Comparison of Prediction Algorithms in Time Series

Nurettin C. Dedetas
Gebze Technical University
Department of Computer Science
Student Number : 171044028
Email: nurettin.dedetas2017@gtu.edu.tr

January 24, 2021

Abstract

In this research I compared 3 time series machine learning algorithms. The algorithms are as follows, FB Prophet, Auto Regression and Exponential Moving Average. The dataset used in this project is daily Bitcoin(BTC) values from 2013 to current day. The goal of the project is to compare which Machine learning algorithm is the best suited for unstable/volatile time series like stocks or cryptocurrencies.

1 Dataset

The Dataset used in this experiment is the real life Bitcoin Price Index from 10/1/2013 to 1/23/2021. The reason this dataset has been chosen is that it is volitile , thus very hard to predict accurately.

2 Algorithms

2.1 FB Prohpet

FB Prohpet proposes a modular regression model with interpretable parameters that can be intuitively adjusted by analysts with domain knowledge about the time series. It describes performance analyses to compare and evaluate forecasting procedures, and automatically flag forecasts for manual review and adjustment. Tools that help analysts to use their expertise most effectively enable reliable, practical forecasting of business time series.
FB Prohpet uses a decomposable time series model (Harvey Peters 1990) with three main model components: trend, seasonality, and holidays. They are combined in the following equation:

$y(t) = g(t) + s(t) + h(t) + t.$
Here g(t) is the trend function which models non-periodic changes in the value of the time series, s(t) represents periodic changes (e.g., weekly and yearly seasonality), and h(t) represents the effects of holidays which occur on potentially irregular schedules over one or more days. The error term t represents any idiosyncratic changes which are not accommodated by the model; later we will make the parametric assumption that t is normally distributed.

2.2 Auto Regression

An autoregressive (AR) model predicts future behavior based on past behavior. It's used for forecasting when there is some correlation between values in a time series and the values that precede and succeed them. The process is basically a linear regression of the data in the current series against one or more past values in the same series.
The AR(p) model is defined by the equation:
$yt = +1yt - 1 + 2yt - 2 + \ldots + pyt - 1 + At$

Where:

- yt-1, yt-2...yt-p are the past series values (lags),
- At is white noise (i.e. randomness),
- is defined by the following equation

$$\delta = \left(1 - \sum_{i=1}^{p} \phi_i\right) \mu ,$$

2.3 Exponential Moving Average (EMA)

The exponential moving average (EMA) is a technical chart indicator that tracks the price of an investment (like a stock or commodity) over time. The EMA is a type of weighted moving average (WMA) that gives more weighting or importance to recent price data. Like the simple moving average, the exponential moving average is used to see price trends over time, and watching several EMAs at the same time is easy to do with moving average ribbons.

The exponential moving average is designed to improve on the idea of a simple moving average (SMA) by giving more weight to the most recent price data, which is considered to be more relevant than older data. Since new data carries greater weight, the EMA responds more quickly to price changes than the SMA.

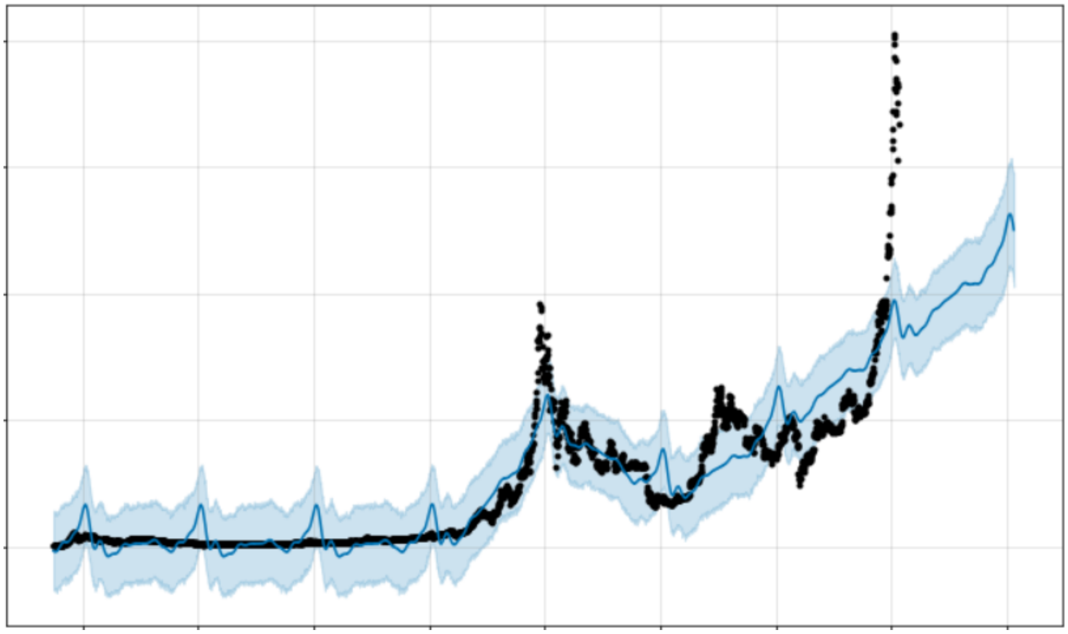
$EMA = \text{Price}(t) \times k + EMA(y) \times (1 - k)$
where:
$t = \text{today}$
$y = \text{yesterday}$
$N = \text{number of days in EMA}$
$k = 2 \div (N + 1)$

I have further modified this algorithm to make predictions.The modifications are as follows. For a given $predictionperiod = x$, the results of the last x EMA calculations are added to the existing dataset, then the EMA for the extended dataset is calculated for the last x entry. Thus predicting the period from the previous data.

