QCoin IV Forecasting – 10s Ahead (Kaggle)

Author: Aman | Date: 2025-08-14

Objective

Predict QCoin implied volatility 10 seconds ahead from high-frequency order book data.

Data & Features

Mid/microprice, bid–ask spread, order book imbalance, depth imbalance (L1–L10), returns, realized-volatility proxies, lags (1/2/3/5/10/20s), rolling stats (5/10/20/60s), and Rolling OFI (10-window).

Validation & Metric

TimeSeriesSplit (5 folds, expanding window). Pearson correlation (primary) and RMSE (secondary).

Model

LightGBM with original parameters preserved. Final model retrained on full train set after CV.

Results Artifacts

OOF Pearson/RMSE printed in notebook. `feature_importance.png` and `oof_true_vs_pred.png` are saved by the notebook.

Trading Considerations

Latency: precompute rolling features; batch predictions. Execution: use microprice for quoting; manage slippage. Risk: scale position sizes with predicted volatility; cap exposure on spikes.

Next Steps

Add peer-asset signals; try shallow Transformer on LOB; add uncertainty calibration (quantile/conformal).