Bitcoin Forecasting Using ARIMA and PROPHET

Işıl Yenidoğan, Aykut Çayır
MIS Department
Kadir Has University, Istanbul, Turkey
{isil.yenidogan, aykut.cayir}@khas.edu.tr

Ozan Kozan, Tuğçe Dağ, Çiğdem Arslan
MIS Department
Kadir Has University, Istanbul, Turkey
{sukruozan.kozan, tugce.dag,
cigdem.arslan}@stu.khas.edu.tr

Abstract— **This paper presents all studies, methodology, and results about Bitcoin forecasting with PROPHET and ARIMA methods using R analytics platform. To find the most accurate forecast model, the performance metrics of PROPHET and ARIMA methods are compared on the same dataset.**

The dataset selected for this study starts from May 2016 and ends in March 2018, which is the interval that Bitcoin values changing significantly against the other currencies. Data is prepared for time series analysis by performing data preprocessing steps such as time stamp conversion and feature selection. Although the time series analysis has a univariate characteristics, it is aimed to include some additional variables to each model to improve the forecasting accuracy. Those additional variables are selected based on different correlation studies between cryptocurrencies and real currencies.

The model selection for both ARIMA and PROPHET is done by using threefold splitting technique considering the time series characteristics of the dataset. The threefold splitting technique gave the optimum ratios for training, validation, and test sets. Finally two different models are created and compared in terms of performance metrics. Based on the extensive testing we see that PROPHET outperforms ARIMA by 0.94 to 0.68 in R^2 values.

Keywords: bitcoin, forcasting

I. INTRODUCTION

Bitcoin [1] is the first and the most popular cryptocurrency in the world. Cryptocurrency is a digital currency form that uses cryptography for security and verifies transactions on the network .

Since cryptocurrency, especially Bitcoin, transaction volumes have become really high and remarkable for the last two years, the accurate forecasting of future values of these are getting more important and interesting area for both researchers and investors.

This study is representing forecasting Bitcoin time series data and comparing forecasting results of two different models: PROPHET [2] and ARIMA [3]. Our interest in this is to find out which model is more successful for Bitcoin forecasting.

This paper is organized as follows. The dataset, preprocessing, and feature selection steps are described in section II. Section III briefly explains the ARIMA and PROPHET forecast models. Experiments and results are presented in section IV, whereas the conclusions are given in section V.

II. DATA PREPROCESSING AND FEATURE SELECTION

A. Dataset

Bitcoin Historical Data set between 2012 and 2018, obtained from the Kaggle website in minute format, is used for the project. The Bitcoin Historical Data set consists of 52685 rows and 8 columns as shown in Fig. 1. The features of this data set are: Timestamp, Open, High, Low, Closed, Volume (BTC), Volume (Currency), and Weighted Price values. The Open, High, Low, and Closed columns in the dataset show the opening, the highest, the lowest, and the closing prices of Bitcoin against US $ in that minute. The Volume (BTC) expresses the transaction volume of the Bitcoin transferred to the pulsed hourly. The Volume (Currency) refers to the trading volume in hours and the Weighted Price is the average of the opening, highest, lowest, and closing prices of Bitcoin.

	Timestamp	Open	High	Low	Close	Volume_BTC.	Volume_Currency.	Weighted_Price
1	1325317920	4.39	4.39	4.39	4.39	0.4555809	2	4.39
2	1325317980	4.39	4.39	4.39	4.39	0.4555809	2	4.39
3	1325318040	4.39	4.39	4.39	4.39	0.4555809	2	4.39
4	1325318100	4.39	4.39	4.39	4.39	0.4555809	2	4.39
5	1325318160	4.39	4.39	4.39	4.39	0.4555809	2	4.39
6	1325318220	4.39	4.39	4.39	4.39	0.4555809	2	4.39
7	1325318280	4.39	4.39	4.39	4.39	0.4555809	2	4.39
8	1325318340	4.39	4.39	4.39	4.39	0.4555809	2	4.39
9	1325318400	4.39	4.39	4.39	4.39	0.4555809	2	4.39
10	1325318460	4.39	4.39	4.39	4.39	0.4555809	2	4.39
11	1325318520	4.39	4.39	4.39	4.39	0.4555809	2	4.39

Fig. 1. An overview of original Bitcoin Dataset.

