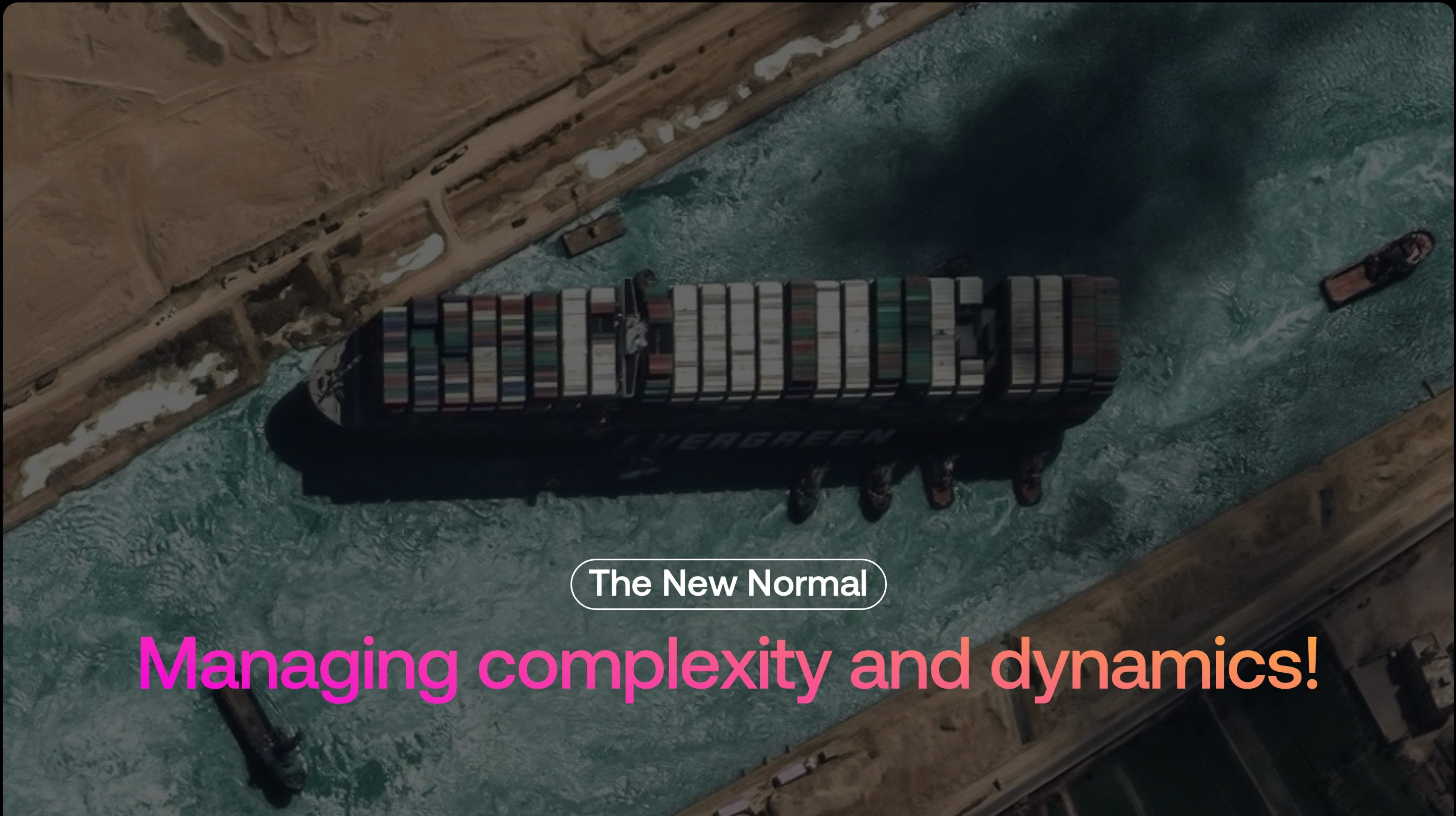
Boosting Time Series Accuracy- The power of Ensemble Methods



The New Normal

Managing complexity and dynamics!

Robert Haase
Lead AI Scientist
paretos

Forecast Data Science Team - Paretos

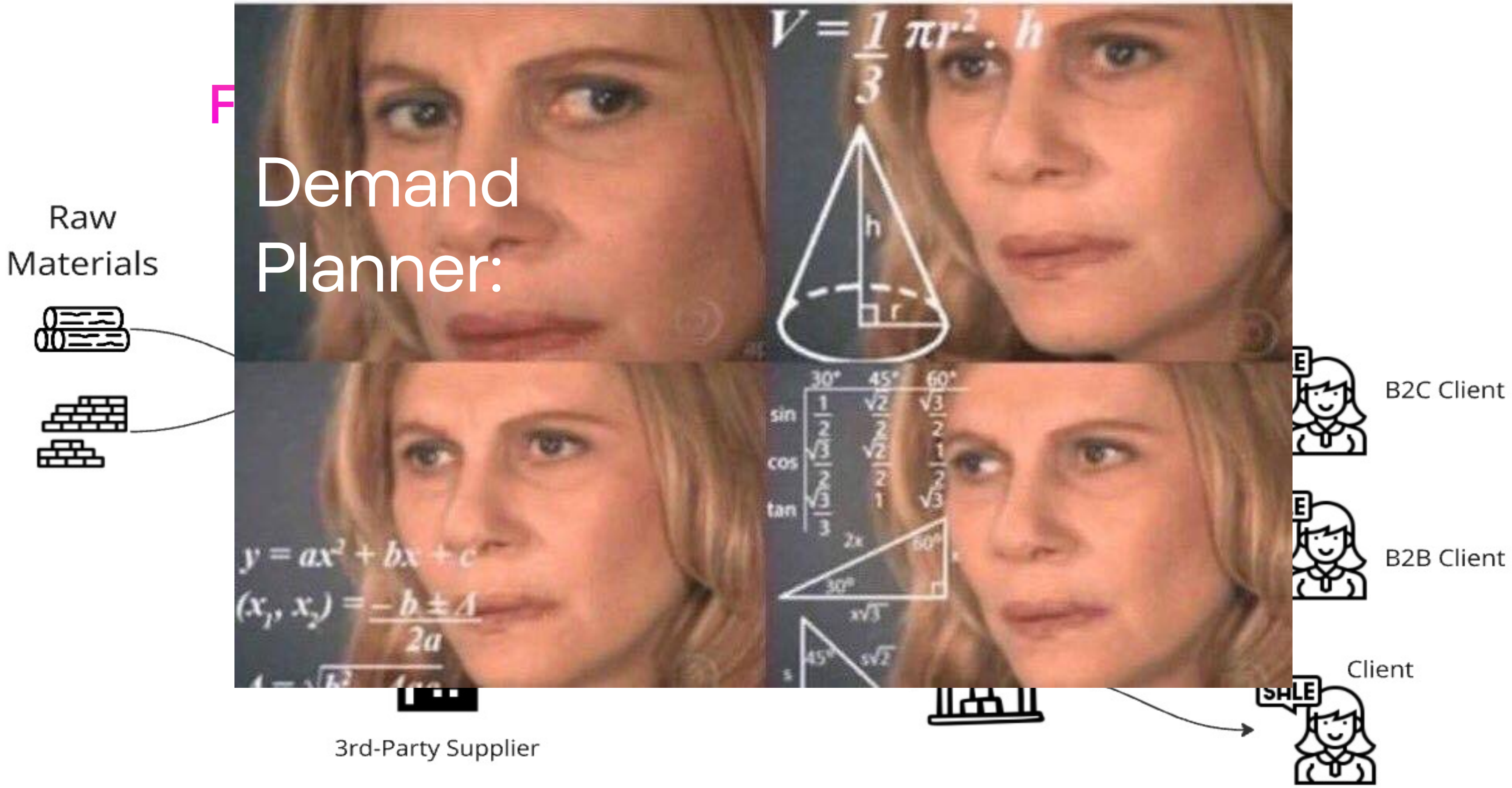
Alexander Meier
Senior Data Scientist

Fabian Bergermann
Senior Data Scientist

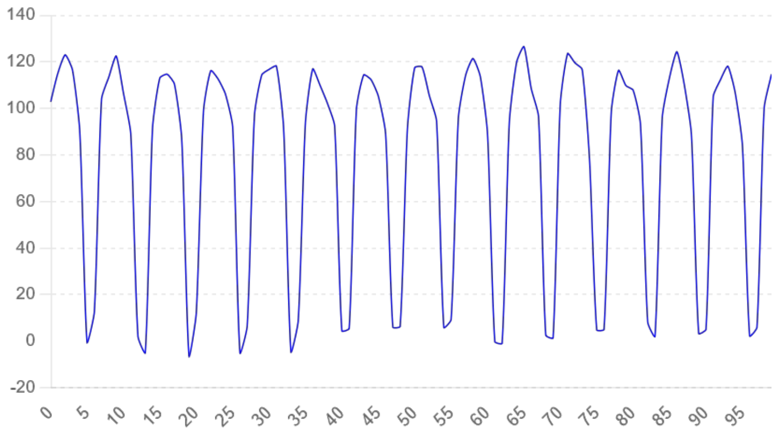
Pedro Harloch
ML Engineer

Daria Mokrytska
Data Scientist

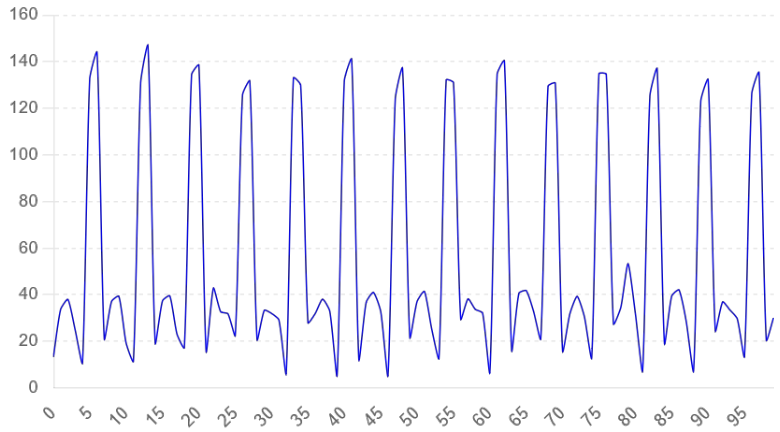
Demand Forecast of a manufacturing plant



