Project Proposal

Comparing Facebook Prophets Forecasts to those an ARIMA Model: A Diebold-Mariano Evaluation of the JSE Top40 Index

Marvelous Mubenesha
Student number : MBNMAR005
mbnmar005@myuct.ac.za
Supervisor: Hanjo Odendaal

Submission date: 9 May 2017

1. Introduction

The problem of forecasting time series of stock returns has attracted the efforts of quantitative analysts' work for decades, the accuracy of forecasting models is pivotal to technical analyses of stock market returns. Traditionally, ARIMA (Autoregressive Moving Average) models have been used to forecast stock returns for their tractability and interpretability. With the evolution of computational power and statistical advancement, other models such as GARCH (Generalized Autoregressive Conditional Heteroscedasticity), ANNs (Artificial Neural Networks) and Bayesian modelling methods have been proposed. However, ARIMA models remain the standard approach which other methods are normally compared against.

Facebook open sourced their time series forecasting tool, Prophet whose main selling point is that it enables "forecasting-at-scale". A term they coined which is defined as "an approach that allows a large number of analysts to forecast a large number and variety of business time series" (Taylor et el, 2017, 3). It has the added advantage of being less 'expensive' than other alternatives. In addition, it is specifically preferable over ARIMA models because it is non-linear, flexible and accommodates varying time intervals (Taylor et el, 2017). Prophet uses a Bayesian Generalized additive model as opposed to the linear stochastic dependence of ARIMA models. This has the added advantage of potentially improving predictions, it translated the problem of forecasting into a curve fitting exercise with an analyst in the loop (Taylor et el, 2017).

The objective of this proposed research is to compare the forecasts generated by prophet which uses intensive Bayesian modelling to those obtained by an ARIMA model through a Diebold-Mariano Evaluation of the JSE Top40 Index. This is particularly useful because analysts from a wide background can be able to easily employ stock price return forecasting using their prior knowledge, without a wholesome understanding of the statistical intricacies involved. The modifiability also means that altering the model to incorporate prior knowledge is relatively easy given the easily interpretable parameters.

If prophet generates significantly improved forecasts for the JSE Top40 Index, the model's forecasts can be compared with those of other advanced models that are more complex but may be less tractable and scalable such as GARCH, Supervised and Unsupervised machine learning methods(Random forests, Clustering) as well as Artificial Neural Networks (ANN).

2. Preliminary Literature review

ARIMA models have widely been used to predict stock returns over the past 40 decades and are normally used as a benchmark against which other methods are compared (Zhang, 2003). Over the past two to three decades, with the evolution of computational power, other approaches have been introduced and consequently compared to ARIMA models, several studies have shown that ARIMA forecasts are as accurate as more expensive approaches (Kohzadi et el, 1996). Though there aren't any similar published studies done in South Africa based on the Top40 Index, studies using both stock return and non-stock returns data have been used to compare the forecasts of ARIMA models to alternatives. In 1996, Kohzadi et el compared the livestock commodity price forecasts of an ARIMA model to those of an Artificial Neural Network (ANN) using the Henrikson-Merton test. They found that ANNs were considerably more accurate than the ARIMA forecasts (Kohzadi et et, 1996). In 2015, Dudek compared Random Forests to ARIMA models to predict short-term electricity load, Kane et el also did a similar study to compare predictions of the avian influenza H5N1 outbreaks in Eygpt (Kane et el, 2014). Both studies concluded that the predictive ability of the Random Forest was superior to the ARIMA model when they compared the Mean Square Errors of the models. Despite such results, ARIMA models are still applied because they are interpretable and less computationally expensive when working on large datasets, a concern which is relevant to stock price information. The studies that have been considered so far have use alternative out-of-sample methods to compare forecasts, in this paper, we use the Diebold-Mariano evaluation method which was intended specifically for comparing forecasts (Diebold, 2013). It is useful because it does not require the use of pseudo-out-of-sample environments to compare forecasts, which means that the results have more power than alternative approaches (Diebold, 2013).