B. Preprocessing

- Preprocessing is divided into two parts: time stamp conversion-integration and removing samples before 2016. First the time stamp is converted into human readable POSIXct year-month-day format.Since the data were in minute format, the overall change was not well understood. The data set was converted to hourly form to reduce transaction complexity.
- The correlation matrix is formed to understand the relations among the features mutually. Since the Volume_BTC and the Volume_Currency features have a weak correlation they are removed from the dataset.
- According to the raw data graph given in Fig. 2, the data before 2016 is removed because the Bitcoin values between these dates were quite stable.

978-1-5386-7893-0/18/$31.00 ©2018 IEEE

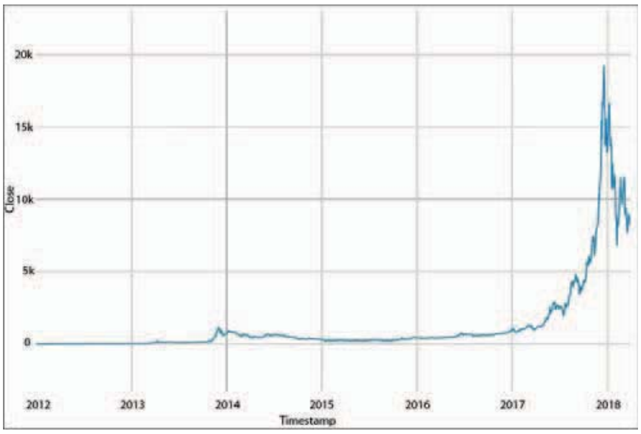


Fig. 2. Bitcoin raw dataset graph.

C. Splitting Data into Training, Validation, and Test Sets

In machine learning studies the data is usually partitioned into training and testing randomly with a ratio of 70 to 30 respectively. However, time series data can not be divided into different parts randomly as a cause of its nature. For a better predictive models to avoid over-fitting, it is best to use the training, testing, and validation sets. In his study, we did not use the general splitting ratios. The optimal ratio for splitting data is determined by observing the changes on R^2, MAPE (Mean Absolute Percentage Error), and RMSPE (Root Mean Square Percentage Error) values for different percentages of splitting.That is, the optimal splitting percentages are obtained by conducting experiments as shown in Fig. 3. Fig. 3 emphasizes that three fitness metrics are getting worse while increasing the split ratio for training set size. Thus, the best splitting for training, validation and test sets are 70, 15 and 15 percent respectively.

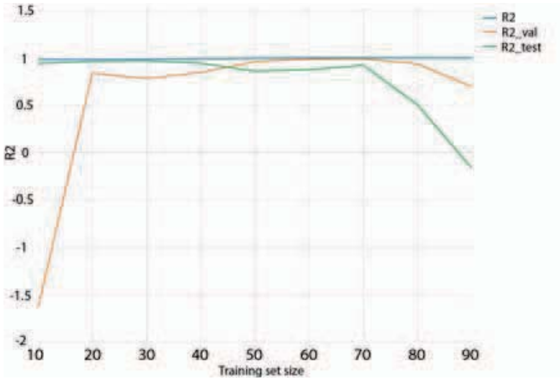
Since we agree with Alex Porcel in his statement: ”When you fit a time series model, you are interested in predicting the future, what is yet to happen. If you test your algorithm with the latest data you will be more confident that the resulting predictions are more accurate” [4], we used the most recent data for test. Because the other currency data which will be used as extra regressors was not available freely before that date the last date intervals for three different splits are chosen as follows :