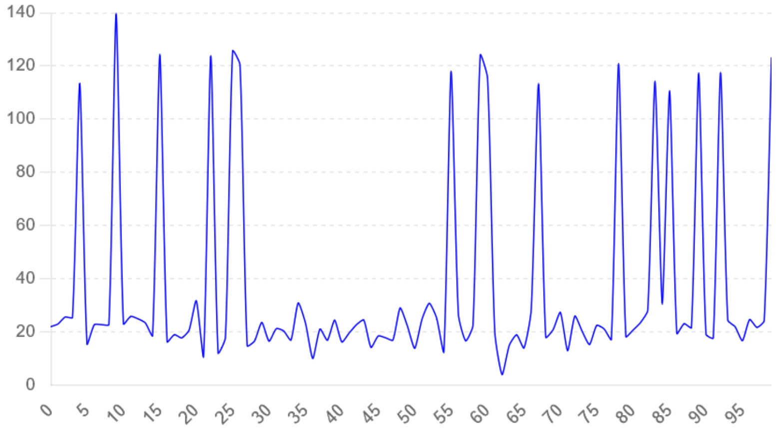
Different Demand Patterns in Manufacturing



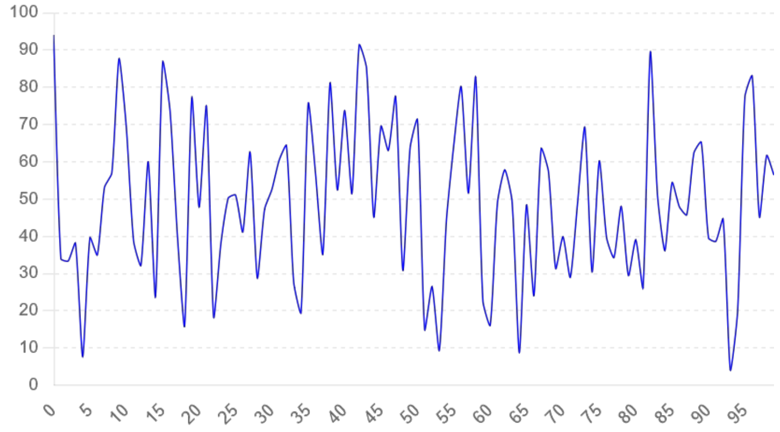
Weekly Pattern



Different weekly pattern



Intermittent Data (B2B Sales)



High noise, low signal

Demand Drivers in Retail

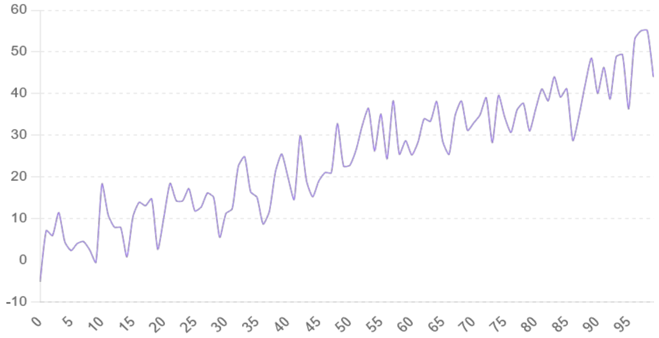
Never out of stock products

New product launches

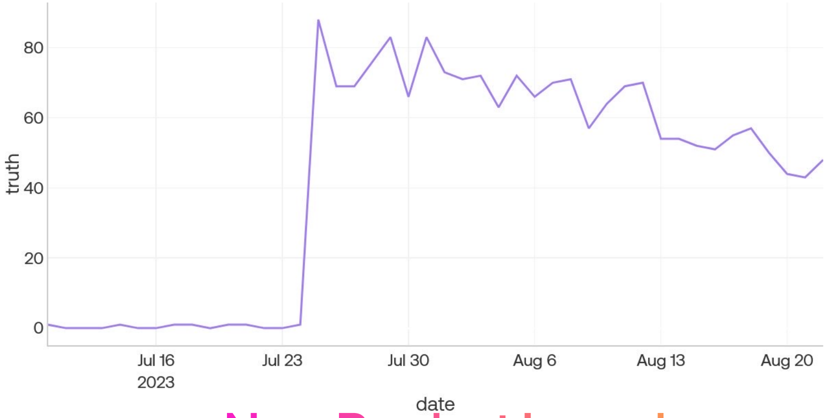
promotions/ special days

Unpopular items

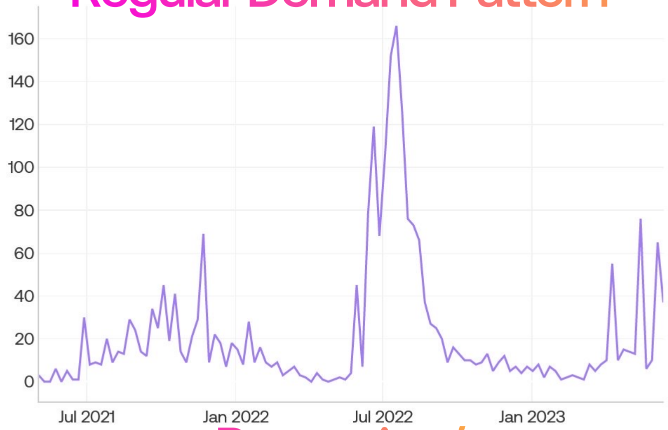
Different Demand Patterns: Retail Example



