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

Marvelous Mubenesha[a]

[a]*Honours Dissertation, University of Cape Town, South Africa*

Keywords: ARIMA Forecasting, Prophet Forecasting, Diebold-Mariano Evaluation

JEL classification

1. Introduction

The problem of forecasting time series of stock returns has attracted the efforts of quantitative analyst's work for decades, the accuracy of forecasting models is pivotal to the techinical analyses of stock market returns. Traditionally, Autoregressive Moving Average(ARIMA) 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 the Generalized Autoregressive Conditional Heteroscedasticity(GARCH)models, Artificial Neural Networks(ANNs) and Bayesian Modelling methods have been proposed. However, ARIMA models remain the standard approach against which other methods are normally compared.

Facebook open sourced their time series forecasting tool, Prophet, whose main selling point is that it enables "forecasting-at-scale". A term coined by the data science team 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). It has the added advantage of being less 'expensive' than other alternatives. In addition, it is specifically prefarable over ARIMA models because of its non-linearity, flexibility and ability to accomodate varying time intervals (Taylor et et 2017).

Email address: mbnmar005@myuct.ac.za (Marvelous Mubenesha)

Prophet uses a Bayesian Generalized additive approach as opposed to the linear stochastic dependence approach that ARIMA models rely on. This distinction could potentially improve prediction accuracy. Essentially, Prophet has translated the problem of forecasting(shouldn't this be "modelling"...figure out how it does parameter estimation) to a curve fitting exerxise 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 'forecast at scale', 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 and Supervised & Unsupervised machine learning methods (Random Forests, Clustering, ANN's).

2. Preliminary Literature Review

ARIMA models have been widely used to predict stock returns over the past 40 decades and are normally used as a bench mark against which other methods are compared (Zhang, 2003). Over the past 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 similar published studies based on the South African JSE Top40 index are publicly unavailable, studies using both stock returns and non-stock returns data have been used to compare 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 ANN using the Henrikson-Merton test. They found that the ANNs forecasts were considerably more accurate than the ARIMA forecasts (Kohzadi et el 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 comapre 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 modesl 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 used alternative out-of-sample methods to compare forecasts. However, in this paper, we use the Diebold-Mariano evaluation method which was itnended specifically for comparing forecasts (Diebold 2013). It is useful because it does not require the use of pseudo-out-of-sample environments to compare forecsts, which means that the results have more statistical 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 issues or papers I would like to consult that in relation to specific concerns I predict might arise as I execute the objectives of this paper.

1. Efficiency of markets in South Africa
2. Optimal model selection criteria for seasonal ARIMA model (AIC or BIC) given the objectives and scope of this paper.
3. If Prophet produces better forecsts than the ARIMA model, to consider comparing Prophet to other Machine Learning and Bayesian Time Series Forecasting methods such as 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.

4. Research Question

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

> "How does the latest prophet forecasting model compare with an ARIMA model through a Diebold-Mariano evaluation of the South African Top40 Index"

A secondary question that I hope the research could answer is whether intensive Bayesian modelling is worth the effort inthe context of forecastign stock returns of the South African Top40.

5. Methodology

The following is a high level outline of how I intend to answer my research question. Note that the latest version (as at May 2017) of R and R Stusio will ne used for model building and forecst comparison.

5.1. Data Collection and Preparation

6. Data Section from original template. Delete when complete

Discussion of data should be thorough with a table of statistics and ideally a figure.

In your tempalte folder, you will find a Data and a Code folder. In order to keep your data files neat, store all of them in your Data folder. Also, I strongly suggest keeping this Rmd file for writing and executing commands, not writing out long pieces of data-wrangling. In the example below, I simply create a ggplot template for scatter plot consistency. I suggest keeping all your data in a data folder.

