M5 Forecasting - Accuracy
Estimate The Unit Sales of Walmart Retail Goods

Peiyuan Liu[1], *Yunqi Li*[2]
[1]*Tsinghua University* [2]*Tsinghua University*

Paper, code,and data are available at:
https://github.com/Hank0626/M5_THU

Peering In

In the challenge, you are predicting item sales at stores invarious locations for two 28-day time periods.

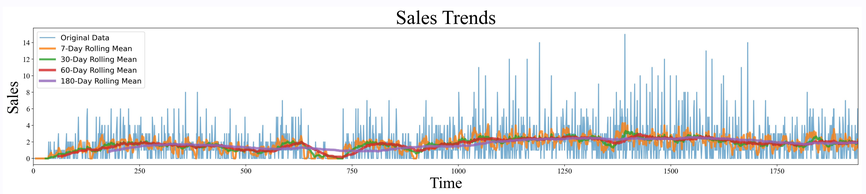
Our contributions：

- We decompose the time series of product sales into long-term and short-term variations based on periodicity. The long-term changes are predicted using Transformer models, while the short-term fluctuations are forecasted using CNNs.
- We construct a decision tree for prediction, utilizing various attributes of the product, including price, calendar, lag, and roll.
- We combined the predictions from deep learning with the results of decision trees using a weighted average approach, ultimately achieving a score of 0.517 on the private test set.

Exploratory Data Analysis

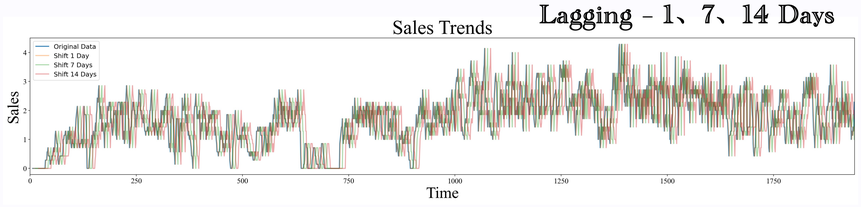
- We visualize the variation in product prices over time, and conduct visual analyses for rolling averages over 7, 30, 60, and 180 days. Rolling operations allow for a smoother representation of the trend in price changes.

Rolling - 7、30、60、180 Days

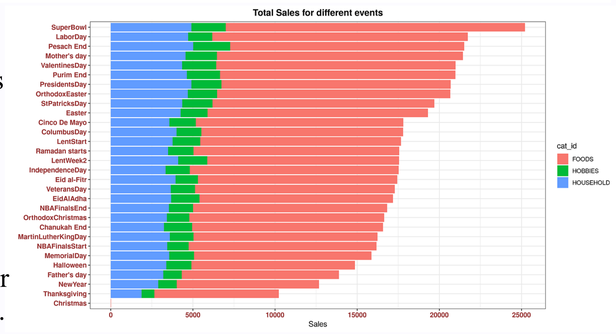


- We also visualize the effects of lagging the prices by 1, 7, and 14 days. Lagging operations aid in analyzing lag effects in time-series data, offering a deeper understanding of price dynamics.

Lagging - 1、7、14 Days

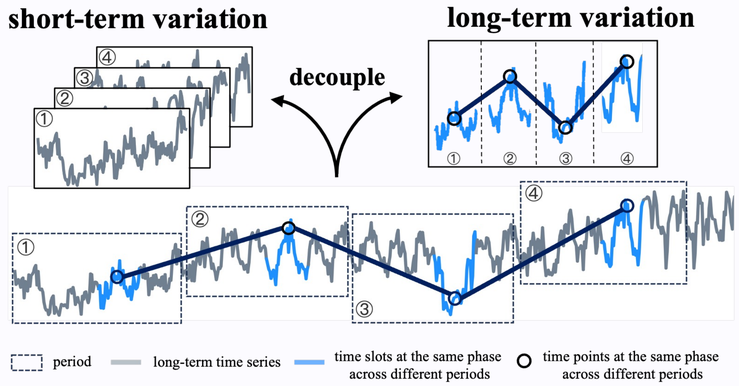


- For the sales volume of products during special festivals, we conduct relevant statistics and incorporated them as feature inputs for subsequent models.

