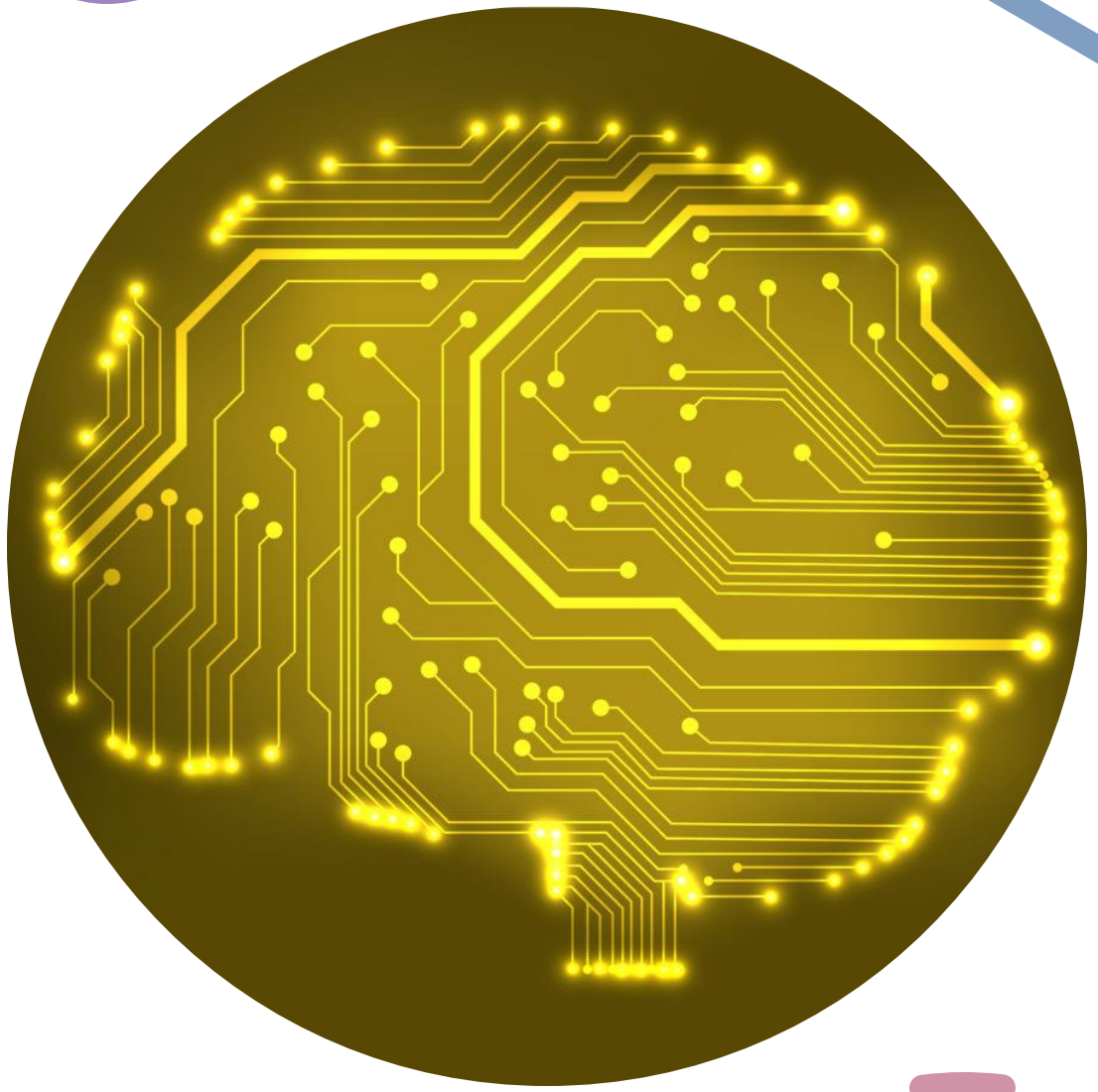


Time Series Forecasting

Matthew Davis, Data Science Manager

Ironside Boston, MA

Objectives

- General Overview of time series forecasting
- Introduction to ARIMA / SARIMA
- Introduction to Prophet
- Introduction to Temporal Fusion Transformers

What is Time Series Forecasting?