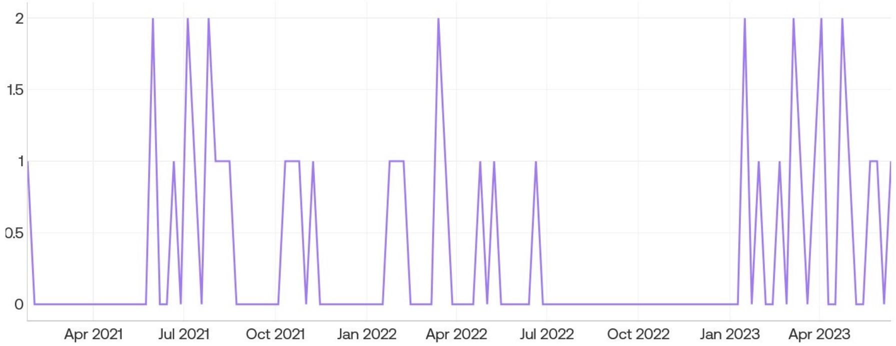
Regular Demand Pattern



New Product Launches

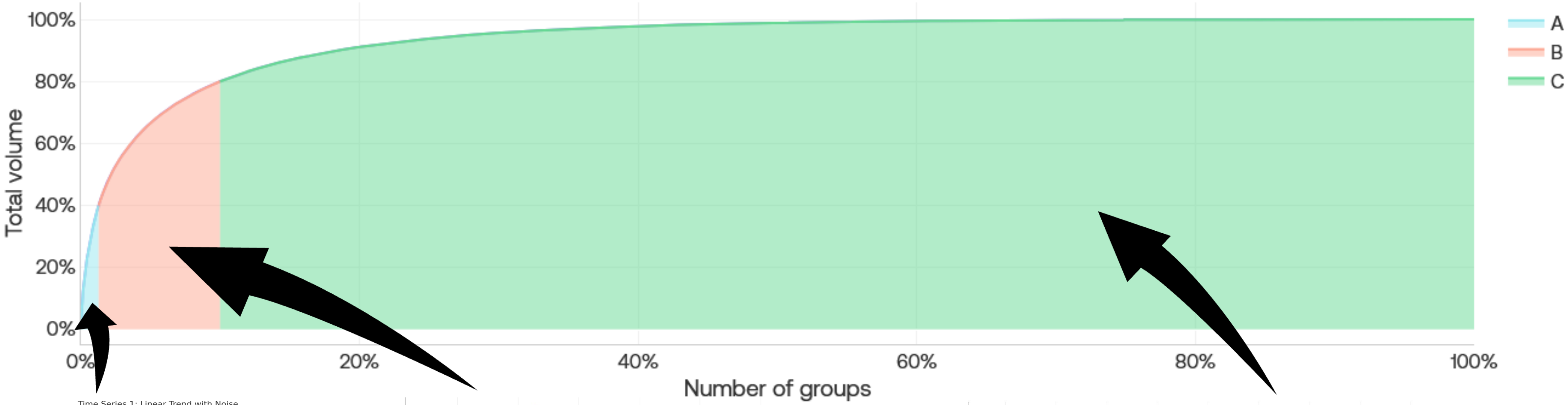


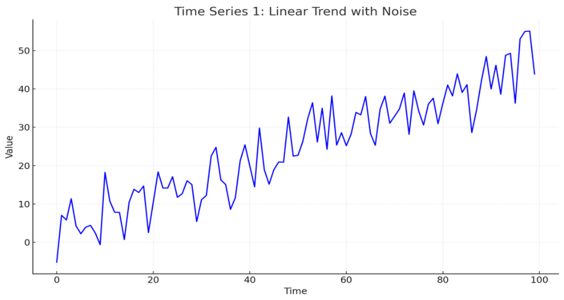
Promotion/
Special Discount Days



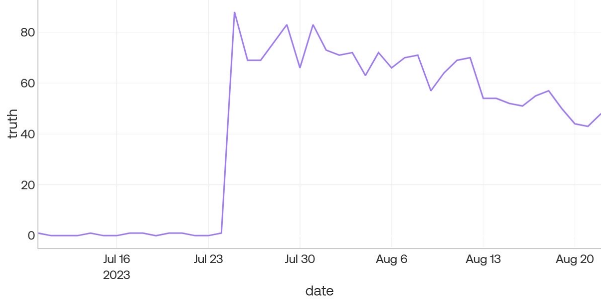
Intermittent Time Series

Data Distribution Retail Example

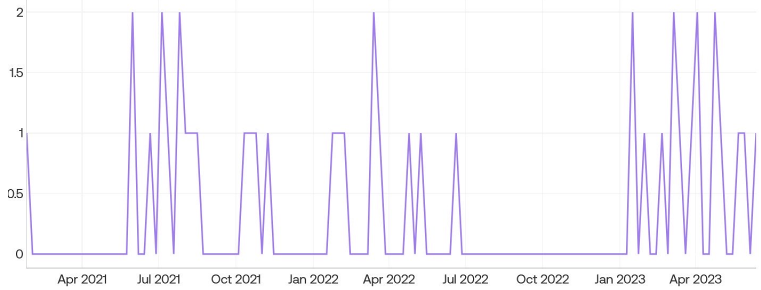




Regular Demand Pattern

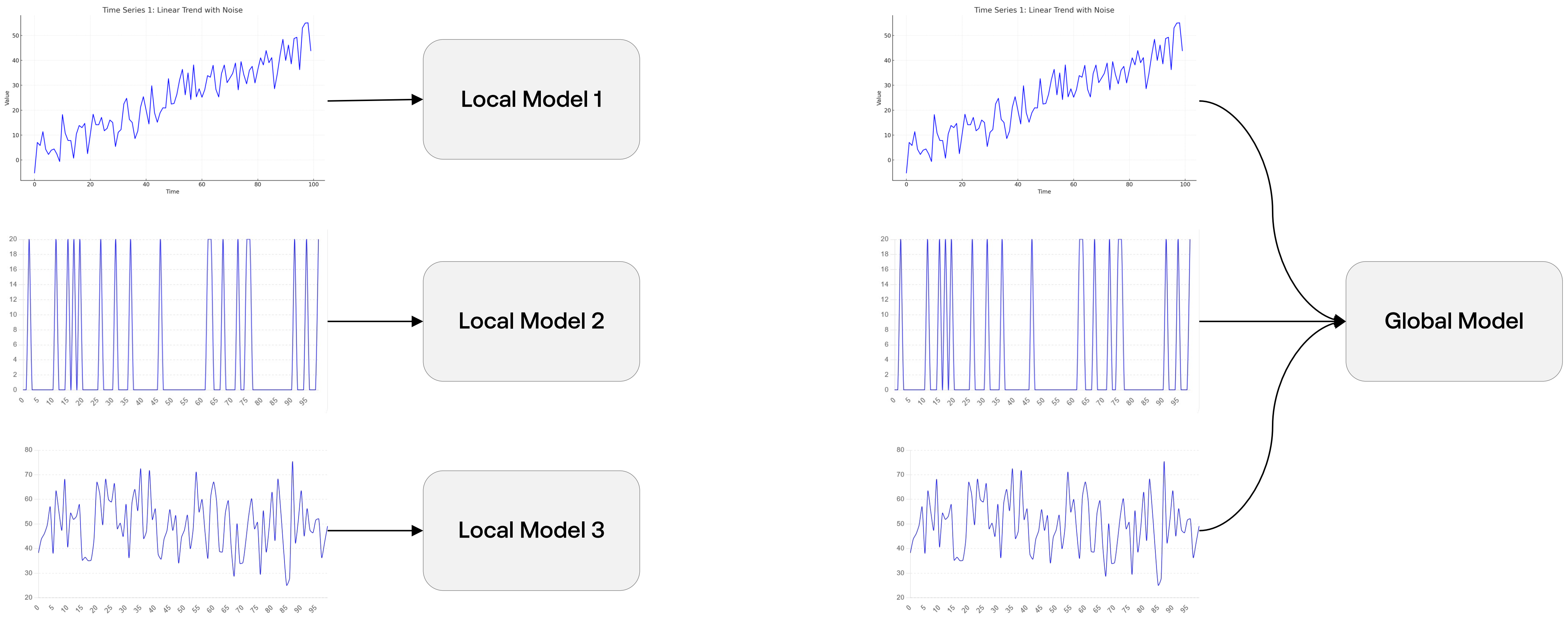


New Product Launches



Intermittent Time Series

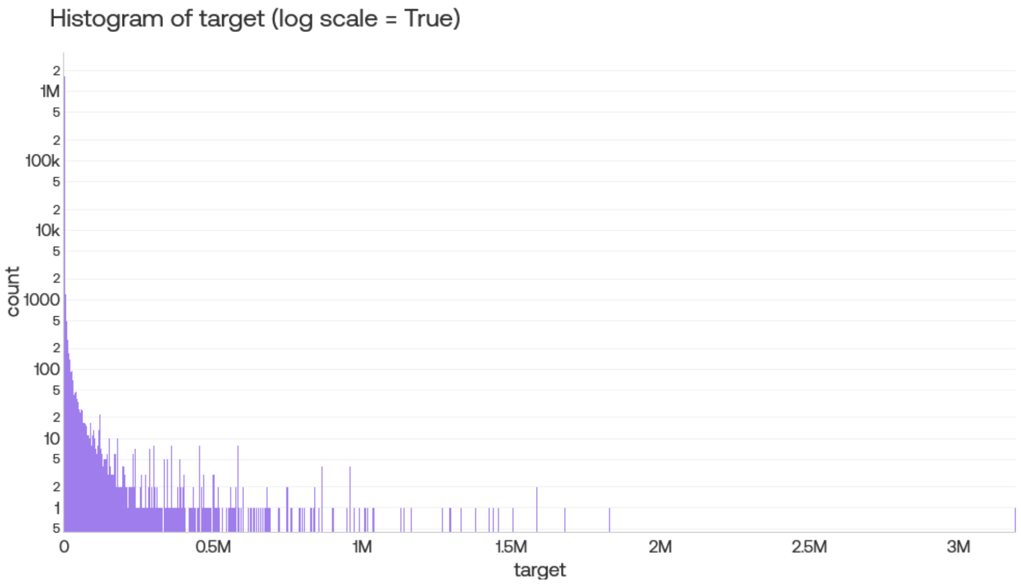
