



Machine Learning Forecasting Models

Pedro Gusmão
Clony Abreu
Bruno Mendes

Neural Networks



- Basically, they consist in many thousands of coefficients (i.e., **weights**) that are adjusted in order to produce the best possible outputs (i.e., **forecasts**).
- These coefficients belong to many calculations that depends on the type of architecture (MLP/LSTM/CNN), hyperparameters (units, regularization, number of layers), etc... chosen. However, that is a topic for another day...

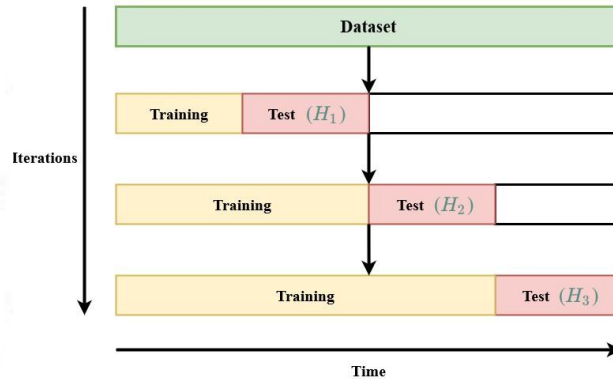
Neural Network (Stages)



- In order to ensure that forecasts are optimized, the models need to go through a **validation** phase.
 - Demands high computation and time (sometimes days).
 - But how frequently is this needed? (ongoing study).
- Even in **production** things might still take a while...

Validation Phase

1. Training with increasing data subsets;
2. Testing: evaluating the performance of the model against other data subsets.
 - a. Average the performance and **retry** with other **hyperparameter** options (see what's best).



Prediction Phase



1. Training with the whole data using the best hyperparameters;
 - a. Use training checkpoints (persistency) every time new data is included to alleviate resource usage and shorten training time.
 - b. Assess if the model needs **revalidation**.
2. Predict using that same model.

Store [0057](#) (40k data samples): predicting the last 30 days with four different LSTM architectures with similar hyperparameters

	1	2	3	4
MAAPE (%)	18.422	19.190	19.279	18.026
Training time (s)	~660	~900	~2150	~1250

These models were not hypertuned.

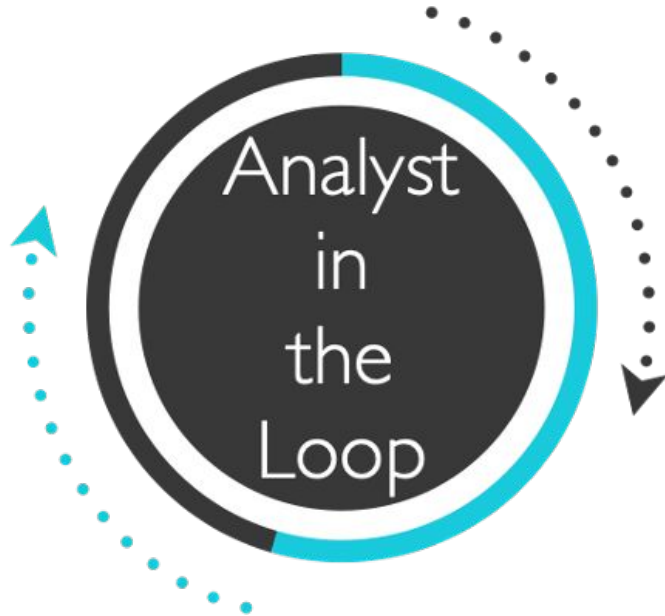


Store [4969](#) (24k data samples): predicting the last 30 days with four different LSTM architectures with similar hyperparameters

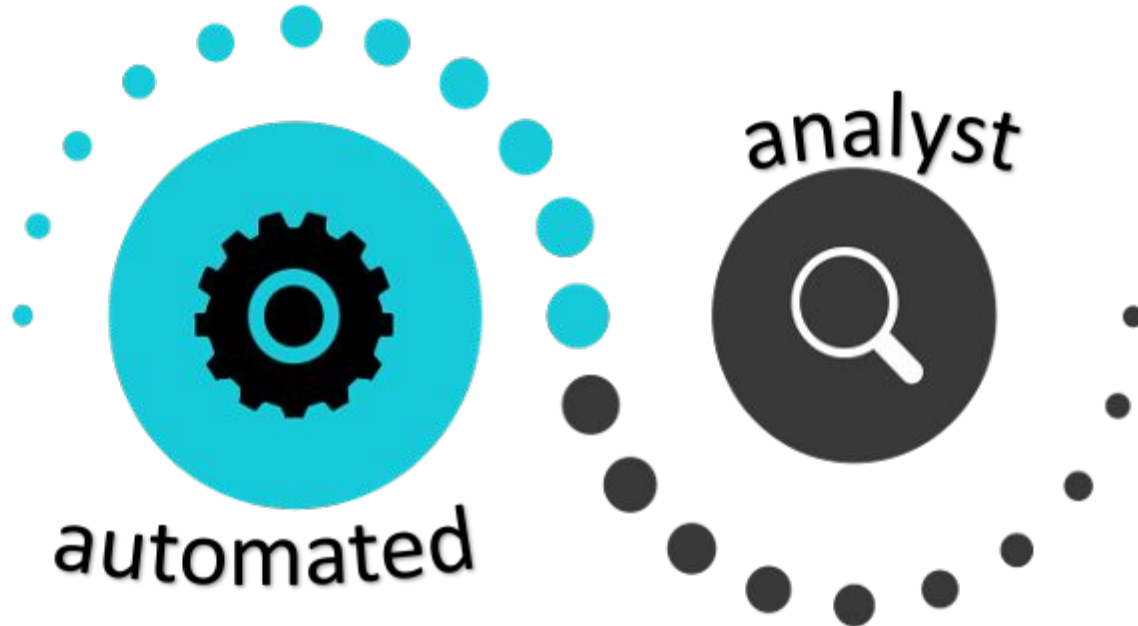
	1	2	3	4
MAAPE (%)	25.609	26.303	25.949	25.875
Training time (s)	~450	~590	~1300	~700

