October 2020

A Statistical Decomposition Based Neural Network For Multivariate Time Series Forecasting

Kenneth C. Lichtendahl Jr.

Haoyun Wu

Weijie Shen

Chi Po Choi

Chris Fry

A Statistical Decomposition Based Neural Network For Multivariate Time Series Forecasting

ABSTRACT

Machine learning based time series forecasting methods are popular and can match the performance of statistical models, in terms of accuracy, scalability, speed, etc. This disclosure presents techniques that incorporate statistical modeling into a neural network framework. The hybrid time series forecasting model described herein is named Seasonality Trend AutoRegressive Residual Yeo-Johnson power transformation Neural Network (STARRY-N). STARRY-N combines the advantages of residual neural network structure (such as N-BEATS) and explainable statistical forecasting models (such as TBATS). The model utilizes a neural network structure with separate stacks for trend, power transformed trend, seasonality, residual correction, and covariate adoption such as holiday effects. STARRY-N has good accuracy and is an explainable forecasting model.

KEYWORDS

- Forecasting
- Multivariate time series
- Deep learning
- Autoregression
- Exponential smoothing
- N-BEATS

BACKGROUND

Time series forecasting is an important business problem with applications across a wide range of domains such as weather forecasting, sales forecasting, capacity planning, inventory management, etc. Traditionally, statistical techniques based on smoothing, autoregression, and decomposition have been used for time series forecasting. Machine learning based time series forecasting methods have become increasingly popular and can match the performance of most statistical models. However, unlike traditional statistical models, many machine learning based techniques suffer from being opaque (black box) and devoid of insights on the modeling dataset. This makes it difficult to materialize the gained accuracy to business intuition and facilitate actionable decisions and to diagnose problems during intervention by analysts.

N-BEATS [1] is a popular machine learning based forecasting model that is accurate, intuitive, scalable, and fast. N-BEATS focuses on basis expansion of a time series and learns the basis coefficients using neural networks across different series. The N-BEATS model is composed of multiple stacks which can be used to represent time series components such as trend or seasonality, or to carry out estimation of expansion coefficients and basis vectors. The input for each stack is the backcast residual from the previous stack and the output from each stack is the forecast residual. The final model forecast is generated by summing up the forecast residuals of all the stacks. Additionally, outputs from individual stacks are also interpretable by a business analyst to gain insights.

The N-BEATS model can achieve a high level of prediction accuracy; however, the underlying models still have risk for extrapolation and explainability. TBATS [2] is a statistical decomposition based model for univariate time series, which features Box-Cox transformation, ARMA errors, exponential smoothing trend ,and trigonometric seasonality.

DESCRIPTION

This disclosure presents techniques that incorporate statistical modeling into a neural network framework for fitting parameters into an architecture similar to N-BEATS. This disclosure describes a hybrid time series forecasting model named Seasonality Trend AutoRegressive Residual Yeo-Johnson power transformation Neural Network (STARRY-N). STARRY-N is an algorithm that combines the advantages of residual neural network structure (such as N-BEATS) and explainable statistical forecasting models (such as TBATS).

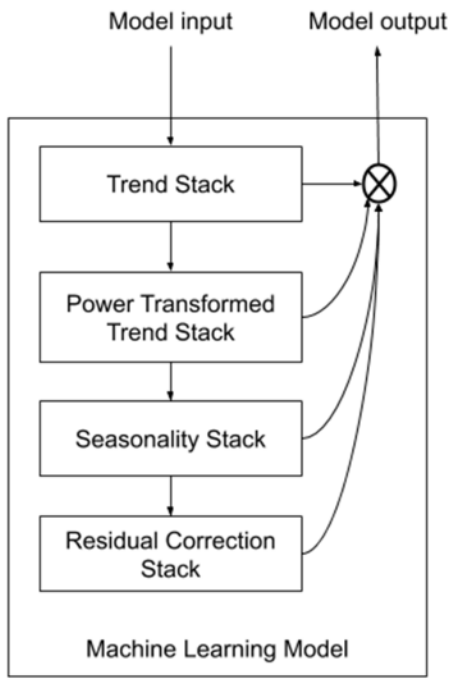

Fig. 1: STARRY-N Architecture

Fig. 1 illustrates the STARRY-N model architecture. The model utilizes a residual neural network structure composed of stacks for trend, power transformed trend, seasonality, residual correction, and a separate stack for covariate adoption. Similar to the N-BEATS model, each stack in STARRY-N consists of multiple stacks (or blocks), where the backcast residual of each stack is provided as input to the next stack. The final model forecast is the aggregated sum of forecast residuals from all stacks.

For the trend stack, the STARRY-N model adopts the state-space model for the trend component within the TBATS model as the underlying model. The trend is modeled as a combination of long term trend and a damped short term trend that converges to zero in the long term. The neural network learns the smoothing parameters and the damping factors for each time series and then generates the fits and forecasts.

For the power transformed trend stack, the Yeo-Johnson power transformation is added on top of the previous trend stack to enable nonlinear fits and forecasts. The neural network learns the appropriate transformation parameter, transforms the input data, and performs fit and forecast operations using the same state-space model in the regular trend stack. The output from this stack is inverse transformed to obtain the nonlinear forecasts.

The seasonality stack consists of a design matrix with binary entries to indicate the dates/weekdays (or other time periods) for each observation. This design matrix is used as the basis function in the neural network which learns the coefficients that are applied to the matrix. An alternative structure is to use Fourier series as the basis function with periods being the seasonality periods of interest.