- Supervised Statistical Technique to predict future values along a time sequence, or time series
- Times Series Data: measurement of a continuous variable over-time (usually at regular time intervals)
- Forecast Horizon: number of time periods ahead you plan to predict
- Granularity Level: Time generally a time frequency (daily or monthly, etc…)
- Seasonality: recurrent patterns that have a fixed period.
- Exogenous Data: External Factors that are know to influence future values of the time sequence
- Metrics: RMSE, R2_Score, Correlation, MAPE, Logloss

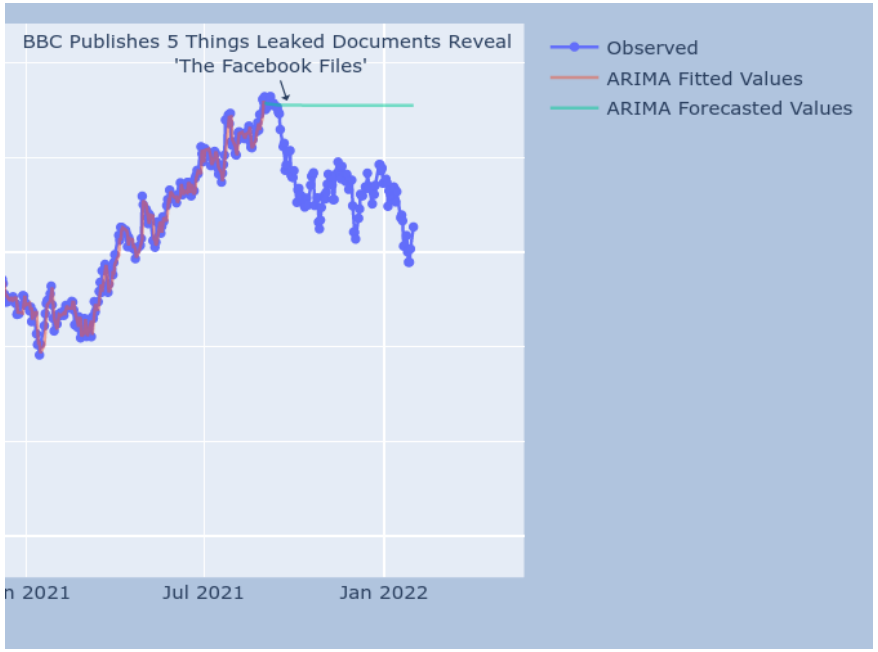
- Overview of time series forecasting: https://medium.com/analytics-vidhya/time-series-forecasting-a-complete-guide-d963142da33f

ARIMA / SARIMA

- Time series forecasting involved carefully picking by hand 7 parameters:
- p: The number of lag observations the model will use (ACF Plot)
- d: The number of times that the raw observations are differenced. (stationarity)
- q: The size of the moving average window. (PACF Plot)

- For seasonal data we need to add also the following:
- P: The number of seasonal lag observations the model will use
- D: The number of times that the seasonal observations are differenced
- Q: The size of the seasonal moving average window.
- m: The number of observations of 1 season

- Rules for picking parameters: https://people.duke.edu/~rnau/arimrule.htm
- Source: https://predictivehacks.com/arima-model-in-python/



ISSUES with ARIMA / SARIMA

- Difficult to automate, yes you can 'for loop' through it but nonsense can result
- Doesn’t handle multiple seasonality well
- Treats every time sequence individually
- Doesn’t handle missing data, or irregular time intervals well
- Needs refitting when new data is available
- External regressors must be known for future prediction dates

Prophet Forecasting tool

- Develop by Facebook (Meta) for the purposes of automating individual time series forecasting
- Handels natively holidays, and multi frequency data with daily/weekly/ monthly, seasonality
- Has change point detection
- Code: https://github.com/facebook/prophet
- Original paper: https://peerj.com/preprints/3190.pdf