Local vs Global Models

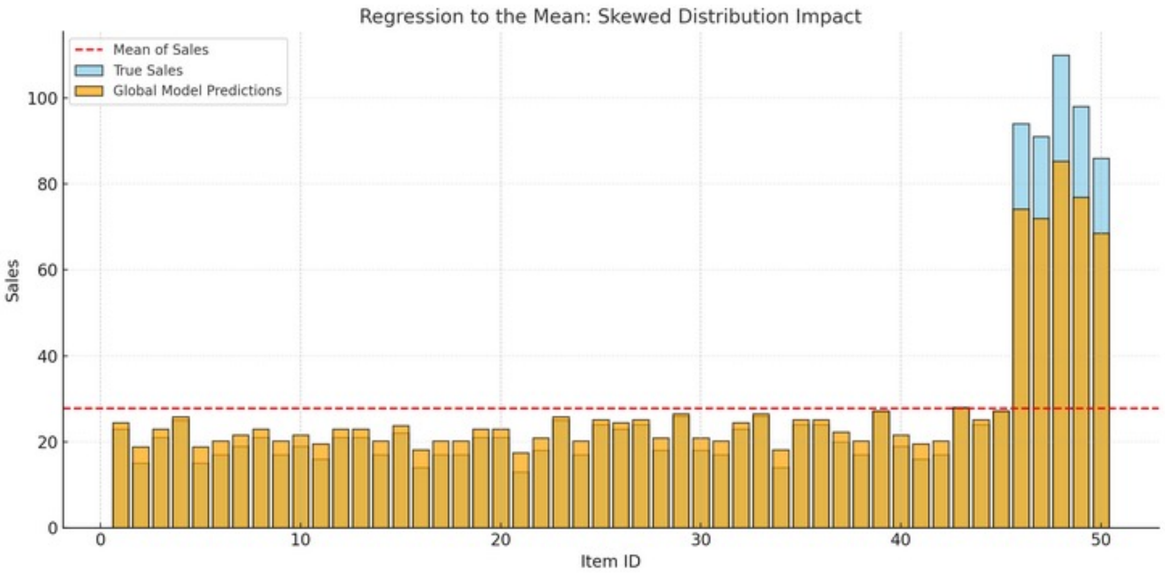


Typically statistical models

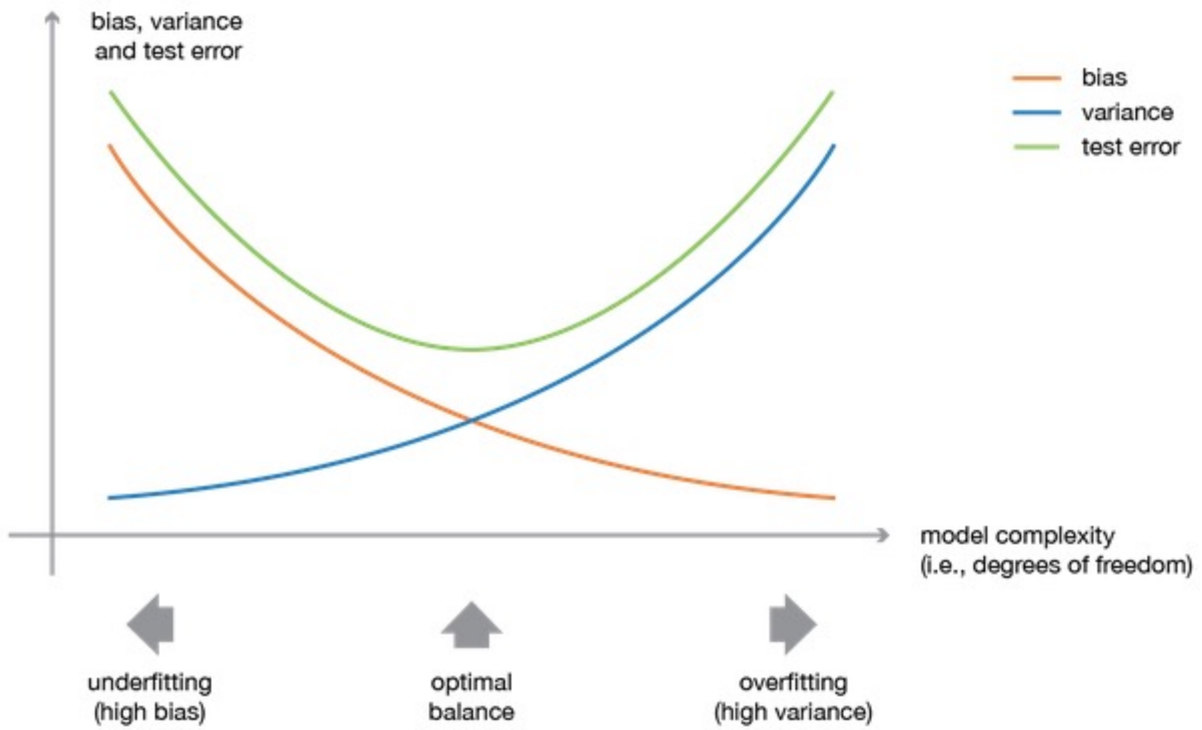
ML models, deep learning models

“Regression to the Mean” for Global Models





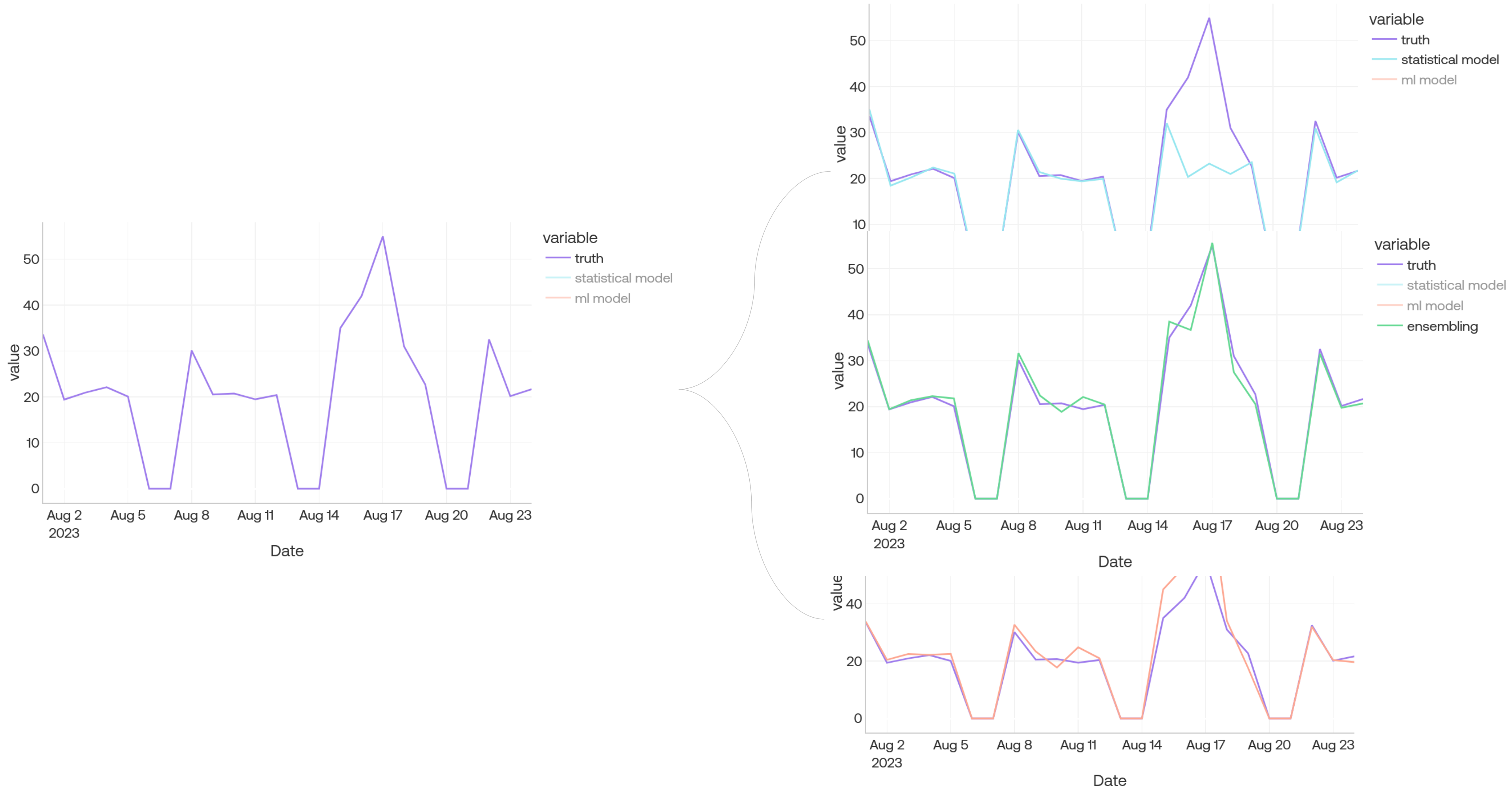
Weak Learners “Bias-Variance-Tradeoff”



https://duchesnay.github.io/pystatsml/machine_learning/ensemble_learning.html

Combining individual approaches to find a robust Bias-Variance trade-off and improve predictive performance