For the residual correction stack, the remainder residuals are modeled as a simple first order autoregressive process to perform short term forecast bias correction. The neural network learns the autocorrelation parameter for the autoregressive process for each time series and then uses it to make forecasts.

For all the above-mentioned stacks, the neural network that learns the parameters is a combination of fully connected layers, normalization layers, and dropout layers. Fig. 2 illustrates an example of the conceptual architecture for the seasonality stack.

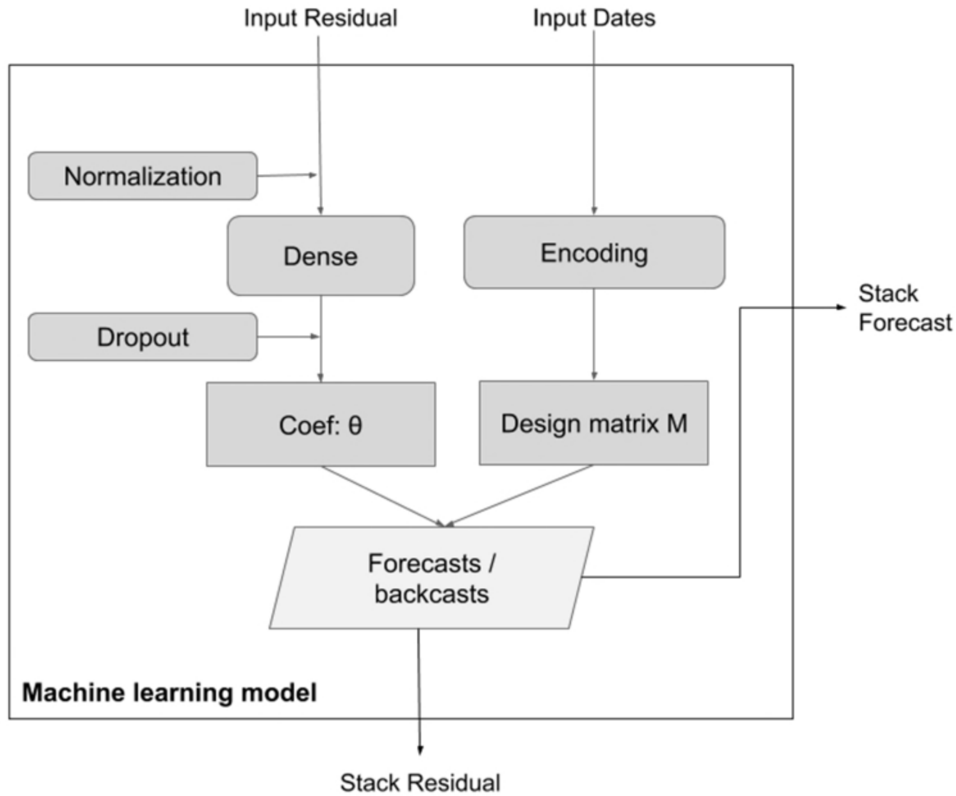

Fig. 2: Seasonality Stack

The seasonality stack takes as input dates for the training dataset and residual from the previous stack (the power transformed trend stack, as shown in Fig. 1). The input residual is passed through normalization layer(s), fully connected layers (“Dense”), and dropout layer(s) to estimate the coefficients to be used for the generation of forecast and backcast. Dates are processed through an encoding block to obtain a design matrix M, which is multiplied with the coefficients (θ) estimated by the neural network. The seasonality stack (like other stacks in STARRY-N) generates a forecast that is an input for the computation of the final forecast and a residual that is fed as an input to the next stack (the residual correction stack, as shown in Fig. 1).

If the input data has covariates, the covariates are passed through a separate covariate adoption block which is used to preprocess the input depending on the data type. Categorical covariates are processed through embedding layers, while numerical covariates are processed through batch normalization and dense layers. After preprocessing, the covariates are combined with the original time series and the combined series is passed through the trend, seasonality, power transformation, and residual correction stacks. The final model forecast is generated by summing up the forecast residuals of all the stacks, as described with reference to Fig. 1.

The STARRY-N model can produce comparable or lower error metrics when compared to N-BEATS. STARRY-N has good accuracy and is an explainable forecasting model.

CONCLUSION

This disclosure presents techniques that incorporate statistical modeling into a neural network framework. The hybrid time series forecasting model described herein is named Seasonality Trend AutoRegressive Residual Yeo-Johnson power transformation Neural Network (STARRY-N). STARRY-N combines the advantages of residual neural network structure (such as N-BEATS) and explainable statistical forecasting models (such as TBATS). The model utilizes a neural network structure with separate stacks for trend, power transformed trend, seasonality, residual correction, and covariate adoption such as holiday effects. STARRY-N has good accuracy and is an explainable forecasting model.

REFERENCES

1. Oreshkin, Boris N., Dmitri Carpov, Nicolas Chapados, and Yoshua Bengio. "N-BEATS: Neural basis expansion analysis for interpretable time series forecasting." *arXiv preprint* arXiv:1905.10437 (2019) available online at http://arxiv.org/abs/1905.10437
2. De Livera, Alysha M., Rob J. Hyndman, and Ralph D. Snyder. "Forecasting time series with complex seasonal patterns using exponential smoothing." *Journal of the American Statistical Association* 106, no. 496 (2011): 1513-1527.