Avocado Price Forecasting Comparative Study

Julisa Bana Abraham
Akhmad Arif Faisal
Universitas Gadjah Mada,
Departement of Electrical Engineering and Information Technology, Yogyakarta
julisa.bana.a@mail.ugm.ac.id
akhmad.arif.f.s@mail.ugm.ac.id

Abstract—**in the digital era, an era of fast growing information and knowledge, an organization must have an advantage in terms of information and knowledge in order to be successful in the industry. Time series forecasting is an information science branch that is integral to numerous decisive moments for an organization. For example, associations over all divisions of industry must take part in scope organization to proficiently allot rare assets and objective setting with the end goal to quantify execution in respect to a pattern. Delivering excellent forecasts isn't a simple issue for either machines or for analyst. In this paper we compare three methods for time series forecasting namely backpropagation multi layered perceptron (MLP), ARIMA, and Prophet. In our experiment Prophet shown a considerable capability of forecasting time series data and would be a good alternative method especially for forecasting time series data as it is easy to implement without any parameter required and yield a great result**

I. INTRODUCTION

In the digital era, an era of fast growing information and knowledge, an organization must have an advantage in terms of information and knowledge in order to be successful in the industry. Time series forecasting is an information science branch that is integral to numerous decisive moments for an organization. For example, associations over all divisions of industry must take part in scope organization to proficiently allot rare assets and objective setting with the end goal to quantify execution in respect to a pattern. Delivering excellent forecasts isn't a simple issue for either machines or for analyst [1], In this paper we compare three methods for time series forecasting namely backpropagation multi layered perceptron (MLP), ARIMA, and Prophet to find the more suitable method for forecasting avocado prices, the dataset was downloaded from the Hass Avocado Board website in May of 2018.

II. METHOD

This paper is focusing on comparing Prophet, the relatively new algorithm for time series forcasting, with two other more well known and estabilished method namely backpropagation MLP and ARIMA. So that we can know the performance of Prophet relative to the other two methods.

The time series dataset that we have been modeled is shown in Fig 1. Represents time series of an average prices of organic avocado throughout United States that will be modeled using three methods backpropagation MLP, ARIMA and Prophet evaluation will be through comparing testing data that is the last 30% of dataset with prediction generated from training or fitting process using the first 70% of dataset.

1. Dataset

Dataset that will be modelled was obtained from Kaggle, it contains historical data on avocado prices and sales volume in multiple US markets. In this paper we're not using all of the data contained in the dataset, but we're just using the time series data of organic avocado average prices to evaluate our model. Our time series dataset has length of 169 representing weekly average price of organic avocado that span from April 2014 until April 2018. The time series data that will be modeled shown in Fig 1

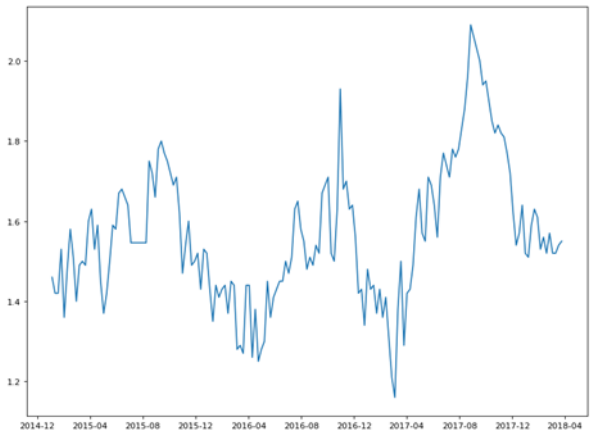

Fig 1. Organic Avocado Prices

2. Multi Layered Perceptron (MLP)

MLP configuration that have been used for this experiment has 3 layers: Input layer that has 1 node, hidden layer with 8 hidden nodes, and output layer that has 1 node and all nodes use rectified linear unit (ReLU) as their activation function. Our MLP structure is shown in Fig 2. The MLP configuration is our own design based on trial and error process that won't be discussed in this paper. Training scheme is based on t+1 prediction. This configuration is implemeted in Keras module in Python 3.5. Our MLP model uses Adam optimizer as the

gradient descent and relu as the activation function on every node.

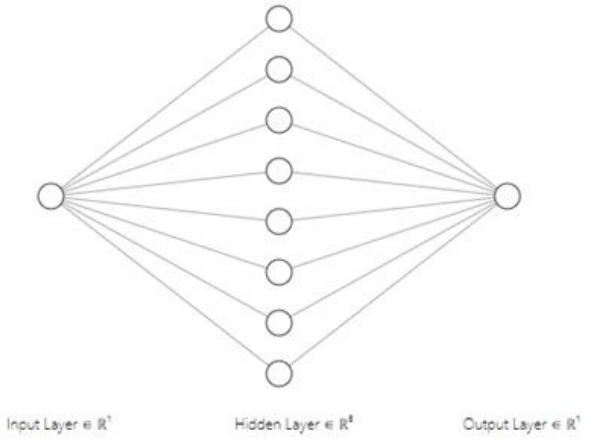

Fig 2. MLP configuration

3. ARIMA

ARIMA model introduced by Box and Jenkins in 1970 and it has been used widely for estimating, indetifying and, diagnosing time series data ever since, espescially in financial forecasting involving price time series [2 , 3, 4] and it have shown efficient capability to produce short-term forecasts and outperformed complex structural models in short-term prediction [5]