Notion of Ensembling

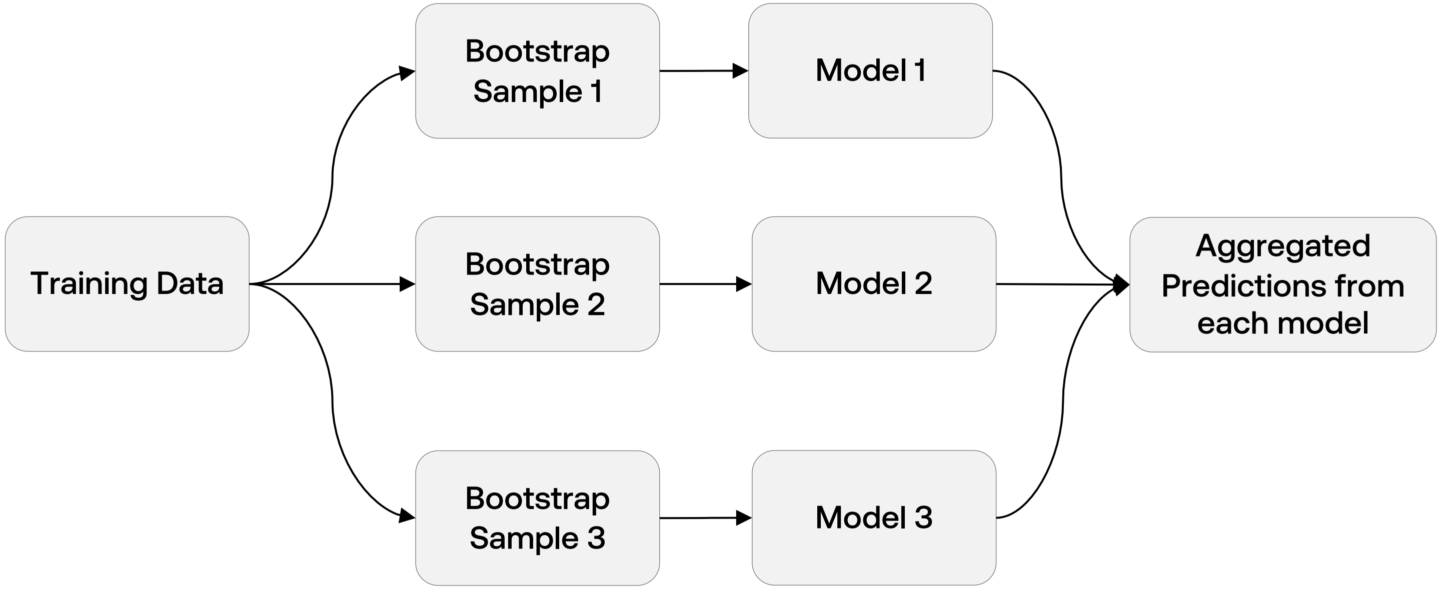


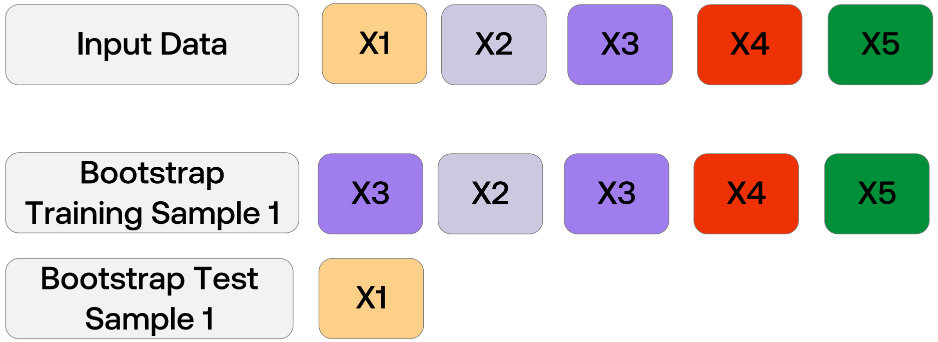
Statistical Model

Ensembled Model

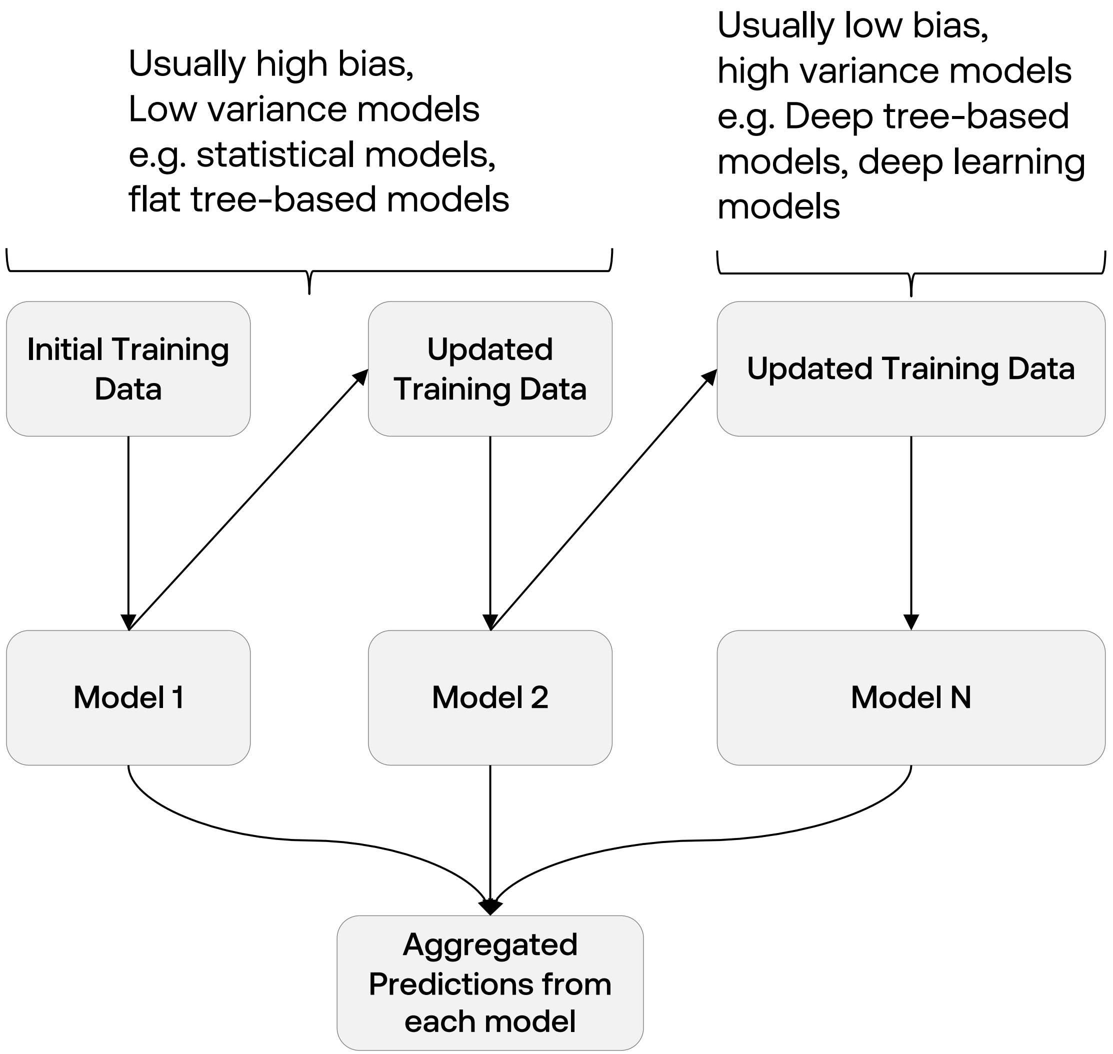
ML Model

Bagging Ensembling

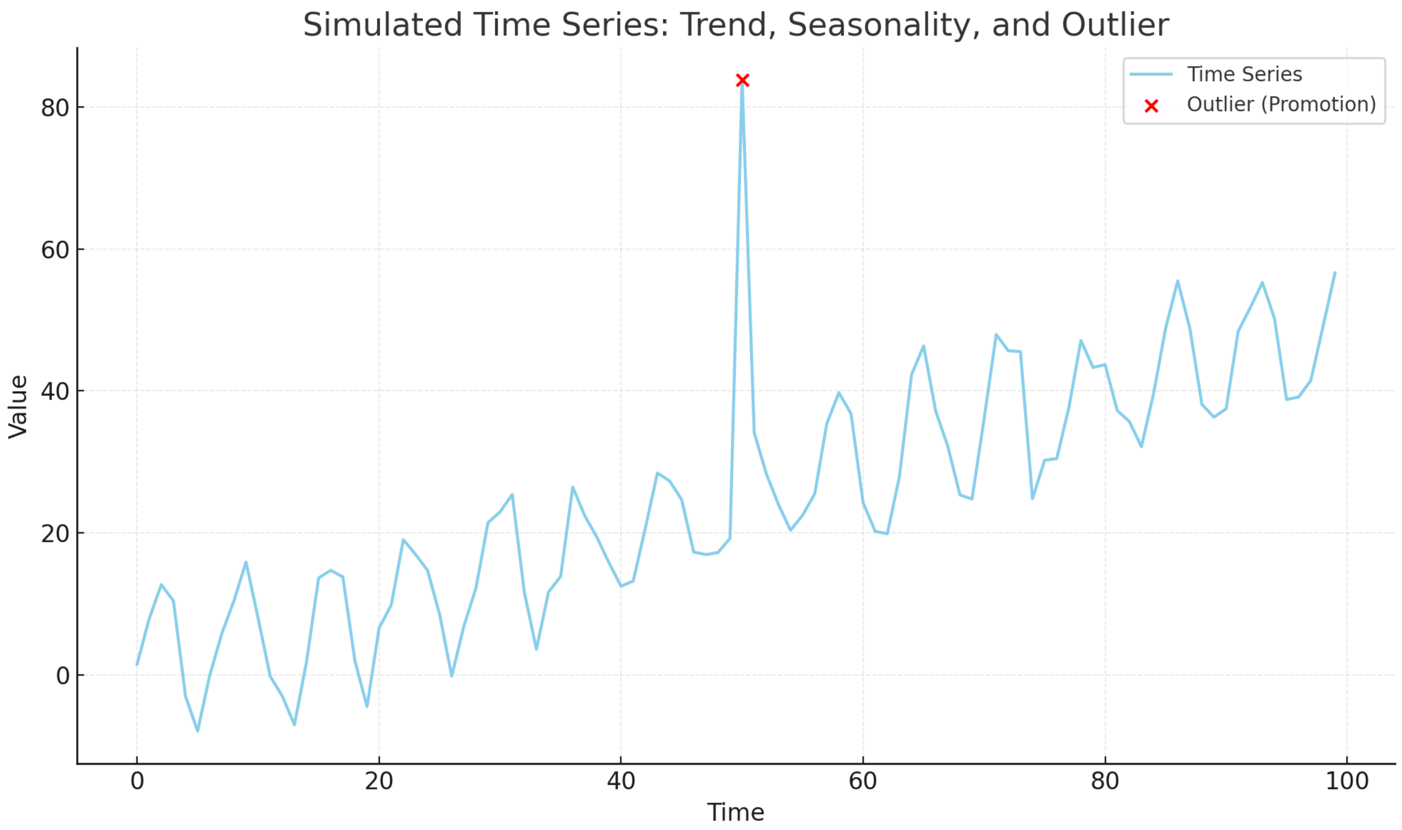




Boosting Ensembling

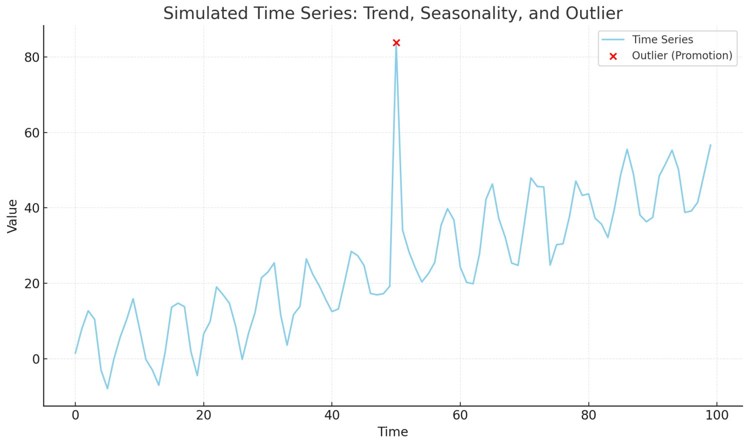


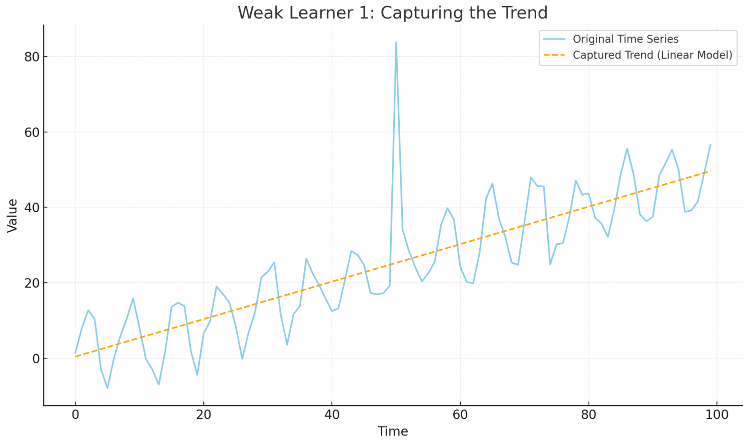
Boosting Ensembling in Time Series Forecasting



