.Net Software Services Pvt. Ltd.

Jan 2022 – Aug 2023

- Collaborated with cross-functional teams to plan and deliver **data-driven solutions** to optimize AdTech performance.
- Set up weekly **10+ A/B testing** environments using in-house tools to statistically evaluate creative templates, improving recommendation accuracy by **50%** and increasing campaign **ROI** by **70%**.
- Automated workflows using **Python**, **Excel**, and **SQL**, reducing manual effort by **50%**, managing ambiguity in data.
- Led **end-to-end project** execution across 5+ initiatives using **Jira** and **Confluence**, ensuring **100%** on-time delivery.
- Mentored 10 junior analysts on **ETL** design, **SQL** querying, and business reporting tools, increasing productivity by **40%**.

Product Analyst, Media.Net Software Services Pvt. Ltd.

Jul 2019 – Dec 2021

- Developed **stored procedures** to automate **ingestion** and **transformation** of **5+ GB** of ad-serving data into **SQL Server** using **Python** and **Excel-based automation**, reducing **batch processing** time by **50%**.
- Collaborated with engineering to integrate new ad platforms using **REST APIs**, increasing impression delivery by **20%**.
- Prioritized delivering weekly campaign performance insights and optimization reports to **10+ clients** using **Tableau/Excel**, improving retention by **10%**.

Certifications

Databricks Data Engineer Professional | PySpark, Delta Lake, Lakehouse, Workflows, Production Deployment, Security
Databricks Data Engineer Associate | Data Modeling & Warehousing, ETL, Spark SQL, Delta Tables, SQL Optimization
Databricks Gen AI Engineer Associate | LLMs, RAG, Vector Search, Embeddings, Model Serving, MLflow, Mosaic AI
Google Associate Cloud Engineer | GCP, Compute Engine, IAM, GKE, Pub/Sub, Cloud Storage, Networking, Security
Microsoft Fabric Data Engineer Associate | Lakehouse, OneLake, Data Factory, Data Pipelines, Power BI, Semantic Models
SnowPro@Core Certification | SQL, Data Warehousing, Data Loading, Data Pipelines, Security, Performance Optimization
SnowPro@Advanced: Data Engineer | Snowpark, Streams, Tasks, Dynamic Tables, Governance, Performance Tuning

Skills & Technologies

AI & GenAI: LLMs, RAG, Agentic AI, LangChain, LlamaIndex, LangGraph, Hugging Face, FAISS, Vector Databases, Embeddings, Prompt Engineering, Context Engineering, Tool Calling, LLM Orchestration, Databricks Genie

ML & MLOps: MLflow, Model Registry, Model Serving, Model Monitoring, Feature Engineering, Hyperparameter Tuning, XGBoost, LightGBM, scikit-learn, PyTorch, TensorFlow, NLP, Time Series Forecasting, Batch Inference, Real Time Inference

Data Engineering: Python, SQL, PySpark, Apache Spark, ETL, ELT, Airflow, dbt, Kafka, Beam, Delta Lake, Delta Live Tables, Databricks Workflows, Data Pipelines, Data Quality, Data Warehousing, Lakehouse Architecture

Cloud & Platforms: Databricks, Snowflake, Microsoft Fabric, AWS, GCP, Azure, AWS S3, Glue, Lambda, RDS, SageMaker, Redshift, BigQuery, Vertex AI, Azure Data Factory, Azure Synapse, Fabric Lakehouse

Databases: PostgreSQL, SQL Server, MySQL, Oracle, MongoDB, Hive, Redis, Elasticsearch, NoSQL, Parquet, JSON, Delta Tables, Stored Procedures, Query Optimization, Dimensional Modeling

BI & Analytics: Power BI, Tableau, Looker, Google Data Studio, Excel, DAX, Power Query, Ontology, Semantic Models, TMDL, PBIP, Dashboard Development, KPI Reporting, Self Service BI, Data Visualization, A/B Testing, Statistical Analysis

Governance & Software: Data Governance, Metadata Management, Data Lineage, Data Security, Data Modeling, SQL Optimization, FastAPI, Streamlit, React, REST APIs, Git, Docker, CI/CD, Agile, Jira

Projects

OntoBricks Graph Chat Reliability | Open-Source Contribution

github.com/databrickslabs/ontobricks

- Contributed a production resilience fix to Databricks Labs OntoBricks that prevents Graph Chat from freezing the app by bounding graph query time and result size, offloading blocking reads from the Uvicorn event loop, auto-sizing worker pools, adding saturation telemetry and admin controls, and validating the changes through 41 passing regression tests.

Multi-Model Inference Router

private repository

- Built a FastAPI inference gateway routing supply chain risk predictions across 3+ Databricks Model Serving endpoints, with validation, batching, caching, and failure handling, while supporting 6 integration patterns, including REST, polling, SSE, WebSockets, and Databricks Jobs, and tracking 10K+ events across 4 Unity Catalog Delta tables to improve reliability by 30%, reduce integration complexity by 40%, and increase debugging efficiency by 50%.

Agent Skills Powered Power BI Dashboard Builder

github.com/bcastelino/powerbi-dashboard-generator

- Built an **Agent Skills** powered **Power BI** Dashboard Generator to reduce manual BI development by converting plain English requests into complete **PBIP projects**, using **10 reusable skills**, **2 human approval** gates, and a **4-stage pipeline** for semantic modeling, visual generation, theming, and packaging.

Stock Market Databricks ETL Pipeline

github.com/bcastelino/stock-dash-e2e

- Built Stock Dash end-to-end to demonstrate **production-style** data engineering by ingesting **Alpha Vantage** stock data into **Databricks**, using **PySpark**, Delta Lake, and a **3-layer Medallion Architecture** to create restartable micro-batches, deduplicated Silver tables, and Gold KPI dashboards for 3 tracked stocks.