3. Further Reading

The preliminary literature review exposed me to the methodology previous studies have used to compare forecasts of different models. The following areas highlight papers I would like to consult that relate to specific issues I predict might arise as I execute the objectives of this paper.

1. Efficiency of markets in South Africa
2. If prophet produces better forecasts than the ARIMA model, consider comparing Prophet to other Machine Learning and Bayesian Time Series Forecasting methods such as Generalized Autoregressive Conditional Heteroskedascticity(GARCH) models, Vector autoregression (VaR)models , Markov Chain Monte-Carlo, Exponential Smoothing and K-Means. The following paper will be considered for guidance;

   Okkels, C.B. 2014. *Financial Forecasting: Stock Market Prediction*. Maters Thesis for University of Copenhagen, Faculty of Science.
3. Optimal model selection criteria for seasonal ARIMA model i.e whether to use the AIC or BIC.

4. Research question

The main objective of this paper is to compare the stock returns forecasts generated by the recently open-sourced Facebooks' Prophet Forecasting package with those of an ARIMA model constructed using the Box-Jenkins methodology. This will be done using the Diebold-Mariano evaluation of the JSE Top40 index. The research is therefore aimed at answering the question;

"How does the latest prophet forecasting model compare with an ARIMA model. A South African Top40 Diebold-Mariano Evaluation"

A secondary question that I hope the research could answer is whether intensive Bayesian modelling is worth the effort.

5. Methodology

The following is a high level outline of how I intend to answer my research question. The latest versions (as at May 2017) of R and RStudio will be used for model building and forecast comparison.

5.1. Data Collection and Preparation

Input data of the returns on the JSE Top40 index will be obtained from the Bloomberg terminal of UCT Libraries. My intention is to use post-crisis data from April 2010 to May 2017. The choice to use April 2010 as the starting point is consistent with the methodology employed by a similar study undertaken by Arabio in 2012, furthermore, the International Monetary Fund(IMF) considers the period after April 2010 to be the post-crisis period (Strauss,K, 2010). This data will be cleaned by detecting and localizing errors then making corrections and imputing if necessary. There will be no need to separate the data into in-sample and out-of-sample data since the Diebold-Mariano evaluation does not require this.

5.2. Model selection and estimation

Prophet forecast applies a quasi-Bayesian technique in the form of prior assumptions about the nature of the components of a time series, namely, the growth, seasonality and holiday components. Prior parameterisation might be unnecessary given that the system is optimized to be automated. My intention is to re-arrange the data into a data frame compatible with Facebook Prophet (i.e a column of dates and another column of log returns on the Top40 index) and then feed this into the estimation function to get results.

The initial intention was to compare the forecasts of the standard ARIMA(1,1) model to Facebook Prophets' forecasts, but to better compare the performance of the models, as has been done in literature for similar studies, I intend on comparing Prophet forecasts with those of a seasonal ARIMA(p,d,q)*(P,D,Q) model. The ARIMA model will be constructed on the full data set using the Box-Jenkins Approach which is explained in Pankratz, 2009. Before running the model, I intend on verifying the stationarity of the series. The BIC will be used as a selection criteria in order to introduce a Bayesian related concept that will then allow us to indirectly evaluate if <u>intensive</u> Bayesian modelling is worth the effort.

5.3. Forecasting, Back-testing and Evaluation of Results

Once the models have been constructed, the full sample data will be used to generate forecasts that will then be compared using the Diebold-Mariano evaluation. The results of which will be used to conclude whether Facebook Prophets' forecasts are significantly different from those obtained by an ARIMA model for the JSE Top40 Index. If results show that Prophet performs better, then I might consider exploring whether it out-performs other methods that have been shown to do better than ARIMA models such as the GARCH and ANN. These approaches are more complex and therefore, rarely used in practice to date, if Prophet does as well as them, then this could be an opportunity for a better performing yet simplistic and interpretable model to be used to forecast stock returns.