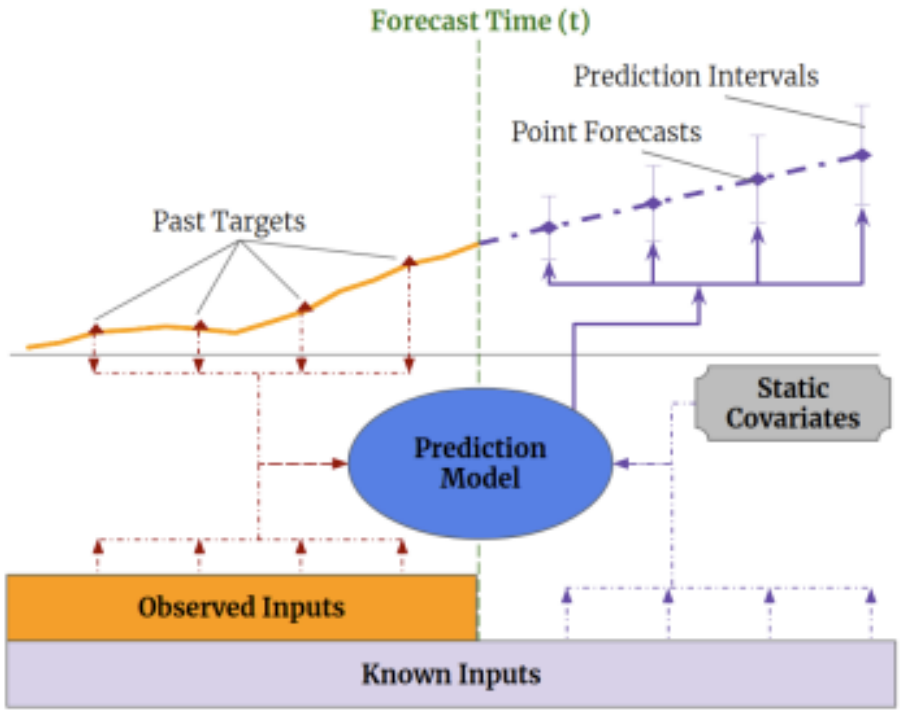
Issues with Prophet

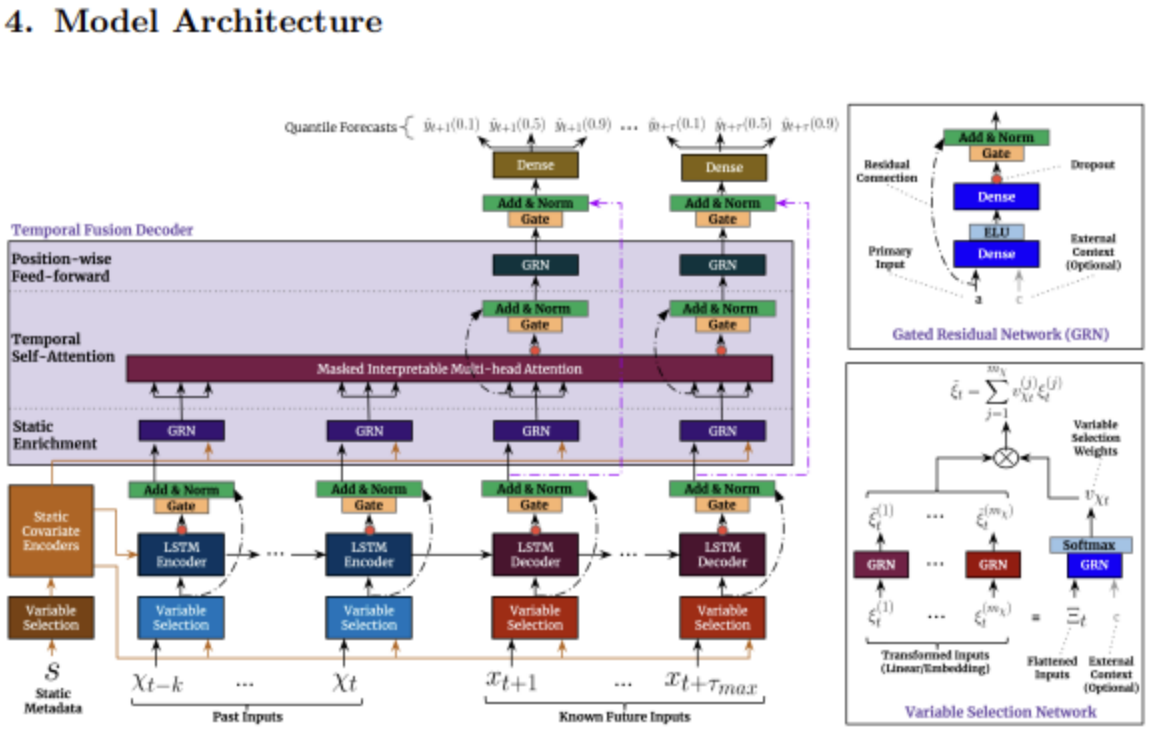
- Needs refitting when new data is available, but can warm start.
- Very sensitive to change point parameters
- External regressors must be known for future prediction dates
- Sensitive to trend and seasonality controls

Temporal Fusion Transformers

- Neural Network Approach, implemented in torch.
- Handels multiple time series in a single model
- Allows for meta data, as well as exogenous variables
- Natively built in uncertainty in forecast method
- Code https://pytorch-forecasting.readthedocs.io/en/stable/tutorials/stallion.html
- Original paper: https://arxiv.org/pdf/1912.09363.pdf

