Presentation for quantitative wealth and investment management $\mathbb{QWIM}$

The Authors

June 2021

Over recent decades, machine learning $\mathbb{ML}$ algorithms have achieved remarkable success in various areas

Given a large representative dataset, ML algorithms can learn to identify complex non-linear patterns and explore unstructured relationships without hypothesizing them in advance.

Thus, $\mathbb{ML}$ algorithms are not limited by assumptions or pre-defined data generating processes, which allows the data to speak for itself.

Literature Review

Some references: Baltas and Scherer (2019), Li et al. (2020), and Rasekhschaffe and Jones (2019)

More references

Metrics for assessing performance

Metric: symmetric mean absolute percentage error sMAPE

$$sMAPE \triangleq \frac{2}{h} \sum_{t=n+1}^{n+h} \frac{\left| Y_t - \widehat{Y_t} \right|}{|Y_t| + \left| \widehat{Y_t} \right|} * 100\,()$$

where Y_t is the value of the time series at point t, $\widehat{Y_t}$ the estimated forecast, h the forecasting horizon, n the number of the data points available in-sample, and m the time interval between successive observations considered for each data frequency, i.e., 12 for monthly, four for quarterly, 24 for hourly and one for yearly, weekly and daily data.

Metric: mean absolute scaled error M

$$sMAPE \triangleq \frac{1}{h} \frac{\sum_{t=n+1}^{n+h} \left| Y_t - \widehat{Y_t} \right|}{\frac{1}{n-m} \sum_{t=m+1}^{n} |Y_t - Y_{t-m}|}$$

where Y_t is the value of the time series at point t, $\widehat{Y_t}$ the estimated forecast, h the forecasting horizon, n the number of the data points available in-sample, and m the time interval between successive observations considered for each data frequency, i.e., 12 for monthly, four for quarterly, 24 for hourly and one for yearly, weekly and daily data.

Results

Comparison was performed against 4 models from M4 competition, each best in their respective model class:

- Best pure ML: best entry among the 6 pure ML models.
- Best statistical: best pure statistical model
- Best ML/TS combination: second best entry
- DL/TS hybrid: winner of M4 competition

The results are summarized in following tables (lower values are better).

In parantheses we show the number of datasets of that type.

Table 1: Performance on M4 test set: sMAPE (top) and OWA and M4 Rank (botom)

	Yearly (23K)	Quaterly (24K)	Monthly (48K)	Others (5K)	Average (100K)
Best pure ML	14.397	11.031	13.973	4.566	12.894
Best statistical	13.366	10.155	13.002	4.682	11.986
Best ML/TS combination	13.528	9.733	12.639	4.118	11.720
DL/TS hybrid, M4 winner	13.176	9.679	12.126	4.014	11.374
NBEATS-G	12.855	9.378	12.130	3.979	11.229
NBEATS-I	12.823	9.418	12.048	4.199	11.203
NBEATS-I+G	12.812	9.372	12.064	4.063	11.190

	Yearly	Quaterly	Monthly	Others	Average	Rank
Best pure ML	0.859	0.939	0.941	0.991	0.915	23
Best statistical	0.788	0.898	0.905	0.989	0.861	8
Best ML/TS combination	0.799	0.847	0.858	0.914	0.838	2
DL/TS hybrid, M4 winner	0.778	0.847	0.836	0.920	0.821	1
NBEATS-G	0.755	0.814	0.823	0.876	0.799	
NBEATS-I	0.753	0.8219	0.820	0.911	0.799	
NBEATS-I+G	0.752	0.814	0.819	0.889	0.797	

As model learns seasonal behavior and dependencies on given covariates across time series, minimal feature engineering needed to capture complex behavior

- DeepAR makes probabilistic forecasts through Monte Carlo samples used to compute consistent quantile estimates for all sub-ranges in the prediction horizon
- By learning from similar items, our method is able to provide forecasts for items with little or no history at all, a case where traditional single-item forecasting methods fail

Table 2: Results for methods implemented in GluonTS

Dataset	AutoARIMA	AutoETS	Prophet	DeepAR
SP500	0.975±0.001	0.982± 0.001	0.985±0.001	0.837±0.002
M4 Daily	0.024±0.000	0.023±0.000	0.090±0.000	0.028±0.000
M4 Monthly	0.097±0.000	0.099±0.000	0.132±0.000	0.135±0.003
M4 Quarterly	0.080±0.000	0.078±0.000	0.123±0.000	0.091±0.001
M4 Weekly	0.050±0.000	0.051±0.000	0.108±0.000	0.072±0.003
M4 Yearly	0.124±0.000	0.126±0.000	0.156±0.000	0.120±0.002