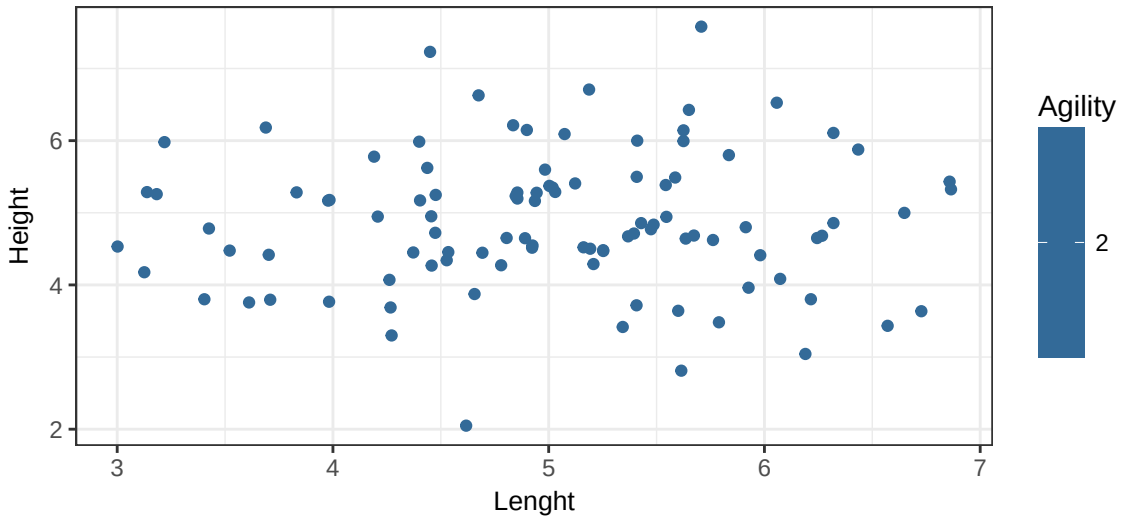

Figure 6.1: Caption Here

To reference the plot above, add a "\label" after the caption in the chunk heading, as done above. Then reference the plot as such: As can be seen, figure 6.1 is excellent. The nice thing now is that it correctly numbers all your figures (and sections or tables) and will update if it moves. The links are also dynamic.

I very strongly suggest using ggplot2 (ideally in combination with dplyr) using the ggtheme package to change the themes of your figures.

Also note the information that I have placed above the chunks in the code chunks for the figures. You can edit any of these easily - visit the Rmarkdown webpage for more information.

Here follows another figure from built-in ggplot2 data:

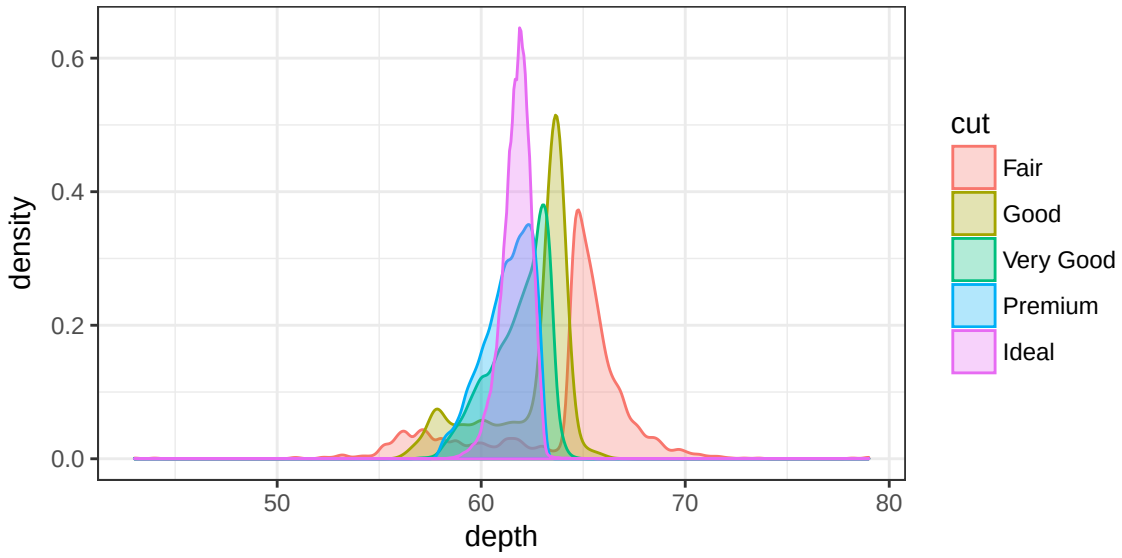

Figure 6.2: Diamond Cut Plot

7. Methodology

7.1. Subsection

Ideally do not overuse subsections. It equates to bad writing.[1]

7.2. Math section

Equations should be written as such:

$$\beta = \sum_{i=1}^{\infty} \frac{\alpha^2}{\sigma_{t-1}^2} \tag{7.1}$$
$$\int_{x=1}^{\infty} x_i = 1$$

[1]This is an example of a footnote by the way. Something that should also not be overused.