The ARIMA model is implemented using statsmodels.tsa module in Python 3.5 stats model have 3 parameters: p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Our model use 1,1,0 as p,d,q respectively based on positive correlation for the first lag and use 1 as d to make differencing is more stationary

4. Prophet

Prophet is a procedure developed by Facebook's data scientist Letham and Taylor in 2017 for forecasting time series data based on model where non linear trends are fit with yearly, weekly, and daily seasonality [1]. Prophet model use a decomposable time series model [6] with three main model components: trend, seasonality, and holidays. They are combined in the following equation:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon t \,. \quad (1)$$

Here g(t) is the trend function which models non-periodic changes in the value of the time series, s(t) represents periodic changes such as weekly and yearly seasonality, and h(t) represents the implications of holidays which can occur potentially on irregular schedules over one or more days. The error term εt represents any idiosyncratic changes which are not accommodated by the model. The implementation of Prophet modeled is using Prophet module installed on Python 3.5 using default parameters.

III. RESULT

In this section we will present the result of our experiment. We will use Mean Absolute Percentage Error (MAPE) as metrics to generalize the result obtained.

1. Multi Layered Perceptron (MLP)

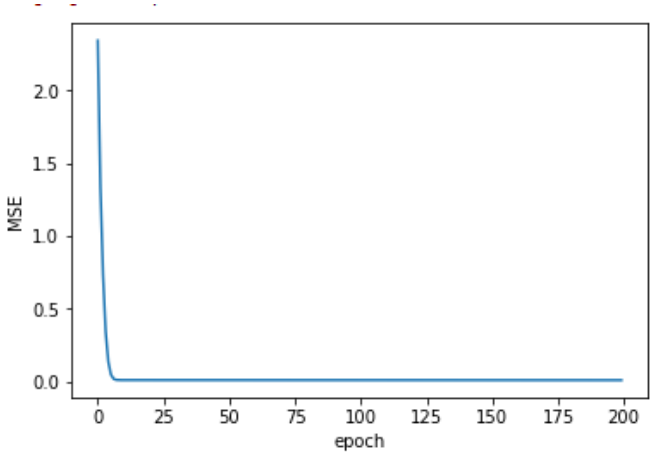

Fig 3. Epoch vs MSE

The training process of our backpropagation MLP model is shown in fig 3. Our model is trained with 200 epochs. The fitting result can be observed in fig 3 In testing process our MLP model has 10.14% Mean Absolute Percentage Error (MAPE)

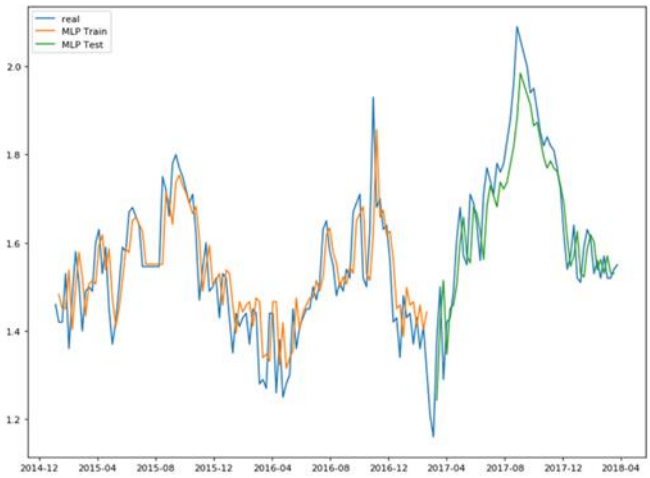

Fig 4. MLP training and testing result

2. ARIMA

The testing result of ARIMA(1,1,0) model is shown in Fig 5. This model has MAPE of 14.7%

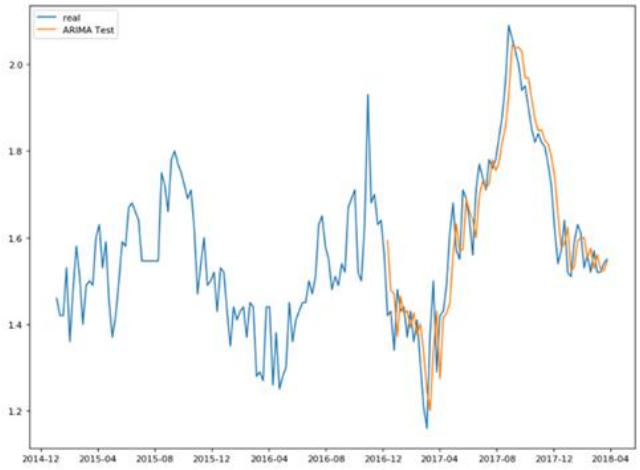

Fig 5. ARIMA testing result

3. Prophet

The result of Prophet model is shown in Fig 6 as the dots represents the actual data, blue lines is the predictions, and the blue shade is error estimations. Our Prophet implementation yields MAPE of 10.11%. This Prophet model generate two model components: trend or g(t) is shown in Fig 7 and yearly periodic changes s(t) shown in Fig 8.

