Learning Lab 38: Introduction to Modeltime for Time Series Forecasting

<https://github.com/business-science/modeltime>

<https://business-science.github.io/modeltime/index.html>

# Theory

|  |  |
| --- | --- |
|  |  |
|  |  |
| One way to improve the modelnig accuracy is to try different time series models on the data and choose the best one. There is no silver bullet. It takes a lot of time. | |
|  | |
|  | |
|  | |
|  | |
| Manually switching back and forth between different data structures. Also since it was done by diferent people, they each handle things a little differently. | |
|  | |
| Advantage 1: Leverages tidymodels. <https://www.tidymodels.org/find/parsnip/> | |
|  | |
| Step 1: Create models  Step 2: Calibrate against test dataset  Step 3: Refit the data to the full data (maybe after removing bad models) | |
| <https://business-science.github.io/timetk/> | |
| Timetk takes care of tidying, visualization and transformation. Modeltime takes care of the modeling part now. | |
|  | |
|  | |
|  | |

# Code

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
|  | | | |
| Plot\_time\_series is from timetk library, takes dates and values | | |  |
|  | | | |
| Cleans up 0’s and inputes with some algorithm. | | |  |
|  | | | Cumulative = True takes the rest as train data. |
|  | | | |
| Fit using Auto ARIMA on the training dataset.  This comes from the forecast package. |  | | |
| It produced a Seasonal ARIMA model. | This comes from Parnsip, but we have t pass some numeric features (it does not take date directly). | | |
|  | This time no trend. | | |
| Prophet | Random Forest and XGBoost | | |
| SVM with various kernels.  Note that we have not tuned the models. That comes in his course. | Boosted models on Prophet and ARIMA. Most ML models don’t know how to deal with data. In this case, it passes date to Prophet and other features to XGBoost. | | |
|  |  | | |
| Workflow | Each cell has the residuals, forecasting error, etc. (nested). | | |
|  |  | | |
| Produces forecasts |  | | |
| Control legend using args | | | |
| Refit to full data | | Same forecast function as before but with h for horizon. | |
|  | | Now we do model averaging | |
|  | | Does not handle ensembling (super learners) right now but that is coming… (see next steps) | |

# What’s next

|  |  |
| --- | --- |
|  | Advanced course has hyperparameter tuning.  Lag analysis, feature engineering. |
|  | Model models coming.  Super learners (advanced ensembling) – being created by tidymodels.  Deep Learning (from Python). |
|  | 3 parts in the course   1. efficient exploration and feature analysis 2. Model building – adjusting parameters, etc., ML algorithms, hyperparameter tuning, ensembling, stacking. 3. Deep Learning – combining R with Python |