These models were not hypertuned.

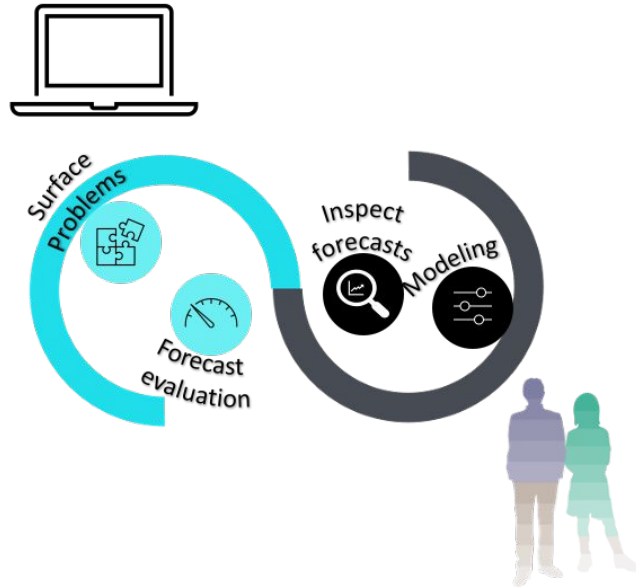
Forecasting with Prophet Concept



Forecasting with Prophet Concept



Forecasting with Prophet Concept





Prediction Phase (Production)

Store [11](#) (44.5k data samples): predicting the last 30 days with Prophet

	1
MAPE (%)	36.578
Prediction and training time (s)	~180

This model were not hypertuned.



Prediction Phase (Production)

Store [11](#) (2.7k data samples): predicting the last 30 days with Prophet

	1
MAPE (%)	32.003
Prediction and training time (s)	~7.6

This model were not hypertuned.



Prediction Phase (Production)

Store [11](#) (1.2k data samples): predicting the last 7 days with Prophet

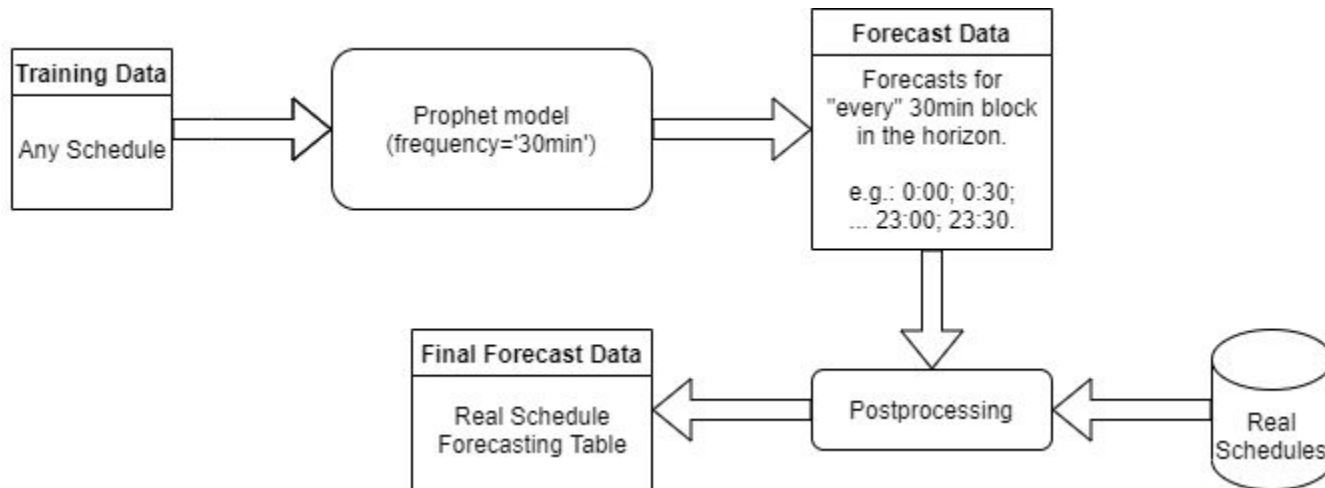
	1
MAPE (%)	18.534
Prediction and training time (s)	~6.9



Schedule Consistency

- Some machine learning models, like Facebook Prophet, need a defined frequency in order to perform a forecast.
 - E.g.: daily, hourly, 30min, 15min
- As for neural networks, these perform better with a consistent number of blocks for each forecast.
- Different schedules must be treated in the preprocessing and postprocessing phases.

Schedule treatment for Facebook Prophet



Schedule treatment for LSTM