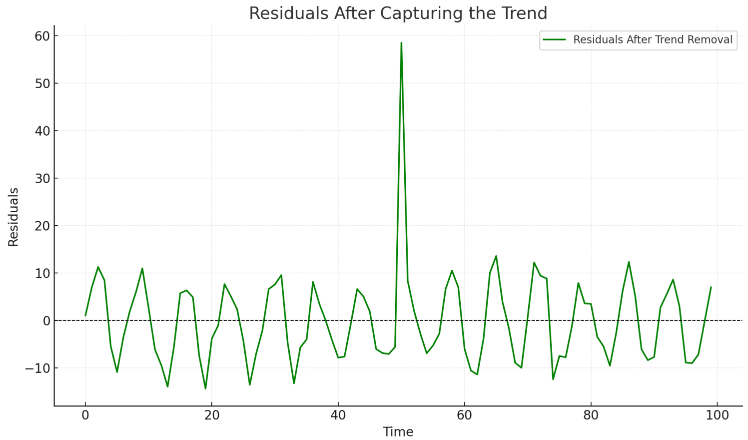
Boosting Ensembling in Time Series Forecasting

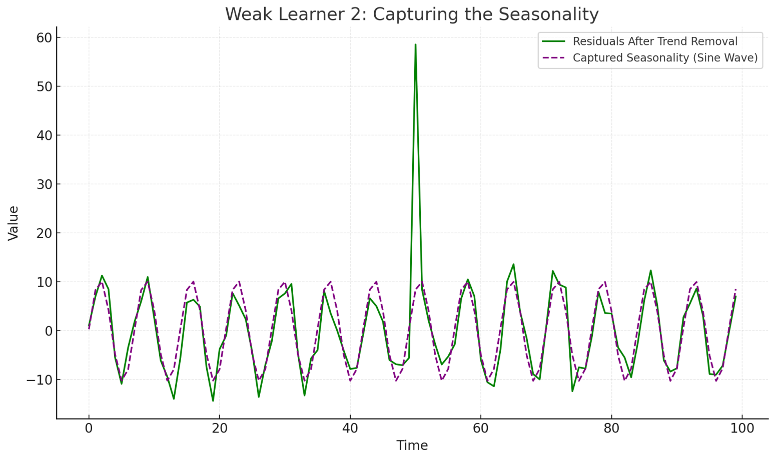
Step 1: Capturing the trend





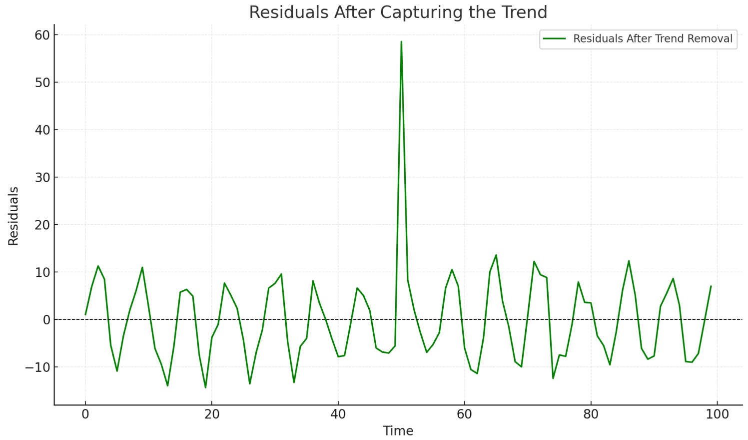
Step 2: Capturing the pattern

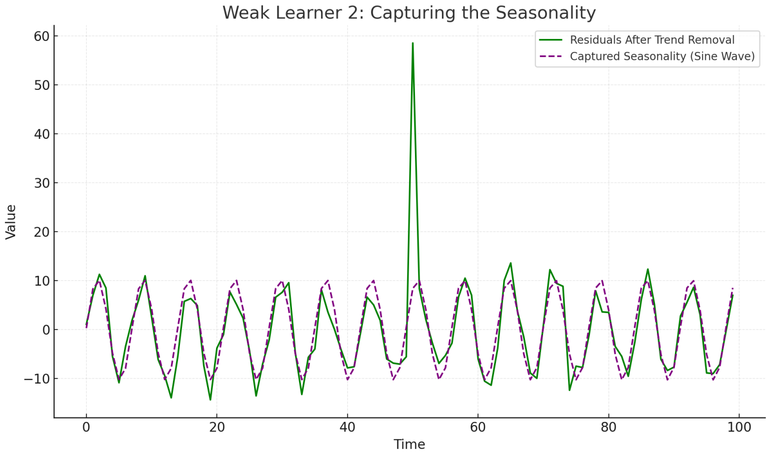




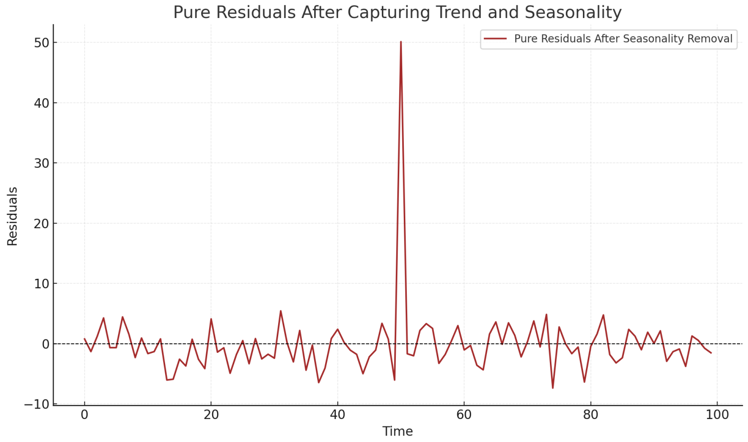
Boosting Ensembling in Time Series Forecasting