TFT Architecture





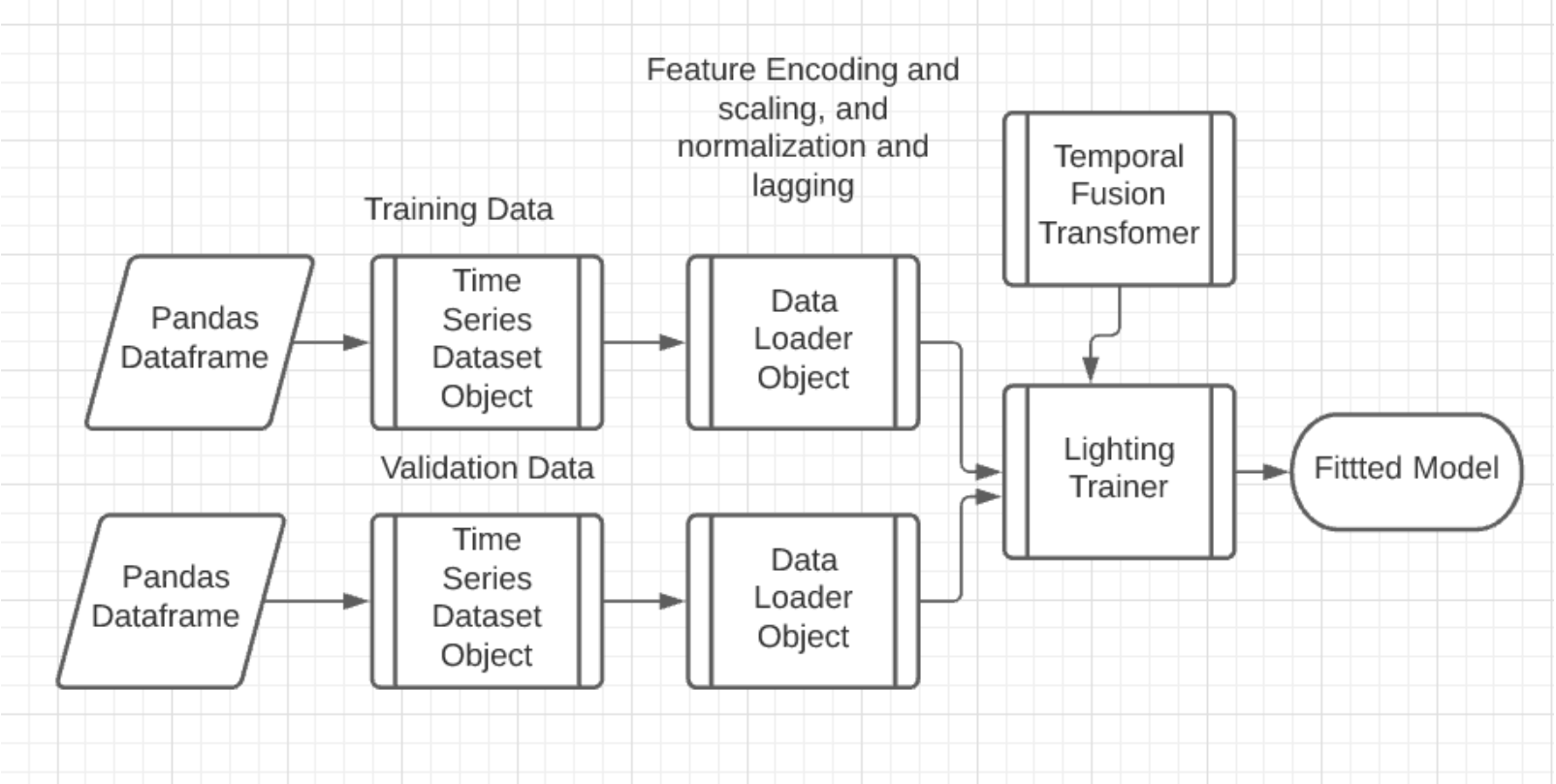
Source: https://arxiv.org/pdf/1912.09363.pdf

TFT Fitting

TFT leverage TimeSeriesDataset Objects to automatically lag, encode and normalized data.

TimeSeriesDataset Objects are used to create data loaders that back in data.

Pytorch lighting trainer trains the TFT model using data loaders as inputs



Issues with TFTs

- Neural Net can easily give nonsense results
- Not robust to outliers, sparsity
- Large search required to optimize parameters

Questions?