Ensemble of Time Series forecasting for Complex Structure

Shanu Agrawal and Balasubramanyam Pisupati

Forecasting is very important for proper management. It becomes difficult when companies grows in terms selling more products or customer. Due to decrease in transparency for each material or customer. And also it becomes difficult to find one forecast model that performs best for all possible series and all horizons. In this paper an approach for forecasting using ensemble model discussed. Ensembling is done using variance based error approach. For validation of forecast model is done using rolling forecast.

1. Brief statement of Research

Forecasting is necessary for better business understanding and decision. When it is done manually, requires lot of effort and time from multiple department like logistics, sales and finance. It also involves lot of gut feeling from experienced people and sometime it might lead to error prone prediction if person is very new in field or not aware of past behavior/ pattern of series. In this paper objective is to automate forecast process such that good accuracy can be achieved and provide this as base model to departments to improve their decision power.

2. Literature Review

In Spyros and Michel's's paper proved that "The accuracy when various methods are combined outperforms, on average, the individual methods being combined and does very well in comparison to other methods". In this paper, various methods for error calculation also discussed like symmetric MAPE, Average Ranking, Median symmetric APE, Percentage Better, and Median RAE.
In 'Forecast combinations in R using the forecast combinations package Manual' written by Eran described various approaches for combination: Simple average, Ordinary Least Squares (OLS) regression, Least Absolute Deviation (LAD) regression, Constrained Least Squares (CLS) regression, Variance-based, or Inverse Mean Squared Error and best Individual model.

In this paper, variance based combination and best Individual model method with symmetric MAPE and MAPE instead of mean square error. This approach assign more weight to more accurate forecasting models according to mean square error. An approach successfully applied in Stock and Watson (2004) for combining forecasts of output growth. Put simply, more accurate forecasting methods (lower MSE) are weighted more heavily. MSE here is computed based on out-of-sample forecasts, sometimes referred to as Mean Squared Prediction Error. The combined forecast is given by.

$$f^c = \frac{\left(\frac{MSE_i}{\sum_{i=1}^{P} MSE_i}\right)^{-1}}{\sum_{i=1}^{P}\left(\frac{MSE_i}{\sum_{i=1}^{P} MSE_i}\right)^{-1}} f_i = \frac{\frac{1}{MSE_i}}{\sum_{i=1}^{P}\frac{1}{MSE_i}} f_i.$$

For study, M3 completion (https://forecasters.org/resources/time-series-data/m3-competition/) monthly data, around 1428 different time series are used shown in table. Number of observation is different for each time series, it varies from 66 to 144 and has different behaviour. The required forecast for each time series is 18 month ahead.

Time interval between successive observation	**Category of data**	**Number of time Series**
Monthly	Micro	474
Monthly	Industry	334
Monthly	Macro	312
Monthly	Finance	145
Monthly	Demographic	111
Monthly	Others	52

3. The Methodology

In this paper for forecasting ARIMA, ETS, Naïve, Snaive, theta, STL used as base model. And then combined them for better forecast for overall.

Normally for checking the efficiency of forecast model, forecasters use rolling forecast approach. For example, for training data from Jan 2011 till Dec 2017 given, require

Training Data	**1**	**2**	**3**	**4**	**5**	**---**	**10**	**11**	**12**
Jan11- Dec15									
Jan11- Jan16									
Jan11- Feb16									
……									
Jan11- Nov17									
Jan11- Dec17									

forecast from Jan 2018 to Dec 2018. Forecast for multiple training data calculated by adding one month at a time as shown in table 1.

For each horizon, forecasters calculate an average mean absolute percentage error to observe the accuracy using the following equation:

$$\mathrm{M} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

Where
n = number of training period
A_t = Actual value
F_t= Forecast value

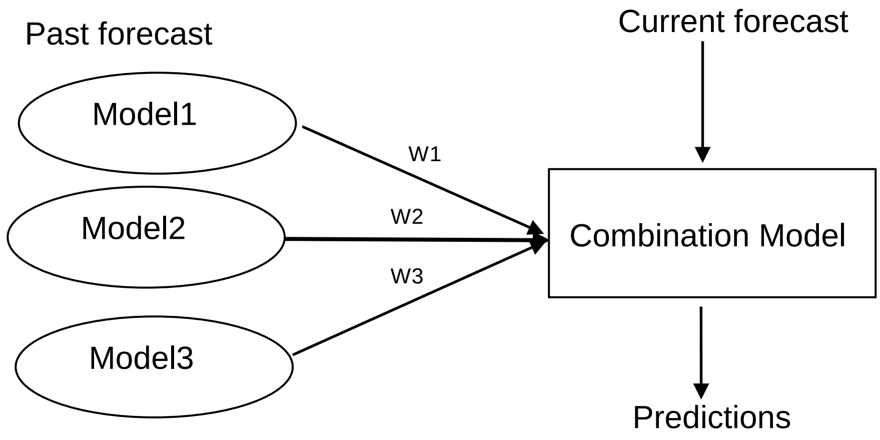
These process can be repeated for all mentioned forecast model. One model might perform well for prediction short term forecast and other model might perform well for long term forecast. It might be difficult for one particular forecast model to predict for all horizon accurately. Some other forecast model might perform better than existing model, when behavior of time series varies. To avoid selection of forecast model manually according to error, we can automate the whole process by assigning weight to each model according to mean square error or any other error method suitable for case.

Build the forecast using all the appropriate forecast model, then combine them in such a way the error will be less. In this study, forecast generated using 12 different time series model for each time series: auto regressive moving average model, exponential smoothing model, state space model, bayesian statistical time series model, theta model, neural network, trigonometric, forecast using gradient boost, multiple aggregation prediction algorithm, naïve and snaïve models used as the base model.

In this paper, single forecast is derived using variance based approach using symmetric MAPE and MAPE. This approach assign more weight to more accurate forecasting models according to error. Errors are computed based on past forecast errors. The combined forecast is given by.

$$f^c = \frac{\left(\frac{MSE_i}{\sum_{i=1}^{P} MSE_i}\right)^{-1}}{\sum_{i=1}^{P}\left(\frac{MSE_i}{\sum_{i=1}^{P} MSE_i}\right)^{-1}} f_i = \frac{\frac{1}{MSE_i}}{\sum_{i=1}^{P} \frac{1}{MSE_i}} f_i.$$

For calculation, it requires past forecast and actual value to derive the weight. This process will be repeated for each horizon. For example, from above table we can generate combine forecast from training data Jan11- Dec16 onwards. For training data Jan 11-Dec16 for horizon 1, all horizon 1 forecast from previous training data, will be used for weight calculation. Once we get weight for each model, current training data forecast will be multiplied to get combined forecast as shown in figure 1. Similarly for horizon 2, using all previous training data horizon 2 forecast weight will be derived for each model and then current period training data forecast will be multiplied to get combined forecast. This process will be repeated for all required horizon. To check the robustness of model rolling approach can be used.



In many applications the time series organized in a hierarchical structure based on dimensions such as region, customer and product type. It is usually forecasted using grouped time series, and rrequires forecast for each combination series, adjusts the forecast such that bottom up forecast matches with top down forecast. Behaviour of each series can be different from other, using one model for all series may not give good accuracy. In these cases also combined forecast model works well.

4. The Findings

As mentioned, for study used M3-competion monthly data used excluding others category. In chart error for two category shown for 1 month ahead, 6 month ahead and 12 month ahead error.

MODEL	ACCURACY 1 MONTH	6 MONTH AHEAD	12 MONTH AHEAD
ARIMA			
ETS			
BSTS			
COMBINATION MODEL			

5. Summary and Conclusions

Researchers have used several techniques and algorithms to address the challenges in predicting.This paper intends to help forecasters in automating the forecast process, which can work for almost all types of time series which are predictable. It is also helpful for planners for preparing quick, informative and effective decisions in predicting using a robust combination model, don’t need to depend on modeller for updating the model for changes. Combined model quickly adopt the new model if observe the changes in behaviour, which is high in accuracy, as model finds best combination according to previous performance.

End Notes

References

Eran Raviv, Forecast combinations in R using the ForecastCombinations package

Rob J Hyndman, George Athanasopoulos, Han Lin Shang, hts: An R Package for Forecasting Hierarchical or
Grouped Time Series