- Train: 2016/05/03 - 2017/08/30
- Validation: 2017/08/31 - 2017/12/12
- Test: 2017/12/13 - 2018/03/27

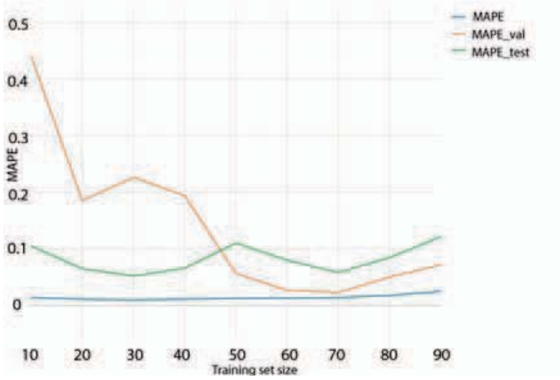
D. Feature Selection

We have figured out that we could not use the features in the dataset as extra regressors to fit our prediction models. Since the opening value of a day is the same as the closing value of the previous day, the model would give meaningless predictions. Niki Shirvani, the former student of MIS Department, told us that she had discovered a relation between Japanese Yen (JPY) and Bitcoin during the study for her graduation project. She recommended us to examine the other currencies in a group meeting. We made another research based on this idea and the correlation matrix between other currencies and Bitcoin is obtained as in Fig. 4.

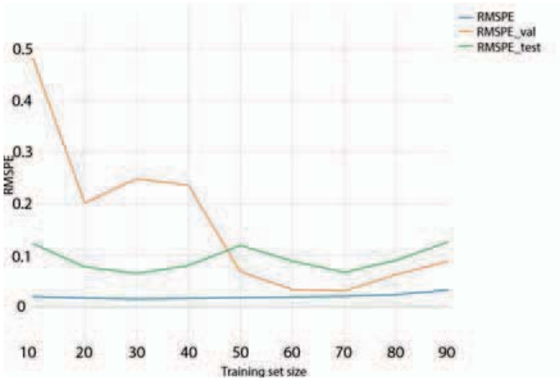
According to the correlation matrix, there is a relation between Bitcoin and GBP, EUR, and JPY. Then ”Price” feature is selected as an extra regressor. The values of those currencies through the time slice of the study are collected and integrated with our original data based on time-series structure.



(a) R^2 values for training, validation and test sets.



(b) MAPE values for training, validation and test set.



(c) RMSPE values for training, validation and test set.

Fig. 3. Finding the best training, validation and test split ratios in terms of regression performance metrics.

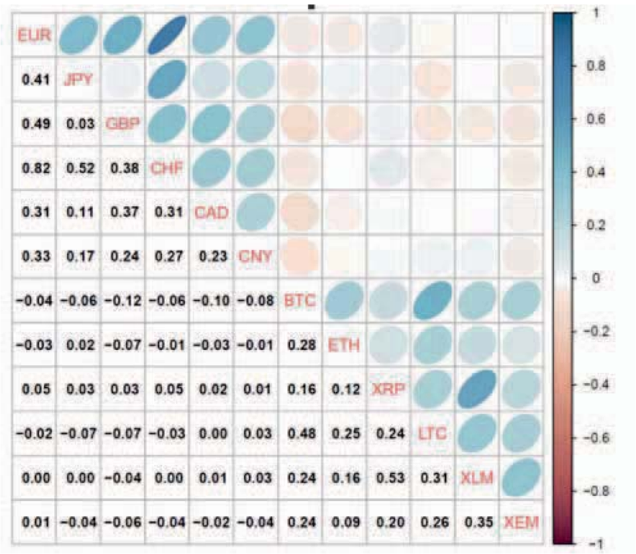


Fig. 4. Correlation matrix of Bitcoin and currencies [5]

.

III. Models

In this section, ARIMA and PROPHET forecast models are explained briefly.