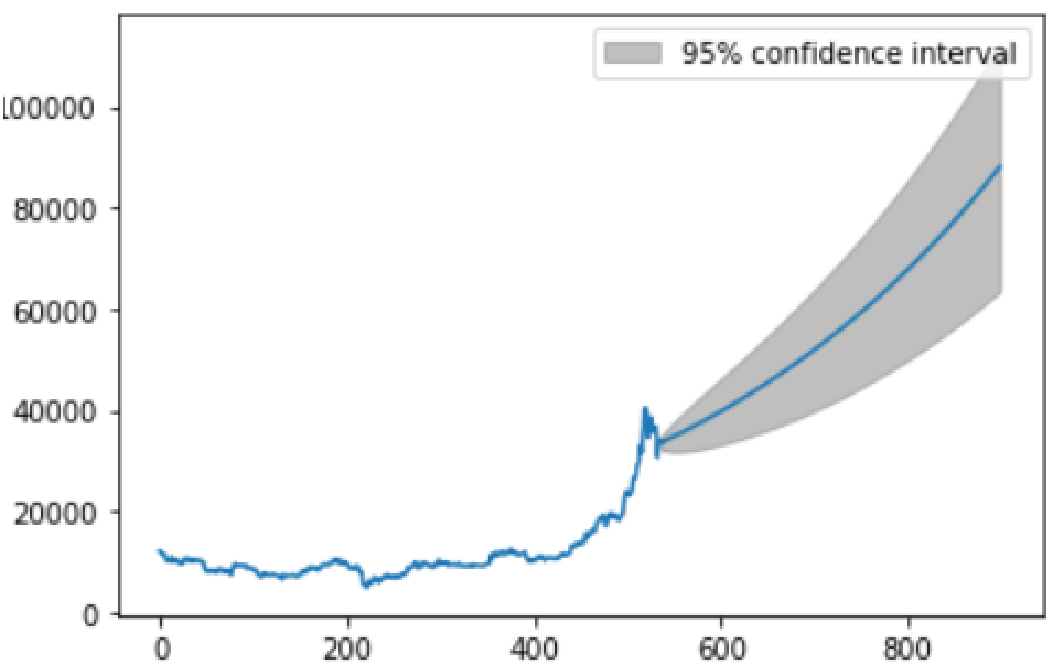
This brings inherent problems.First of all, the datasets last x entries are now assumed values, thus if there was an error with the values, the error will grow exponentially. Second of all, this approach lacks seasonality unlike the last two approaches.

3 Results

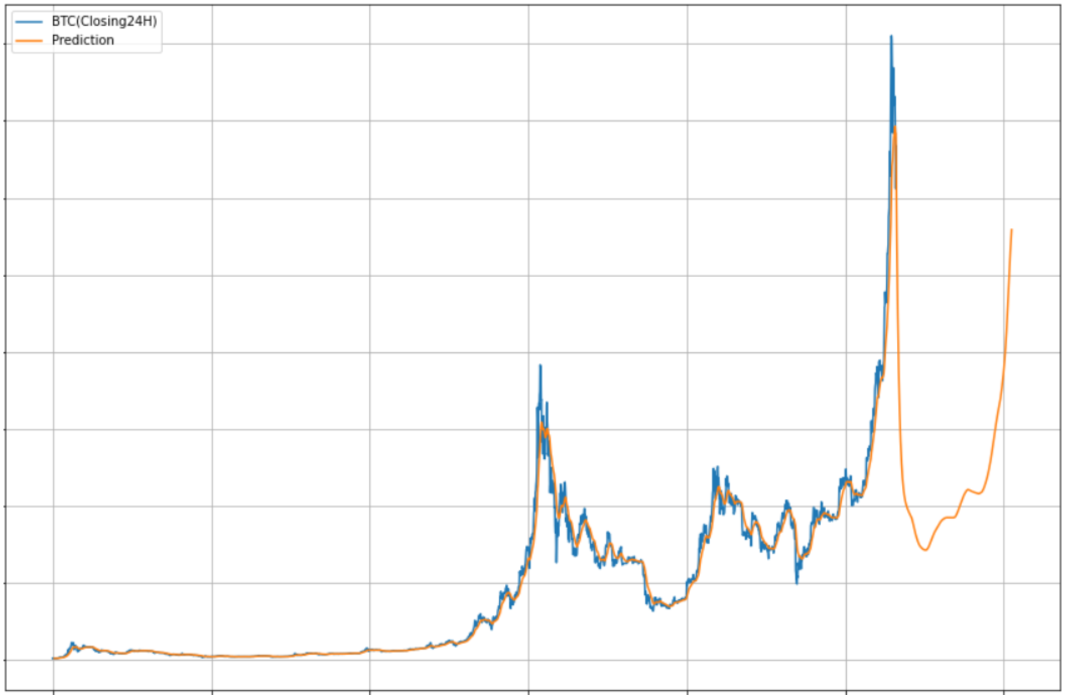
3.1 FBProphet



3.2 Auto Regression



3.3 EMA



4 Summary

Compared to FBP and AR, the EMA puts more emphasis on the more current data and thus misses the overall seasonality and fluctuates too much. The FBP and AR are pretty close to each other but Auto regression model has a pretty linear growth unlike FBP which still has some volatility.