Geometric Log-Space Ensembling with Orthogonal Neural Forecasters for Large-Scale Retail Revenue Prediction

Daniel Dragonevskiy
High School of Economics University, Moscow, Russia

Abstract:

Forecasting monthly store-level turnover across a network of tens of thousands of physical retail locations is a canon-ical Big-Data-in-Retail problem: the cross-sectional panel is wide and heterogeneous, the signal-to-noise ratio varies dramatically across store formats and geographic clusters, and the dominant variance components are multiplicative in nature – driven by recurring calendar effects, regional macro-economic cycles, and local competitive dynamics. Accurate prediction at this scale has direct operational consequences: a one-percent shift in forecasting error maps to material misallocations in procurement, replenishment logistics, and working-capital planning.

We present a full-stack forecasting pipeline developed for X5 Group, the largest grocery retailer in Eastern Eu-rope, operating more than 18,000 convenience and super-market points-of-sale. The task is to predict each store's rev-enue (RTO, retail turnover) for March 2025, given 25 months of daily sales history together with store metadata, a regional macro-economic panel, and a calendar of national and cultural events. Model performance is evaluated with the strictly multiplicative metric [illegible] = 100 (1 - MAPE/100)^2, which quadratically penalises proportional prediction error and is therefore minimised by forecasters that are geometrically unbiased – i.e., unbiased on the logarithmic scale – rather than arithmetically unbiased.

The central methodological contribution is a rigorous ex-ploitation of this geometric structure. We derive that, for log-normally distributed targets, the geometric (log-space) ensemble is the MAPE-optimal aggregator, and confirm this analytically and empirically. The proposed pipeline combines three deliberately heterogeneous base learners: (i) a Light-GBM Tweedie model that captures macro-level and regional salary effects via a ratio reparameterisation of the target; (ii) a CatBoost regressor trained on log-RTO with MAE loss and March-aligned calendar features, including a binary in-dicator for the Saturday placement of International Women's Day; and (iii) an N-BEATSx neural forecaster trained on GPU, whose residuals are nearly orthogonal to the tabular models (Pearson correlation 0.046), providing a genuinely independent correction signal. The three constituents are fused via a weighted geometric blend in log space, subject to a global mass-preservation constraint that stabilises the pre-dicted revenue total and eliminates a systematic scale-drift artefact.

The pipeline advances the competition metric from 90.28 (naive calendar baseline) through 90.68 (best single model) to 90.76 (final three-way ensemble), placing within 0.06 of the top-1 solution. We additionally document an extensive abla-tion study covering negative results – linear blends, multi-plicative scale shifts, semantic re-weighting, and the Chronos zero-shot neural model (residual correlation 0.995, no ensem-ble gain) – and discuss why log-space averaging is the natural geometry of MAPE in retail revenue forecasting.

Keywords:
Retail forecasting, MAPE, gradient boosting, N-BEATSx, en-sembling, geometric mean, log-normal targets, mass preser-vation.