A. ARIMA Forecast Model

ARIMA stands for Auto Regressive Integrated Moving Average. It is a well known forecasting model in financial and data science time series applications [6]. ARIMA has two different types of model such as seasonal and non-seasonal. Nonseasonal ARIMA means removing seasonal effects like holidays and term effects from dataset and using pure trend lines. Unlikely, seasonal ARIMA uses holidays and other term effects directly on the dataset. There is no specific smoothing or removing term effects from the dataset.

Seasonal and nonseasonal ARIMA models have different parameters. Seasonal ARIMA model can be described with three parameters as the following:

- $P =$ number of seasonal autoregressive terms
- $D =$ number of seasonal differences
- $Q =$ number of seasonal moving-average terms

Nonseasonal ARIMA model can also be described with three parameters following:

- $p =$ number of autoregressive terms
- $d =$ number of nonseasonal differences
- $q =$ number of moving-average terms

In this study, we use nonseasonal ARIMA model for forecasting on Bitcoin dataset. The hardest part of training of the model is to find nonseasonal ARIMA parameters p, d and q. In R statistical programming language there is a framework called *forecast* to infer these parameters automatically. The function used in this framework is called *auto.arima* [7]. We have used *auto.arima*() to dedect these parameters automatically.

B. PROPHET Forecast Model

PROPHET is an open source software that is available in Python and R for forecasting time series data. PROPHET is published by Facebook's Core Data Science team. It depends on a contribution model where non-linear trends are fit with weekly and yearly seasonality and plus holidays. PROPHET is strong to missing data, capturing the shifts in the trend and large outliers. In addition, it gets a reasonable estimate of the mixed data without spending manual effort [2].

Purely automatic prediction techniques are not flexible to combine useful assumptions because they are fragile. Furthermore, high quality estimates are not easy to make, requiring special data science skills. All these are determined as working motivation for PROPHET because it wants to make the high-quality predictions easier.

PROPHET is optimized for business forecast that are observed on Facebook. For example, time, daily, weekly observations of history, within a year, large outliers, trend changes, missing observation and trends that are non-linear growth curves [8].

PROPHET framework has its own special data frame to handle time series and seasonality easily. The data frame needs two basic columns. One of these columns is ”ds” and this column stores date time series.The other column is ”y” and it stores the corresponding values of the time series in the data frame. Thus, the framework can work on seasonal time series quiet well and it provides some options to handle seasonality of the dataset. These options are yearly, weekly and daily seasonality. Due to providing these options, a data analyst can choose the available time granularity for the forecast model on the dataset [8].

IV. Experiments and Results

This paper presents two different forecast models on Bitcoin dataset. Thus, our experiment has been performed on the dataset for ARIMA and PROPHET models. All scripts have been written in R statistical programming language [9].

As mentioned before, training set is 70% of the all dataset and validation and test sets are 15% of the all dataset using the threefold splitting method for the time series as found in Fig. 3. To measure performance of models and to find the best splitting ratios, we have used different metrics for all models. These metrics can be listed as below:

- $R^2 = \frac{\sum(\hat{y_i}-\bar{y})^2}{\sum(y_i-\bar{y})^2}$, where $\hat{y_i}$ is predicted value and $\bar{y}$ is the mean of the real values. R^2 measures goodness of fitting.
- $MSE = \frac{1}{N}\sum(y_i - \hat{y_i})^2$, where N is the number of samples, $\hat{y_i}$ is predicted value and y_i is the real value.
- $MAE = \frac{1}{N}\sum|y_i - \hat{y_i}|$, where N is the number of samples, $\hat{y_i}$ is predicted value and y_i is the real value.

RMSE, RMSPE and, MAPE are derived from MSE and MAE.