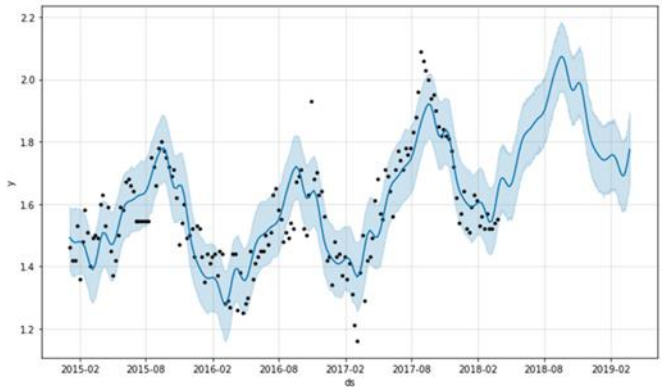

Fig 5. Prophet testing result

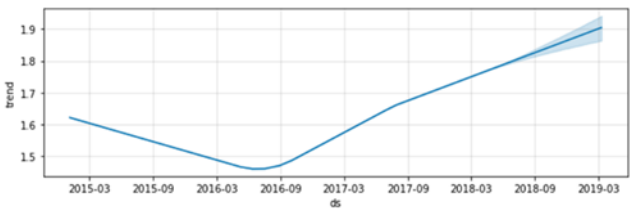

Fig 6. Trend component

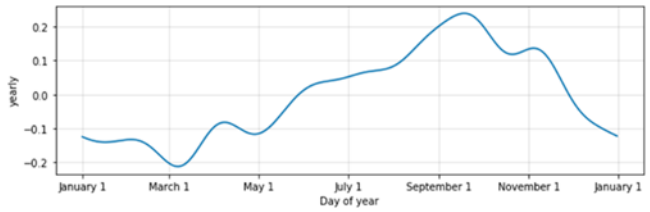

Fig 7. Yearly changes component

4. Forecasting Comparison

We can see the forecasting comparison of the three methods that we have been implemented in Fig 5. Interestingly enough our MLP model is might not have the capability to forecast the time series data as we can see our MLP model forecast is stuck at 1.45 and it is to be expected because the training scheme is just to predict t+1 time span. Prophet and our parameterized ARIMA have different forecast trend and we think both result is reasonable enough to be considered.

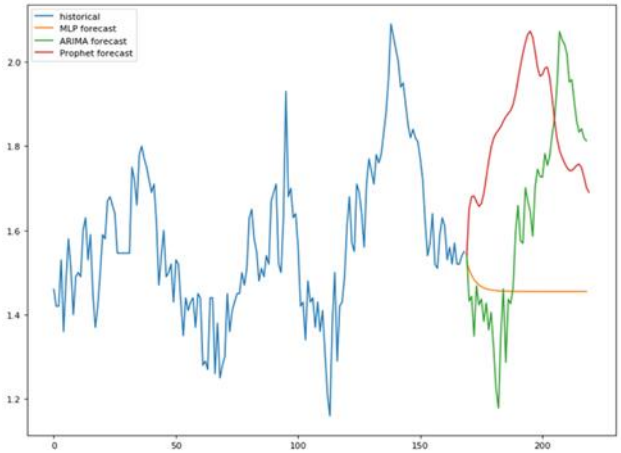

Fig 7. Forecasting comparison

IV. CONCLUSION

This paper presents the comparison between Prophet with two more well known methods for forecasting namely backpropagation MLP and ARIMA, both with our own configurations. In our experiment Prophet shown a considerable capability of forecasting time series data and would be a good alternative method especially for forecasting time series data as it is easy to implement without any parameter required and yield a great result.

REFERENCES

[1] Taylor SJ, Letham B. 2017. Forecasting at scale. *PeerJ Preprints* 5:e3190v2

[2] P. Pai and C. Lin, "A hybrid ARIMA and support vector machines model in stock price prediction", Omega vol.33 pp. 497-505, 2005.

[3] Rangan and N. Titida, "ARIMA Model for Forecasting Oil Palm Price", Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications, Universiti Sains Malaysia, 2006.

[4] Merh, V.P. Saxena, and K.R. Pardasani, "A Comparison Between Hybrid Approaches of ANN and ARIMA For Indian Stock Trend Forecasting", Journal of Business Intelligence, vol. 3, no.2, pp. 23-43,2010

[5] A. Meyler, G. Kenny and T. Quinn, "Forecasting Irish Inflation using ARIMA Models", Central Bank of Ireland Research Department, Technical Paper, 3/RT/1998.

[6] A. & Peters, S. (1990), 'Estimation procedures for structural time series models'Journal of Forecasting 9, 89–10