If you would like to see the equations as you type in Rmarkdown, use $ symbols instead (see this for yourself by adjusted the equation):

$$\beta = \sum_{i=1}^{\infty} \frac{\alpha^2}{\sigma_{t-1}^2} \int_{x=1}^{\infty} x_i = 1$$

Note again the reference to equation 7.1. Writing nice math requires practice. Note I used a forward slashes to make a space in the equations. I can also align equations using `\&''`, `and set to numbering only the first line. Now I will have to type`begin equation" which is a native LaTeXcommand. Here follows a more complicated equation:

$$\begin{aligned}
y_t &= c + B(L)y_{t-1} + e_t \\
e_t &= H_t^{1/2} z_t; \quad z_t \sim N(0, I_N) \quad \& \quad H_t = D_t R_t D_t \\
D_t^2 &= \sigma_{1,t}, \ldots, \sigma_{N,t} \\
\sigma_{i,t}^2 &= \gamma_i + \kappa_{i,t} v_{i,t-1}^2 + \eta_i \sigma_{i,t-1}^2, \quad \forall i \\
R_{t,i,j} &= diag(Q_{t,i,j}{}^{-1}).Q_{t,i,j}.diag(Q_{t,i,j}^{-1}) \\
Q_{t,i,j} &= (1 - \alpha - \beta)\bar{Q} + \alpha z_t z_t' + \beta Q_{t,i,j}
\end{aligned} \tag{7.2}$$

Noten that in 7.2 I have aligned the equations by the equal signs. I also want only one tag, and I create spaces using "quads".

See if you can figure out how to do complex math using the two examples provided in 7.1 and 7.2.

8. Results

Tables can be included as follows. Use the *xtable* (or kable) package for tables. Table placement = H implies Latex tries to place the table Here, and not on a new page (there are, however, very

many ways to skin this cat. Luckily there are many forums online!).

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.00	6.00	160.00	110.00	3.90	2.62	16.46	0.00	1.00	4.00	4.00
Mazda RX4 Wag	21.00	6.00	160.00	110.00	3.90	2.88	17.02	0.00	1.00	4.00	4.00
Datsun 710	22.80	4.00	108.00	93.00	3.85	2.32	18.61	1.00	1.00	4.00	1.00
Hornet 4 Drive	21.40	6.00	258.00	110.00	3.08	3.21	19.44	1.00	0.00	3.00	1.00
Hornet Sportabout	18.70	8.00	360.00	175.00	3.15	3.44	17.02	0.00	0.00	3.00	2.00

Table 8.1: Short Table Example

To reference calculations **in text**, *do this:* From table 8.1 we see the average value of mpg is 20.98.

Including tables that span across pages, use e.g.:

