Melissa Wong (mwong3)

Real-Time Crypto Volatility Detection – Scoping Brief

Use Case & Business Context
Cryptocurrency markets are highly volatile. This project builds a real-time pipeline to detect short-term volatility spikes, enabling:

- **Risk Management**: Early alerts for traders and risk teams
- **Trading Triggers**: Signals for algorithmic strategies
- **Market Surveillance**: Live monitoring of market conditions

Target Users: Retail/institutional traders, risk teams, market makers
Data Source: Coinbase Advanced Trade WebSocket API (public data only)

Prediction Objective
Predict if a volatility spike will occur in the next **60 seconds** using real-time BTC-USD ticker data.

- **Spike Definition**: Rolling std. dev. of midprice returns ≥ 90th percentile threshold (taken from EDA)
- **Features**: Log return volatility, return statistics, bid-ask spread, trade intensity, rolling stats (30s/60s/300s)
- **Output**: Binary classification (0 = normal, 1 = spike)

Success Metrics

- **Primary**: PR-AUC ≥ 0.60 (handles class imbalance)
- **Secondary**:
    - Precision > 20% (minimize false alarms)
    - Recall > 50% (capture majority of true spikes)
    - Latency < 120s (2x real-time)

Risks & Constraints

- **Technical**: WebSocket drops, API latency (<100ms), feature drift (weekly retraining)
- **Data**: Tick rate ~1–10/sec; time-based train/test split
- **Operational**: No trade execution; real-time Kafka pipeline; reproducible via replay script
- **Modeling**: Baseline Z-score rule; Logistic Regression (primary), XGBoost (optional); walk-forward validation