Trading Competition Report

Generated on 2026-02-19 18:18

Executive Summary

- The pipeline is submission-ready and produces a valid file with 13,098 rows.
- Rolling time validation (3 recent monthly folds) gives MAE 15.65 and RMSE 39.83.
- Main observation: error increases in later folds (possible regime shift).
- Model used: HistGradientBoostingRegressor with robust loss (absolute error).

What Was Done

1. Loaded train/test data and parsed delivery timestamps.
2. Engineered temporal, cyclical, and weather interaction features.
3. Applied robust preprocessing: median numeric imputation + one-hot encoding for market.
4. Validated with rolling month-based folds to simulate real forecasting.
5. Trained on full training data and generated final submission predictions.

Validation Metrics

Fold	Train Rows	Val Rows	MAE	RMSE
2025-06	119,360	4,320	12.1252	28.4339
2025-07	123,680	4,464	16.3345	45.9321
2025-08	128,144	4,464	18.4967	45.1217
Average	**-**	**-**	**15.6521**	**39.8293**

Charts

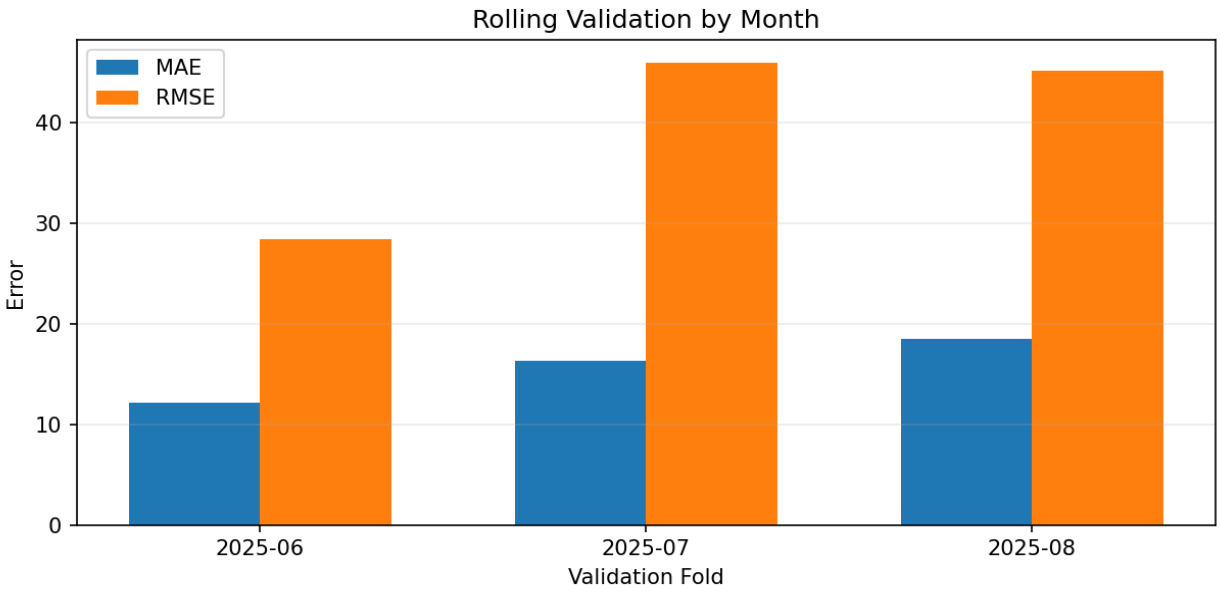

Figure 1. MAE and RMSE for each rolling validation fold.