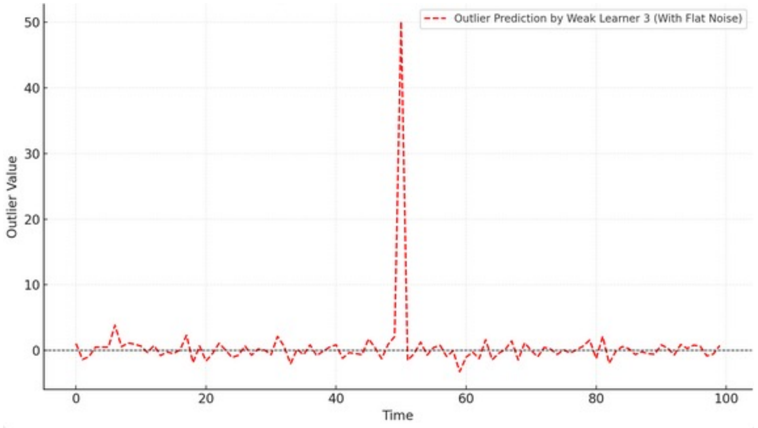
Step 2: Capturing the pattern



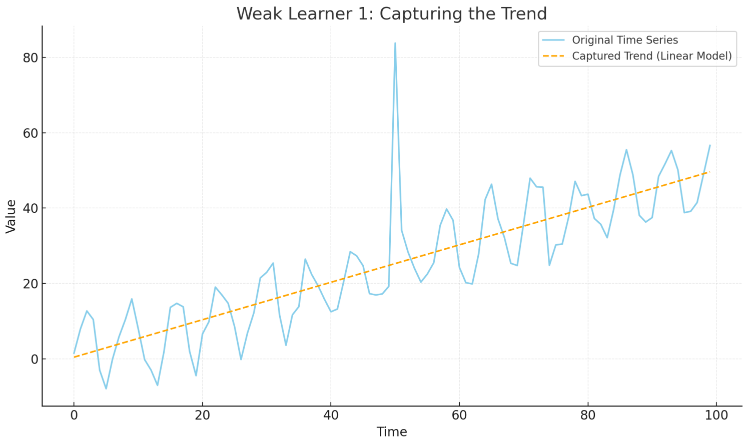


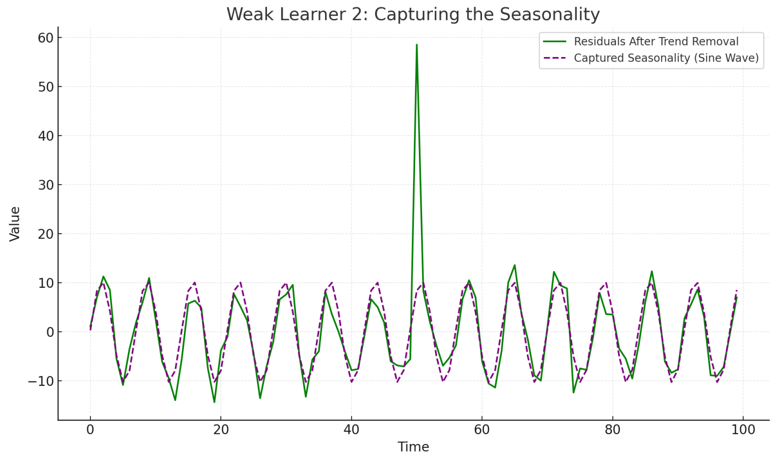
Step 3: Capturing the outliers

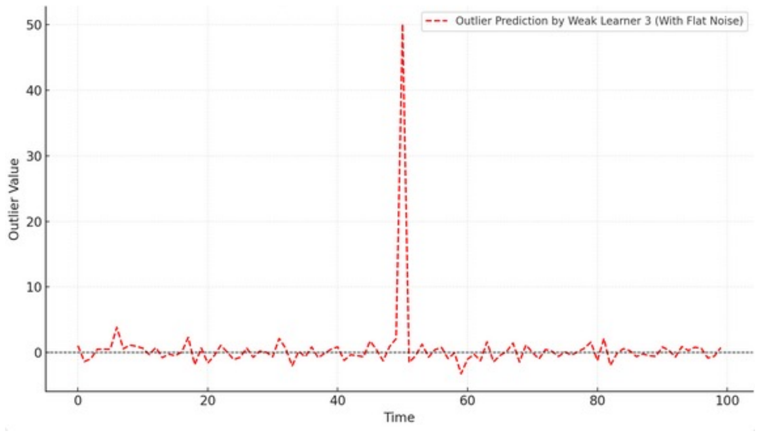


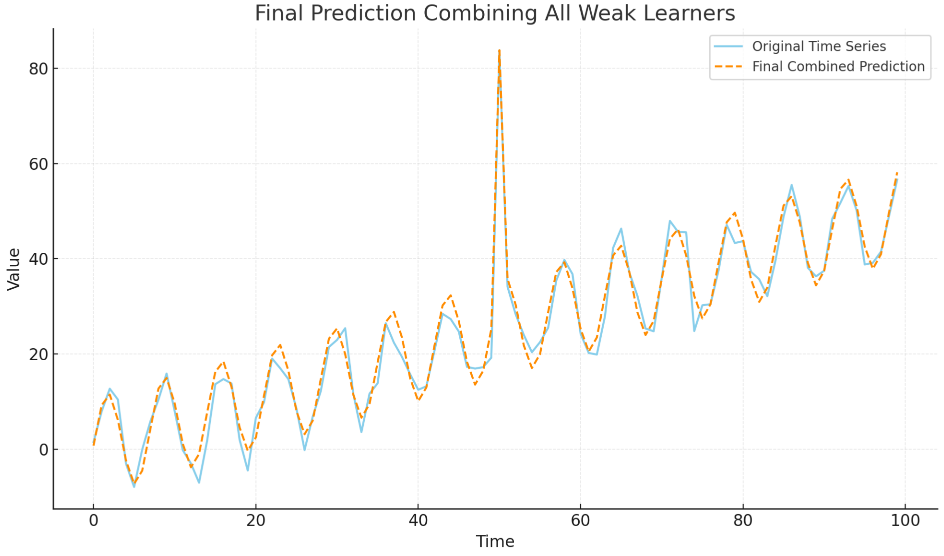


Boosting Ensembling in Time Series Forecasting

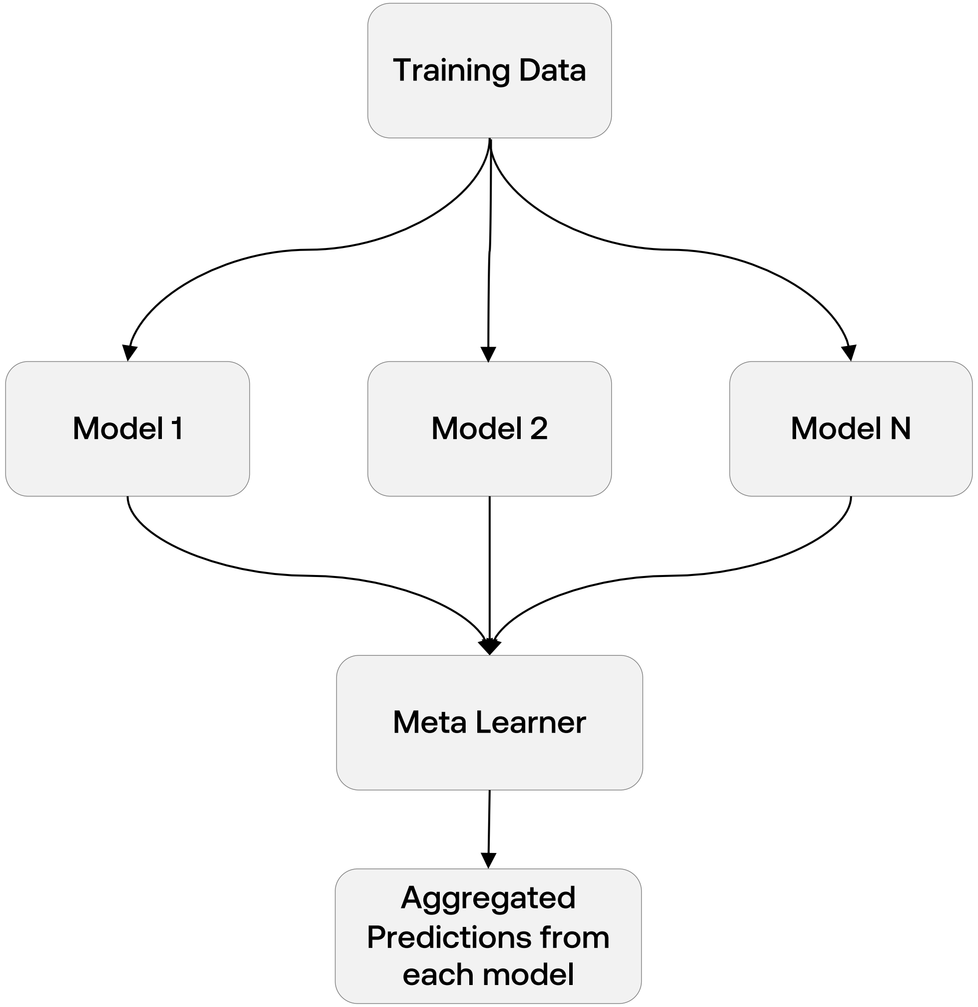




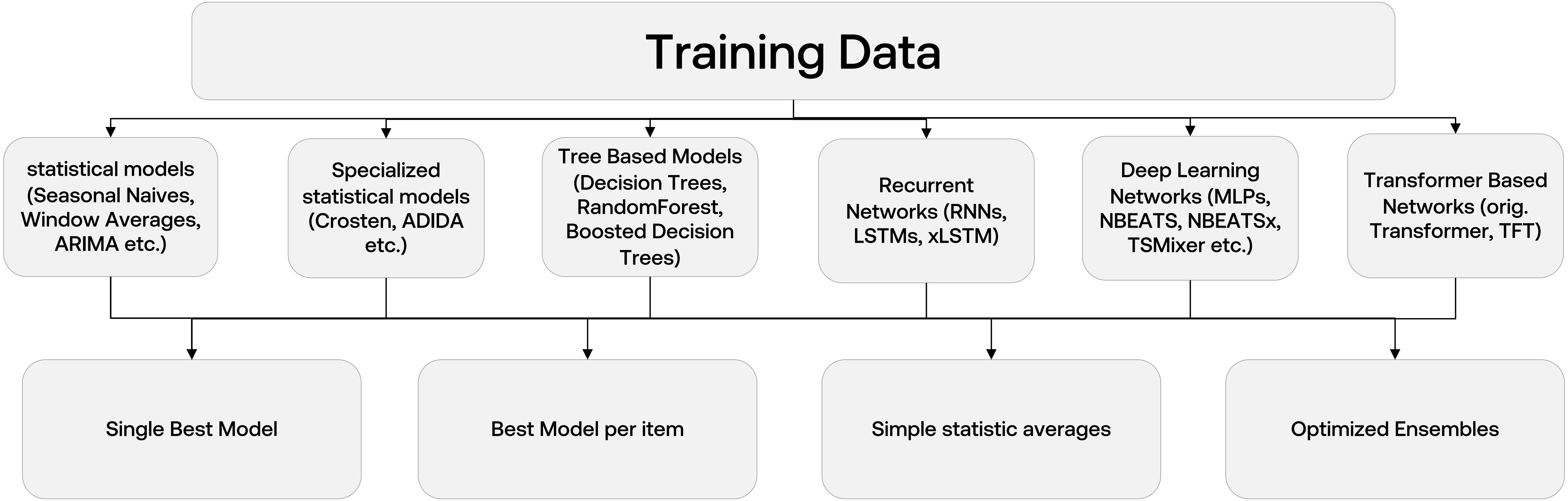




Stacking Ensembling



Power comes through flexibility in Stacking Ensembling Approach

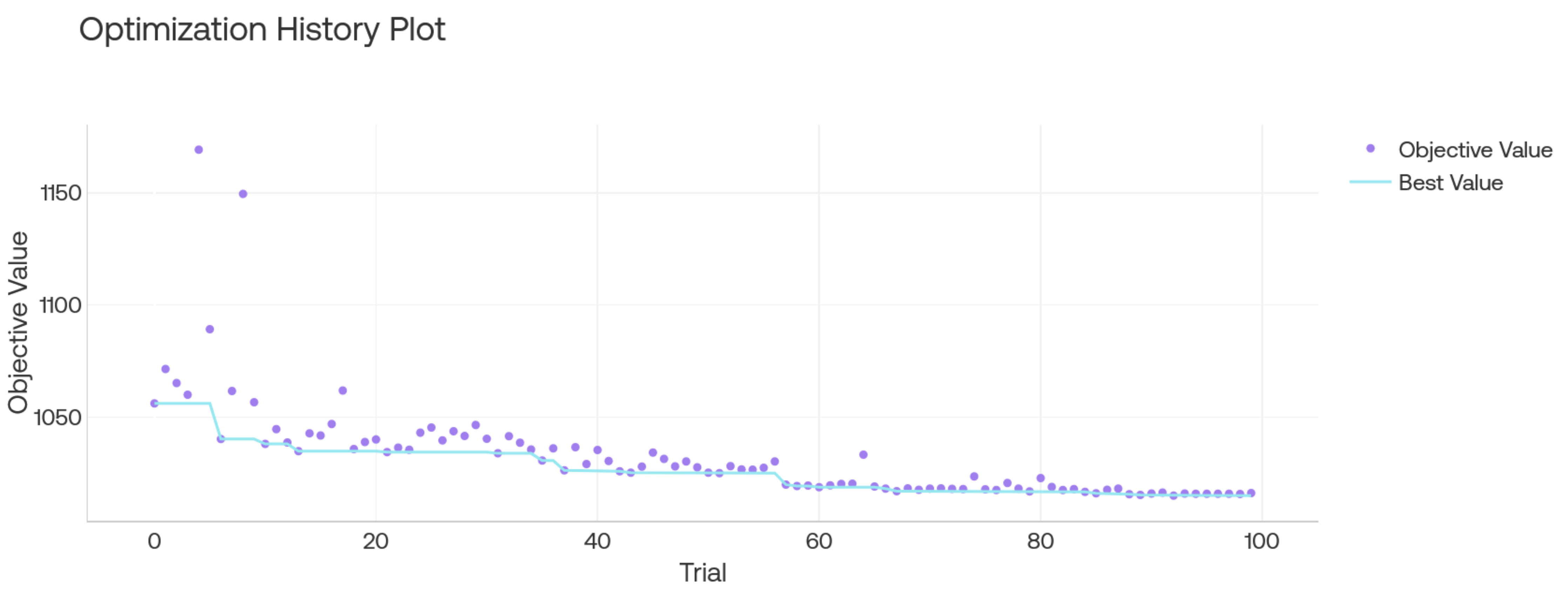


No “classical” Meta Learner Model due to higher complexity in handling additional ML Model

