Gregory Burns

burnsgregm@gmail.com | 831-226-6907 | San Francisco Bay Area
linkedin.com/in/gregburns | github.com/burnsgregm | burnsgregm.netlify.app/portfolio.html

FraudShield V3 — One-Page Summary

Author: Greg Burns

Role: Machine Learning Engineer • MLOps Builder

Tech Stack: Python • GCP (Pub/Sub, Dataflow, Feature Store, Vertex AI) • FastAPI • XGBoost • Isolation Forest • Terraform • BigQuery

Overview

FraudShield V3 is a cloud-native, real-time fraud detection system built on Google Cloud Platform. It combines streaming feature engineering, a hybrid supervised + unsupervised model, and a low-latency online scoring API to identify anomalous or fraudulent card transactions at scale.

Problem Addressed: Financial institutions require accurate, millisecond-level fraud detection that adapts to emerging fraud patterns without human intervention.

Key Capabilities

Hybrid Fraud Model (CPR Architecture)

- XGBoost: supervised fraud probability
- Isolation Forest: unsupervised anomaly detection
- Weighted ensemble: unified fraud score
- Supports versioning, rollback, and endpoint-level traffic control

Real-Time Streaming Features

- Raw transactions ingested through Pub/Sub
- Dataflow streaming pipeline with 10-minute event-time sliding windows
- Velocity-based aggregates per tenant_id and card_id
- Real-time features stored in Vertex AI Feature Store
- Enriched events archived to BigQuery for analytics and retraining

Online Scoring API (FastAPI)

- POST /v3/score
- Retrieves live features from Feature Store
- Sends full feature vector to Vertex AI Endpoint
- Returns fraud score + component scores + metadata

MLOps & Infrastructure

- Terraform provisioning of all GCP components
- Model artifacts stored in Cloud Storage
- Hybrid CPR deployed to Vertex AI Endpoint
- BigQuery-based monitoring: predictions, drift, feature snapshots

Business Impact

FraudShield V3 reduces false positives, improves detection of novel fraud patterns, and enables real-time, high-throughput fraud decisioning for financial systems. It delivers faster investigations, richer observability, and production-ready MLOps practices suitable for enterprise fintech, banking, and payments platforms. This project demonstrates my ability to architect, deploy, and operationalize end-to-end machine learning systems on GCP.