Dataset	NPTS	Transformer	CNNQR	DeepAR
SP500	0.832±0.000	0.836±0.001	0.907±0.006	0.837±0.002
M4 Daily	0.145±0.000	0.028±0.000	0.026±0.001	0.028±0.000
M4 Monthly	0.233±0.000	0.134±0.002	0.126±0.002	0.135±0.003
M4 Quarterly	0.255±0.000	0.095±0.003	0.091±0.000	0.091±0.001
M4 Weekly	0.296±0.001	0.075±0.005	0.056±0.000	0.072±0.003
M4 Yearly	0.355±0.000	0.127±0.004	0.121±0.000	0.120±0.002

Some images

Figure 1: First time series

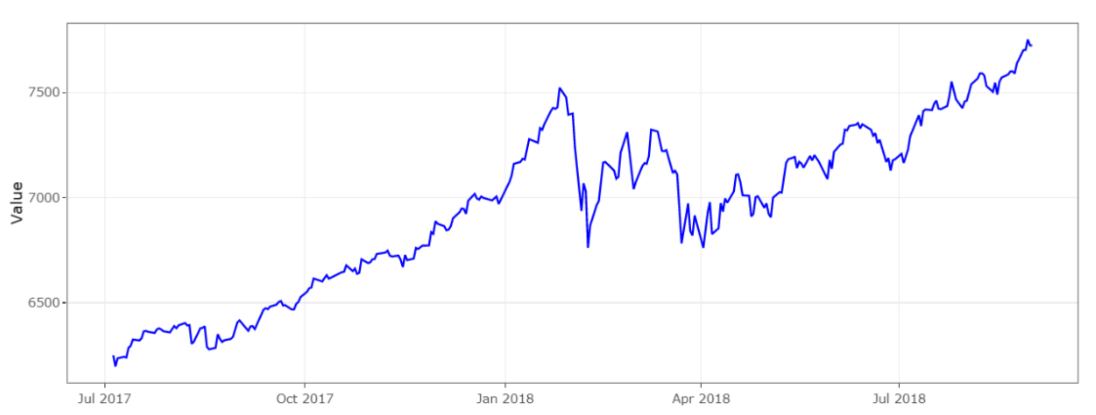
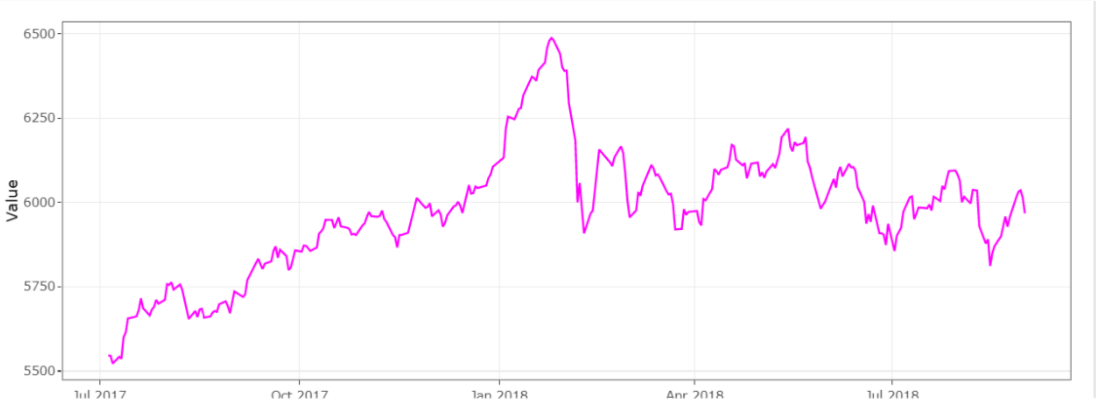


Figure 2: Second time series



References

Baltas, N. and Scherer, B. (2019). "Tail Risk in the Cross Section of Alternative Risk Premium Strategies." In: The Journal of Portfolio Management.

Li, Y., Turkington, D., and Yazdani, A. (2020). "Beyond the Black Box: An Intuitive Approach to Investment Prediction with Machine Learning." In: The Journal of Financial Data Science.

Rasekhschaffe, K. and Jones, R. (2019). "Machine Learning for Stock Selection." In: Financial Analysts Journal.