Dominic Bankovitch

dbankovitch@berkeley.edu | dominicbankovitch.com | linkedin.com/in/dominic-bankovitch/ | github.com/dominic738

Education

University of California, Berkeley — Berkeley, CA
Data Science & Applied Mathematics; Minor: Computer Science; GPA: 4.00/4.00 — *2024 – 2027*

University of Notre Dame — Notre Dame, IN
Applied Mathematics & Statistics; Science Honors Program; GPA: 3.96/4.00; 2x Dean's List — *2023 – 2024*

Technical Skills

Languages: Python, SQL, Java, TypeScript, JavaScript, HTML/CSS
Libraries/Frameworks: pandas, NumPy, scikit-learn, XGBoost, LightGBM, CatBoost, SentenceTransformers, Optuna, PyTorch, TensorFlow, Keras, Flask, FastAPI, React, Tailwind CSS, OpenCV
ML/AI: Classification, Neural Nets, NLP (Transformers, SBERT), Feature Engineering, A/B Testing, RAG
Tools/Databases: Git, Jupyter, GCP, Vertex AI, MySQL, BigQuery, Redis, Upstash, Pinecone, Docker, Tableau, Qlik

Professional Experience

Data Science Intern — Summer 2025
Kohl's — *San Francisco, CA*

- Imputed e-commerce attributes across **6M+ products**, enriching inputs for downstream models.
- Built binary gender classifier (XGBoost, CatBoost, LightGBM), achieving **0.98 AUC** with group CV.
- Extended to multi-label gender prediction via One-vs-Rest LightGBM (**0.86 Jaccard**, **0.87 F1** across 6 classes).
- Embedded **3.5M+ product titles** with SBERT + PCA to encode semantic hierarchy for imputation.
- Trained Bayesian-optimized Keras neural network on hierarchy embeddings to classify **2,018** e-commerce paths.
- Achieved **96% Top-5 accuracy** on hierarchy prediction; deployed results to BigQuery via scalable batch pipeline.
- Processed **22M+ SKUs** with GPU-parallelized pipelines, optimizing joins/chunked computation for scalability.
- Designed loyalty features from VOC insights (savings tracking, tiered incentives).
- Built Qlik Sense dashboard to track return forecast accuracy, merging model outputs with actuals.
- Automated pipelines via Qlik SaaS scheduler ensuring continuous KPI access.

Consulting Extern – Deloitte — Spring 2024
Student International Business Council — *Notre Dame, IN*

- Collaborated on AV market simulation, researching adoption strategy, sensor tech, and competitors.
- Presented final analysis to Deloitte, informing their view of the AV industry and key decisions.

Project Leader — Spring 2024
Quant Club — *Notre Dame, IN*

- Developed automated MLP prop odds scraper across 5 sportsbooks, analyzing **70+** lines daily.
- Computed EV and detected mismatches, identifying arbitrage opportunities weekly.

Markets Intern — Summer 2022
X-Change Financial Access — *San Francisco, CA*

- Applied options strategies (straddles, spreads, volatility plays) and drafted trade tickets to support floor decisions.
- Observed and assisted traders, analyzing live trades and providing real-time insights to guide execution.

Selected Projects

Spotify Lyrics Sentiment Analysis | *FastAPI, SBERT, Groq API* — Present

- Built Python pipeline projecting songs onto semantic axes using SBERT lyric embeddings.
- Integrated Genius & Spotify APIs with regex cleaning to preprocess lyrics.
- Enabled user-defined semantic axes enriched via generative AI prompting (Groq API).
- Deployed FastAPI site via Railway + Netlify for interactive visualization, driving richer insights into songs.

Spotify Playlist Analyzer | *Flask, MySQL, Spotify API* — Present

- Developed Flask web app to maintain a continuously growing database of **6,000+ Spotify tracks**.
- Implemented Spotify API integration to auto-fetch metadata and detect missing album tracks.
- Designed SQL queries to uncover insights, such as best N-track runs and albums with the most rating variance.

Kaggle Competitions — 2024 – Present

- Top **5%** in Titanic Survival Prediction competition with feature-engineered XGBoost models.