

Experimenting with Scaleable Forecasting Models

Matthew Coghlin Pelumi Dacosta Joseph Despres Aneesh Gahi

Joseph Sigler

December 2, 2021

1 Introduction

The ability to generate accurate forecasts is of great interest organizations that need to plan for uncertainty. Governments, businesses, and non-profit organizations are all forced to make forecasts when planning their various activities. As datasets grow, organizations require more and more items forecast. A time-series expert is generally required to carefully tuning ARIMA model parameters (Taylor & Letham, 2018). Although, this approach is riggerois, it does not scale. Analysts are being asked to generate high quality forecasts for thousands, and even hundreds of thousands of series at a time (m3). The goal of this project is to implement algorithms that generate high quality forecasts with minimal involvement, validate them with a training and testing partitian, and generate ensemble predictions of the different methods.

2 Data

The data obtained for this project are provided as part of a kaggle challenge which we are participating. The competition challenges participants to forecast daily retail sales demand (*Store Item Demand Forecasting Challenge*, 2018). As contestants, we are given 5 years of training

data, with the daily sales of 50 different products from ten different stores. This is a total of 913,000 data points to train forecasting models. The goal is to forecast the next 90 days for each of the 50 products and 10 stores. Judged by Mean Absolute Scaled Error (MASE)

$$SMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{(|A_t| + |F_t|)/2}$$

The data required very little preparation. There were several datapoints that were zero we switched to a one because the statsmodels implimentzation of vector autoregression did not support zero values. We only did that once. After that we combined the stores and items into one string column to avoid nesting loops when iterationg over stores and items.

EXPLOATORY PLOTS AND DESCRIPTIVE STATISTICS These data are highly seasonal with a slight upward trend. These data have significant noise. the vast majority of data involving human activities have a seasonal and trend component (Taylor & Letham, 2018).

After that, we separated into a training and testing partitians to test. We used the first 1279 (80

3 Experiments

The project will use that clean data and input it into a feature detection algorithm which quantifies various features of each series. Then Takes the data and runs six different forecasting algorithms to generate a prediction for the remainder of the testing data. Then analyze forecasting errors. Then attempt to use machine learning methods to select the best model for the data given features

FLOW CHART

Running Multiple Forecast Models

Although there are many forecasting models to choose from, there is not much research on when a given forecasting model outperforms another. This is an open field and we are going to test different ones and evaluate them strictly on their performance on testing data. Due to the data having strong weekly swings, we implement several models with an autoregressive terms.

Exponential Smoothing

Exponential smoothing, or this is sometimes known as the holt winters method. decomposes the timeseries into three components: Seasonality, Trend, and slope. With the effect of seasonality and trend to be linear.

XGBoost

XGBoost is an implementation of a tree boosting system. This uses decision tree regression, and fits an ensemble of models fit to the data, then the residuals, then fits the residuals residuals. This ensemble of tree boosting is quite robust and is useful for a variety of different regression and

classification tasks.

Facebook's Prophet

Facebook released a forecasting library designed specifically to meet the challenges of generating many high quality forecasts. The model prophet is a General Additive model, that consists of three functions, trend which fits an aperiodic logistic population growth model (we did not limit the growth, however that is a parameter), seasonality is a fourier series fit to the remaining seasonal component, and a holiday parameter which is a vector of user specified holiday periods, the holiday periods saw a drop in sales, however not enough to justify us specifying specific dates.

Vector Auto Regression

The first algorithm we implement, is an autoregressive model. This takes the first 5 lag positions and uses them as regressors, then using timeseries decomposition, it models the seasonality. This model is commonly used to forecast economic variables.

NeuralProphet

NeuralProphet, is a forecasting library that expands on facebook prophet and includes an autoregressive term in the general additive model and uses neural networks to generate the autoregressive terms in the model.

Seasonal Auto Regressive Distributed Lag

ARDL models add to the above autoregressive model, however in addition to seasonality with is fit with a vector of indicator variables. and trend, in this case we are adding an explanatory variable of time and fitting the model to lags of time.

Feature Detection

There are many different shapes and patterns a timeseries plot can take. Tsfeatures (Montero-Manso, Athanasopoulos, Hyndman, & Talagala, 2020) quantifies different features such as the number of times it crosses the median, degree of seasonality, the number of flat spots, entropy, and many more.

We ran this algorithm on each series collecting 30 quantified features. The central question of this study is to determine if we can determine which forecast will perform the best given these features.

Predict Which Day the Model Has the Smallest Error Given the Features

Using different models, we tried knn, logistic regression, XGBoost, and an artificial neural network.

Due to the way kaggle scores, final Predictions could not be a weighted average of the results. We selected which was most likely to be the final.

4 Conclusions

The models ensemble was able to outperform each of the individual. In practice, it is likely that we are too use these methods to forecast a months sales. this would be aggregated.

References

Montero-Manso, P., Athanasopoulos, G., Hyndman, R. J., & Talagala, T. S. (2020). Fforma: Feature-based forecast model averaging. *International Journal of Forecasting*, 36(1), 86–92.

Store item demand forecasting challenge. (2018).

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45.