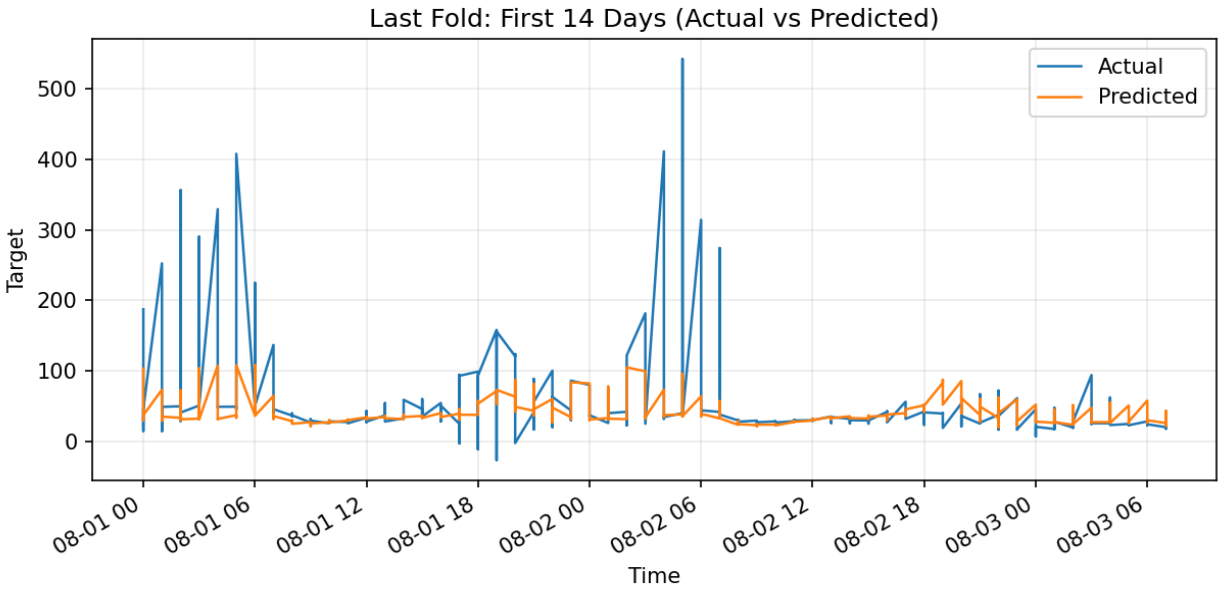

Figure 2. Last fold time-series comparison between actual and predicted values.

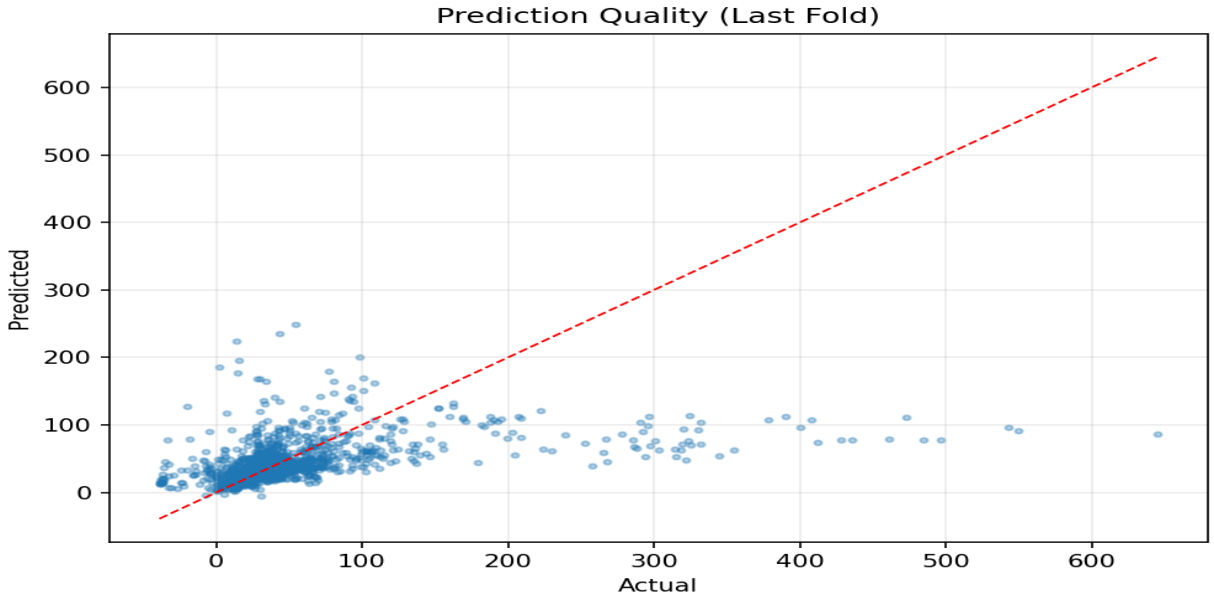

Figure 3. Scatter of actual vs predicted values in the last fold (red line = perfect fit).