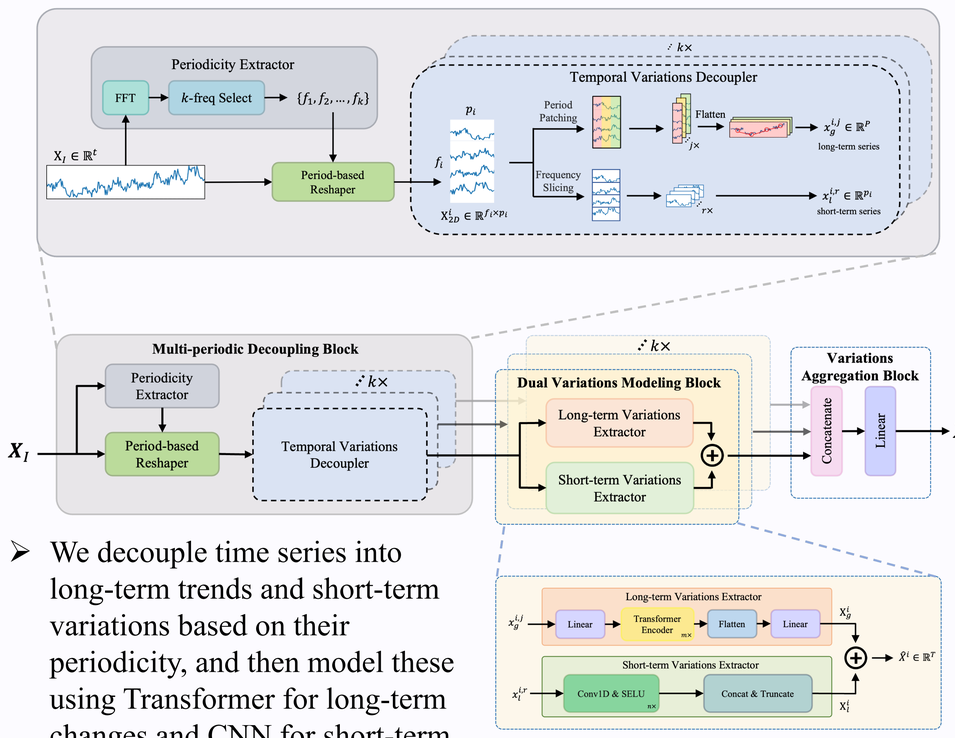


Framework – Deep Learning

- Decoupling time series into long-term and short-term variations.



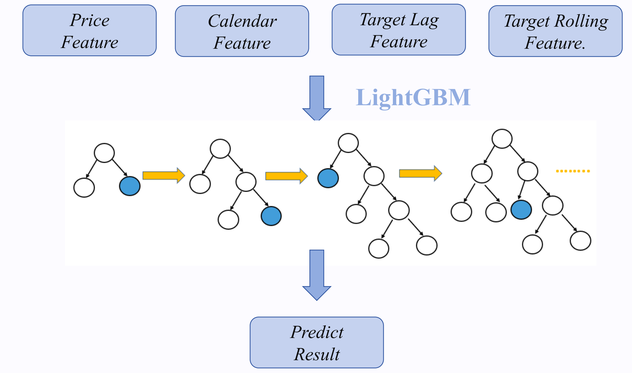
- We utilize the initial 1,942 days of data for each product as the training and testing set, employing the historical 336-day period to forecast the subsequent 28 days.



- We decouple time series into long-term trends and short-term variations based on their periodicity, and then model these using Transformer for long-term changes and CNN for short-term changes.

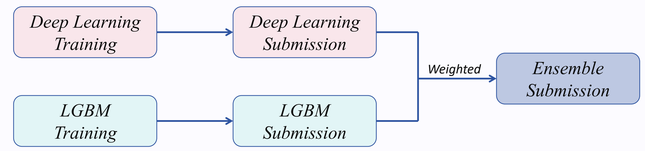
Framework – LightGBM

- We utilize four features — price, calendar, target lag, and target roll — as inputs to construct decision trees for prediction using LightGBM as the model.



Mix of Expert (MoE)

- Final Submission is the ensemble of deep learning and LGBM



Predict Results

Method	Private	Time
Deep Learning	**0.79156**	**～1h**
LGBM	**0.55078**	**～2h**
Ensemble	**0.51747**	**～3h**
Kaggle 1^{st} solution	**0.52043**	**～12h**

- Our approach achieved a score of 0.517 on the private test set, marking a 6% improvement over the first-place solution on Kaggle's leaderboard.
- Our approach, by integrating deep learning with the results of LGBM, effectively reduced the training time. The entire process of training and testing was completed in approximately 3 hours.