Forecast has been performed using Bitcoin's relation to JPY, Euro and, Pound. Although the time series analysis has a univariate characteristics, it is aimed to include some additional variables to each model to improve the forecasting

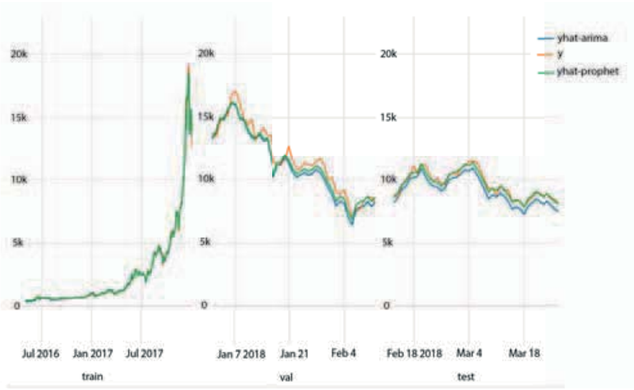


Fig. 5. Visualization of Fitting for Both Models

accuracy. These variables were JPY, EUR, and GBP based on the correlation matrix in Figure 4. The problem was that neither of the additional variables had future values to use as extra regressors. We made several individual forecasting models to fill in the future values for each currency (JPY, EUR, and GBP). After that we use these predicted values for the two main models (Arima and Prophet) which forecast the future value of Bitcoin. The models predict values that could be considered successful in the 90-day forecast that they had never seen their real value before.

TABLE I: R^2 AND ERROR VALUES FOR PROPHET FORECAST MODEL

Set	R^2	MSE	RMSE	RMSPE	MAE	MAPE
Train	0.997	35605	188.69	0.036	71.384	0.026
Val	0.941	425344	652.18	0.054	520.503	0.044
Test	0.945	60069	245.09	0.024	196.652	0.020

TABLE II: R^2 AND ERROR VALUES FOR ARIMA FORECAST MODEL

Set	R^2	MSE	RMSE	RMSPE	MAE	MAPE
Train	0.997	38498	196.21	0.057	85.254	0.043
Val	0.907	667508	817.01	0.071	678.807	0.059
Test	0.681	352600	593.80	0.062	541.946	0.056

Table I and II show PROPHET and ARIMA fit on training and validation sets very well. However, ARIMA is worse than PROPHET on the test set. Figure 5 visualizes goodness of fitting for both models and this plot supports the results are reported on table I and II.

V. CONCLUSIONS

The objective of this study was to compare Bitcoin's future values using the PROPHET and ARIMA methods. To this end, the bitcoin data set has been examined throughly. Taking into account the correlation values with other currencies (GBP, EUR, JPY), are also investigated for a possible regressor. Regression was performed using Bitcoin's relation to Dollar, Euro and Pound. Models were trained on the data belonging to May 2016 and January 2018 time interval. The models predicted values that could be considered successful in the 90-day forecast. In this respect, the ARIMA and PROPHET packages in R were were found to be quite important factors for the job at hand. Having obtained the Japanese Yen as a regressor we then used all the data with these two packages, the success of the model in the time series prediction was investigated and compared.

While the PROPHET model makes predictions quite close to reality, that is up to 94.5% precision, the ARIMA model provided only 68% precision.

REFERENCES

[1] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.
[2] Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, (just-accepted), 2017.
[3] Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, 18(3):1014–1020, 2003.
[4] Alex Porcel. What % is the best train/test split for Time Series Data? . Available at `https://stats.stackexchange.com/questions/328176/what-is-the-best-train-test-split-for-time-series-data`, accessed Feb. 2018.
[5] Correlation of Bitcoin and Other Currencies. `https://data.bitcoinity.org/`, 2018. [Online; accessed April 2018].
[6] G Peter Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175, 2003.
[7] Rob J Hyndman and Yeasmin Khandakar. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3):1–22, 2008.
[8] Prophet Forecasting At Scale. `https://research.fb.com/prophet-forecasting-at-scale/`, 2018. [Online; accessed April 2018].
[9] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.