Optimized Stacking Ensembling

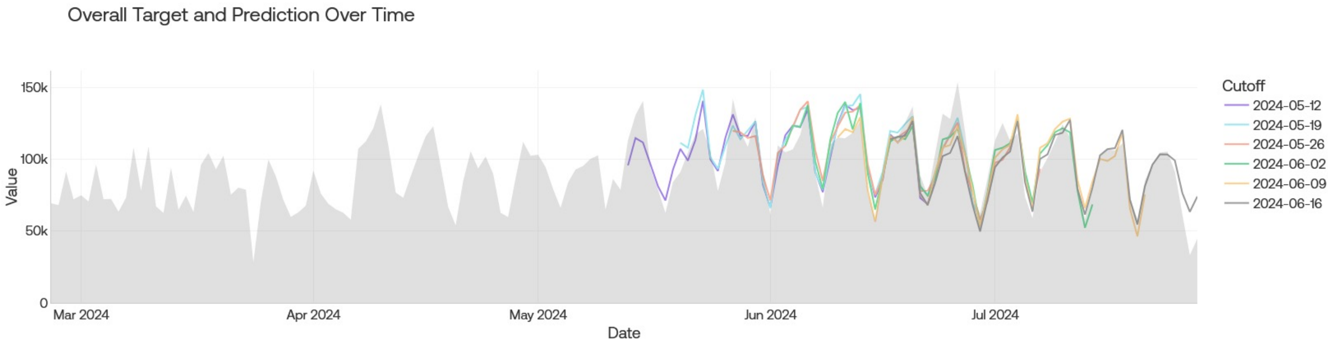
model	weight
auto_arima_12	0.013828020840354142
auto_arima_12_wout_exog	0.3353154367612936
AutoETS	0.021110262674598177
HistoricAverage	0.043168518344653635
Naive	0.017570250879700867
seasonal_smoothing_12	0.041571323576030474
SeasonalNaive	0.033941562483361574
WindowAverage	0.15306872721268203
XGBRegressor	0.3404258972273256

Optimized Stacking Ensembling



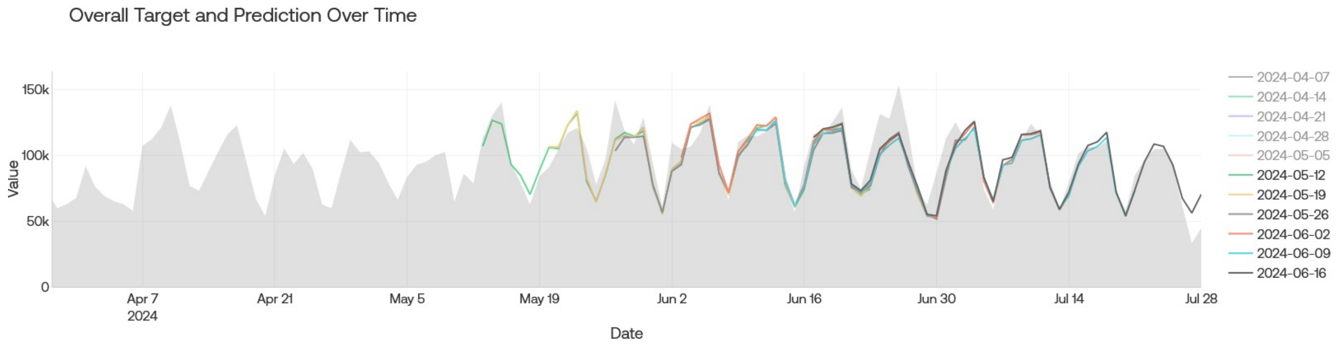
Improved Accuracy

single weak learner (XGBoost)



best_bias	-12.14028793905214
best_mae	437.95906201775983
best_mape	1.1269096668069234
best_marre	38.04149514952431
best_model	XGB
best_rmse	568.381550016413
best_wape	0.2217976684049948

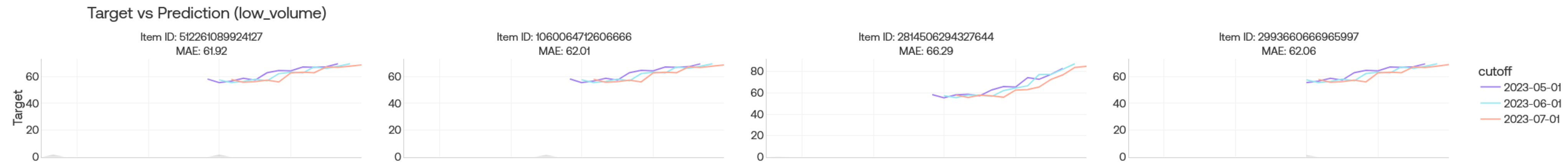
Optimized stacking ensemble (statistical models, ML models)



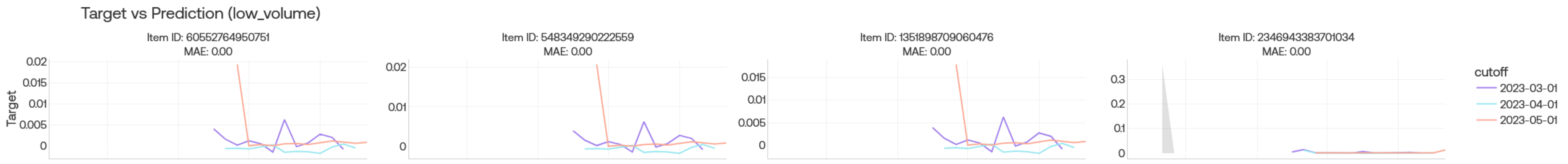
best_mae	360.02462166080727
best_mape	0.4099874387792668
best_marre	15.530316471656414
best_model	Ensemble[XGB_xlarge_5,XGB_xl...
best_rmse	478.80815704347435
best_single_model	XGB_xlarge_4
best_wape	0.2029841262541356

Reduction of “Regression to the mean” phenomena for low volume groups

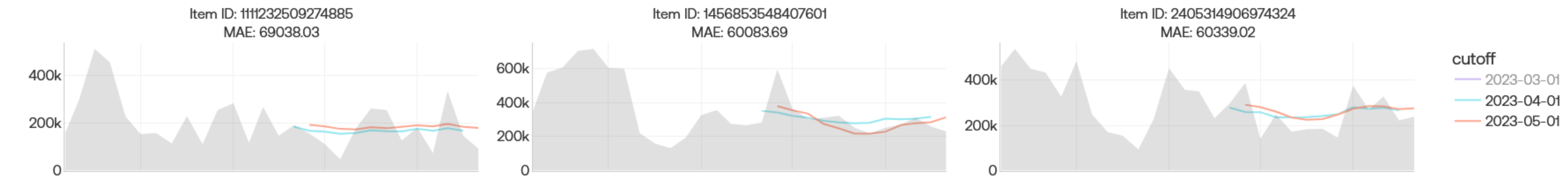
single weak learner



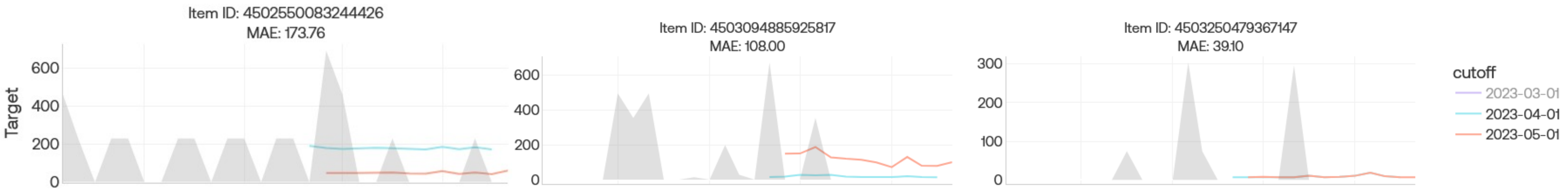
Optimized stacking ensemble



Handling heterogenous target patterns



model	weight
auto_arima_12	0.01718390315...
auto_arima_12_wout_exog	0.00257300159...
AutoETS	0.05815134832...
AutoNHITS	0.71862674816...
HistoricAverage	0.01018542298...
Naive	0.05865703863...
seasonal_smoothing_12	0.04158351043...
SeasonalNaive	0.02228818959...
WindowAverage	0.07075083711...



model	weight
auto_arima_12	0.00426378414...
auto_arima_12_wout_exog	0.59569270056...
AutoETS	0.06083338539...
HistoricAverage	0.13196745308...
Naive	0.09684248053...
seasonal_smoothing_12	0.05876083456...
SeasonalNaive	0.05163936171...

No more bad decisions

Thanks for listening!

Robert Haase, Lead AI Scientist, paretos