6. Proposed Structure of the paper

The following is an outline of how I intend to structure the paper.

1. Title Page and Abstract
2. Introduction: Brief overview of the paper and its purpose.
3. Background : A review of theory from the literature review that will be used in the paper
4. Methodology and data
5. ARIMA Model formulation: Box-Jenkins Approach

   5.1. Mathematical specification of the model: Description of the mathematical model and approach

   5.2. Model building: This section will illustrate the Box-Jenkins methodology that will be followed to build the ARIMA model. The following is an outline of the procedure.

   - Differencing to achieve stationarity
   - Model Identification
   - Parameter Estimation
   - Diagnostic Checking(Unit root tests)
   - Final Model Selection

5.3. Forecasting with Final ARIMA model

6. Prophet Forecasting Model

   6.1. Model formulation: Here, I will deconstruct the mathematical formulation of the quasi-Bayesian time series model through component analysis (Growth, seasonality, holidays and events).

   6.2. Model fitting: A description of the estimation procedure and the results obtained.

   6.3. Forecasting with Facebook Prophets' model

7. Diebold-Mariano Evaluation

This section will cover the comparison of the two forecasting models. I intend on breaking it down into the following sections;

   7.1. Brief description of the Diebold-Mariano Evaluation method and the

   the model assumptions that will be made.

   7.2. Results of Diebold-Mariano Evaluation when applied to ARIMA

   and Prophet forecasts obtained.

8. Discussion and Conclusions: A response to the research question given the results obtained in the study. Thereafter, I might consider what the implications of the results are given what has been observed and propose further areas of study.
9. References
10. Appendix

7. Timeline

The timeline below outlines key milestones & target dates in the research process.

Date	Task
9 May	-Submit proposal. -Continue reading literature and add to preliminary literature review. -Modify methodology when new information comes to light.
20 May	-Be comfortable with using markdown and dplyr. -Work on writing abstract, introduction and background using Markdown -Collect data, clean it and test coding skills with small sample. Identify issues and address them.
12 June	-Submit literature review.
23 June	-Exams
20 July	-Resume working of draft paper -Do final model building, forecasting, and verify results. -Do further reading if necessary. -Complete draft paper, engage with supervisor where necessary
11 September	-Submit draft paper. Maybe even earlier.
16 October	-Receive feedback on draft paper. -Modify paper based on feedback
6 November	-Deadline for submission of final paper.

8. References

Diebold, F.X., 2015. *Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests*. Journal of Business & Economic Statistics, 33(1), pp.1-22.

Dudek, G., 2015. *Short-term load forecasting using random forests*. Intelligent Systems' 2014 (pp. 821-828). Springer International Publishing.

Pankratz, A., 2009. *Forecasting with univariate Box-Jenkins models: Concepts and cases* (Vol. 224). John Wiley & Sons.

Kane, M. Price, N. Scotch, M.Rabinowitz, P. 2014. *Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks.* BMC Bioinformatics 15(1):276.

Kohzadi, N., Boyd, M.S., Kermanshahi, B. and Kaastra, I., 1996. *A comparison of artificial neural network and time series models for forecasting commodity prices*. Neurocomputing, *10*(2), pp.169-181.

Strauss-Kahn, D. 2010. *"Economic Policy Challenge In the Post-Crisis Period, A Speech BY The Managing Director Dominique Strauss-Kahn".* As prepared for delivery. International Monetary Fund.[Online]. Viewed 8 May 2017. Available at : http://www.imf.org/external/np/speeches/2010/041010.htm.

Taylor, S. Letham, B. 2017. *Forecasting at Scale.* Facebook Datascience. [Online]. Viewed 9 May 2017. Available at : https://facebookincubator.github.io/prophet/static/prophet_paper_20170113.pdf.

Zhang, G.P., 2003. *Time series forecasting using a hybrid ARIMA and neural network model.* Neurocomputing, *50*, pp.159-175.