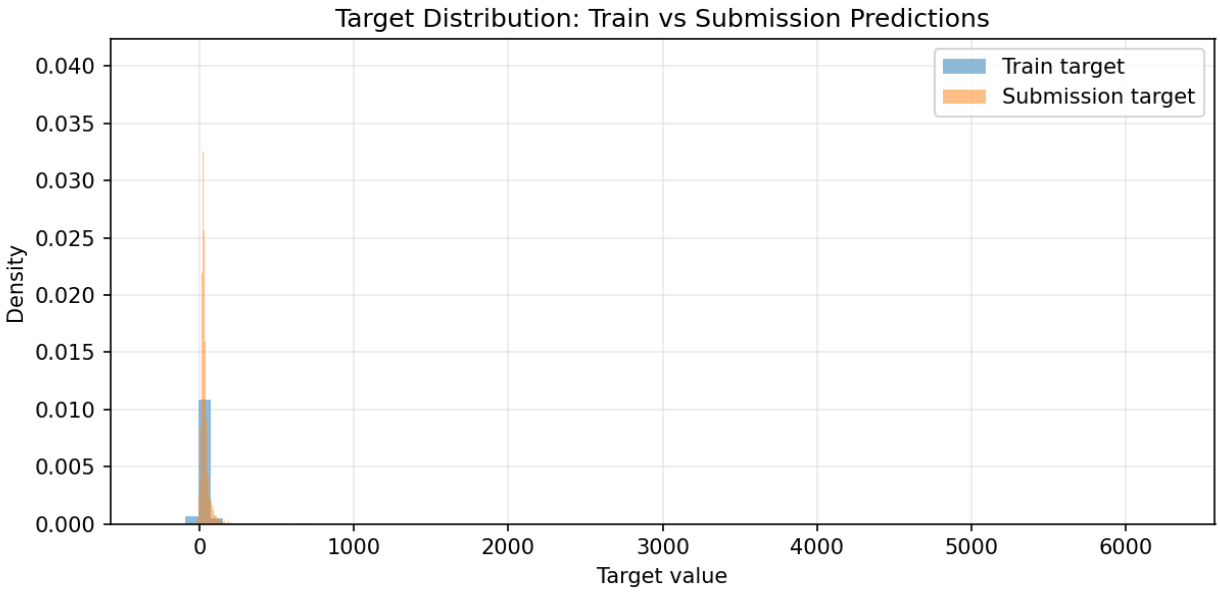

Figure 4. Distribution check: train targets vs final submission predictions.

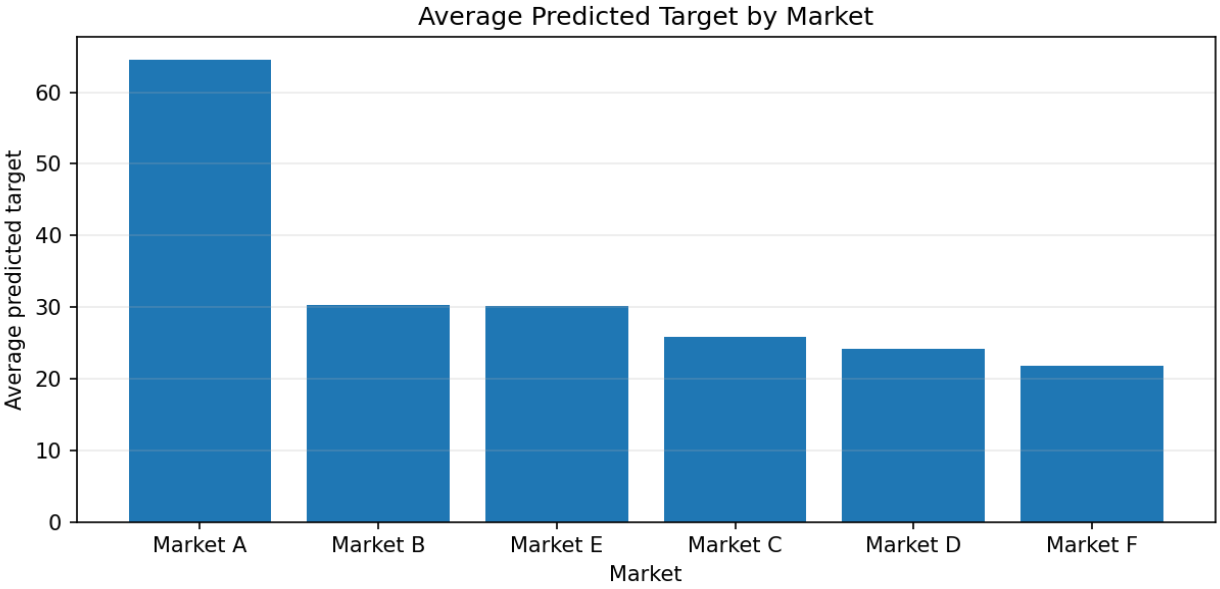

Figure 5. Average predicted target by market.

Key Conclusions

- The solution is production-ready for this competition and the submission format is valid.
- Performance degradation across newer folds suggests some distribution shift in summer months.
- The model captures central behavior well but may still under-react on extreme spikes.
- Best next improvement: add a second complementary model and build a weighted ensemble based on time-CV folds.

Output Files

- C:\Users\Kajetan\Desktop\Trading comp\complete_project_pipeline.py
- C:\Users\Kajetan\Desktop\Trading comp\submission_complete_hgb.csv
- C:\Users\Kajetan\Desktop\Trading comp\project_summary.pdf