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.00	6.00	160.00	110.00	3.90	2.62	16.46	0.00	1.00	4.00	4.00
Mazda RX4 Wag	21.00	6.00	160.00	110.00	3.90	2.88	17.02	0.00	1.00	4.00	4.00
Datsun 710	22.80	4.00	108.00	93.00	3.85	2.32	18.61	1.00	1.00	4.00	1.00
Hornet 4 Drive	21.40	6.00	258.00	110.00	3.08	3.21	19.44	1.00	0.00	3.00	1.00
Hornet Sportabout	18.70	8.00	360.00	175.00	3.15	3.44	17.02	0.00	0.00	3.00	2.00
Valiant	18.10	6.00	225.00	105.00	2.76	3.46	20.22	1.00	0.00	3.00	1.00
Duster 360	14.30	8.00	360.00	245.00	3.21	3.57	15.84	0.00	0.00	3.00	4.00
Merc 240D	24.40	4.00	146.70	62.00	3.69	3.19	20.00	1.00	0.00	4.00	2.00
Merc 230	22.80	4.00	140.80	95.00	3.92	3.15	22.90	1.00	0.00	4.00	2.00
Merc 280	19.20	6.00	167.60	123.00	3.92	3.44	18.30	1.00	0.00	4.00	4.00
Merc 280C	17.80	6.00	167.60	123.00	3.92	3.44	18.90	1.00	0.00	4.00	4.00
Merc 450SE	16.40	8.00	275.80	180.00	3.07	4.07	17.40	0.00	0.00	3.00	3.00
Merc 450SL	17.30	8.00	275.80	180.00	3.07	3.73	17.60	0.00	0.00	3.00	3.00
Merc 450SLC	15.20	8.00	275.80	180.00	3.07	3.78	18.00	0.00	0.00	3.00	3.00

Cadillac Fleetwood	10.40	8.00	472.00	205.00	2.93	5.25	17.98	0.00	0.00	3.00	4.00
Lincoln Continental	10.40	8.00	460.00	215.00	3.00	5.42	17.82	0.00	0.00	3.00	4.00
Chrysler Imperial	14.70	8.00	440.00	230.00	3.23	5.34	17.42	0.00	0.00	3.00	4.00
Fiat 128	32.40	4.00	78.70	66.00	4.08	2.20	19.47	1.00	1.00	4.00	1.00
Honda Civic	30.40	4.00	75.70	52.00	4.93	1.61	18.52	1.00	1.00	4.00	2.00
Toyota Corolla	33.90	4.00	71.10	65.00	4.22	1.83	19.90	1.00	1.00	4.00	1.00
Toyota Corona	21.50	4.00	120.10	97.00	3.70	2.46	20.01	1.00	0.00	3.00	1.00
Dodge Challenger	15.50	8.00	318.00	150.00	2.76	3.52	16.87	0.00	0.00	3.00	2.00
AMC Javelin	15.20	8.00	304.00	150.00	3.15	3.44	17.30	0.00	0.00	3.00	2.00
Camaro Z28	13.30	8.00	350.00	245.00	3.73	3.84	15.41	0.00	0.00	3.00	4.00
Pontiac Firebird	19.20	8.00	400.00	175.00	3.08	3.85	17.05	0.00	0.00	3.00	2.00
Fiat X1-9	27.30	4.00	79.00	66.00	4.08	1.94	18.90	1.00	1.00	4.00	1.00
Porsche 914-2	26.00	4.00	120.30	91.00	4.43	2.14	16.70	0.00	1.00	5.00	2.00
Lotus Europa	30.40	4.00	95.10	113.00	3.77	1.51	16.90	1.00	1.00	5.00	2.00
Ford Pantera L	15.80	8.00	351.00	264.00	4.22	3.17	14.50	0.00	1.00	5.00	4.00
Ferrari Dino	19.70	6.00	145.00	175.00	3.62	2.77	15.50	0.00	1.00	5.00	6.00
Maserati Bora	15.00	8.00	301.00	335.00	3.54	3.57	14.60	0.00	1.00	5.00	8.00
Volvo 142E	21.40	4.00	121.00	109.00	4.11	2.78	18.60	1.00	1.00	4.00	2.00

Table 8.2: Long Table Example

9. Lists

To add lists, simply using the following notation

- This is really simple

  - Just note the spaces here - writing in R you have to sometimes be pedantic about spaces...

- Note that Rmarkdown notation removes the pain of defining LaTeXenvironments!

10. Conclusion

I hope you find this template useful. Remember, stackoverflow is your friend - use it to find answers to questions. Feel free to write me a mail if you have any questions regarding the use of this package feel free to email me. To cite this package, simply type citation("Texevier") in Rstudio to get the citation for Katzke (2017) (Note that uncited references in your bibtex file will not be included in References).

References

Katzke, N.F. 2017. *Texevier: Package to Create Elsevier Templates for Rmarkdown*. Stellenbosch, South Africa: Bureau for Economic Research.