Tesla Stock Price Prediction Using ARIMA-based Model

Danning Lao, Zhanpeng Zhou, Zhengfei He

UM-SJTU Joint Institute

Abstract

Based on the fetched Tesla's stock price and potentially related economic data from trustful databases, we aimed to predict Tesla's future stock price from 7^{th}*Dec* to 11^{th}*Dec* using explainable regression models implemented by `python`. Untill now, We compared different modelling result and developed *Seasonal ARIMA-GARCH* based on:

1. Addictive sub-series assumption.
2. *Autoregressive Integrated Moving Average model* (ARIMA)
3. *Generalized Autoregressive Conditional Heteroskasticity model* (GARCH)

This model explains a linear relationship between future and past data and partially addresses time dependent conditional variance.

Introduction

Tesla is an American electric vehicle and clean energy company which is founded in 2003 but achieve increasingly success and popularity these days. It's meaningful and challenging to forecast Tesla stock precisely and reliably (historical stock price shown in Fig 1).

Approximate entropy (ApEn) is a technique used to quantify regularity and unpredictability over time-series data, in this case, the ApEn of Tesla Stock = 0.569, indicating some regularity and predictability.

Figure 1:Candlestick Grpah of Tesla's Stock Price from Dec 2019 to Dec 2020

Mathematical Fomulations of Seasonal ARIMA-GARCH Model

First, we decomposed the original actual time seriesy_t into sub-series of trend T_t, seasonalityS_t and residualR_t, assuming an additive relationship

$$y_t = T_t + S_t + R_t$$

We preprocessed the trend and residual series to be stationary series $T(T_t)$ and $T(R_t)$, then we apply ARIMA-GARCH model to the transformed series. Where the relationship can be expressed as:

ARIMA: [1]

$$W_t = E[T(T_t)] = \sum_{i=1}^{p} \Phi_i W_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

GARCH: [2]

$$\epsilon_t = \sigma_t Z_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{n} \beta_j \sigma_{t-j}^2$$

(*) p, q, m, n will be selected by grid search using criteria of BIC.

Here we have two assumptions:

$$Z_t \sim N(0,1)$$

$$\epsilon_t \ stationary$$

And we repeat the process with $P_t = E[T(R_t)]$, and obtain the final model:

$$Y_t = T^{-1}(W_t) + T^{-1}(P_t) + S_t$$

In the modelling process, ARIMA explain a linear relationship between past and future data, while GARCH address volatility clustering problems, reducing heteroskedasticity.

Design FlowChart

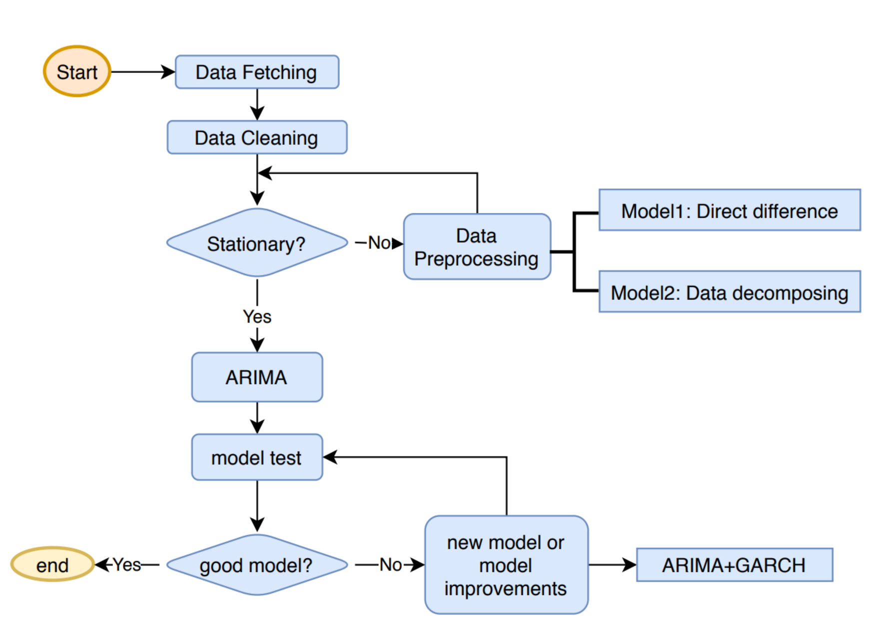

Figure 2:Flowchart of Modelling Process

Experimental Results

Setup EDA: After seasonal decomposition, Tesla stock data from 2020/9 to 2020/11 show a generally increasing trend and a clear seasonality.

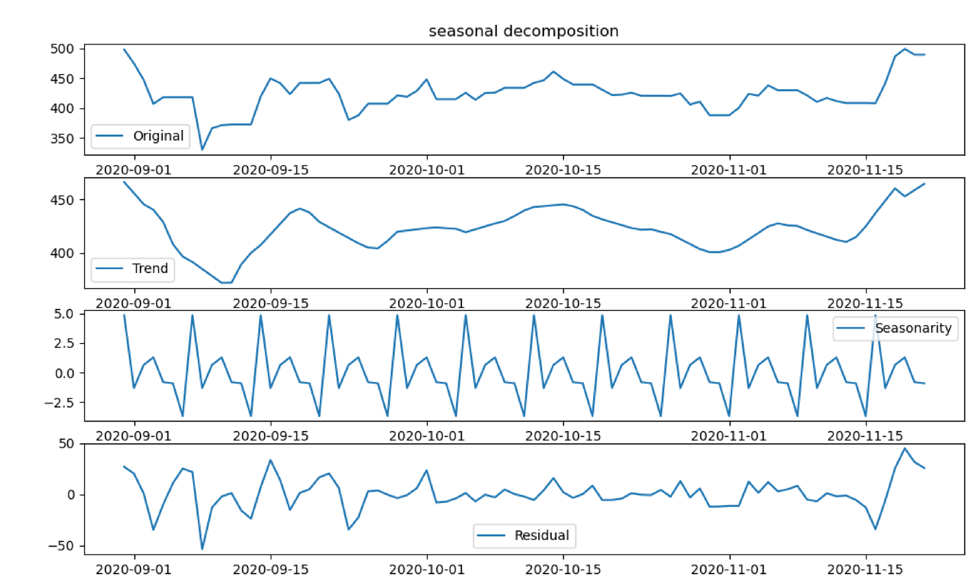

Figure 3:data decomposition

Fit-Result: We perform the modeling of different kinds of ARIMA model.

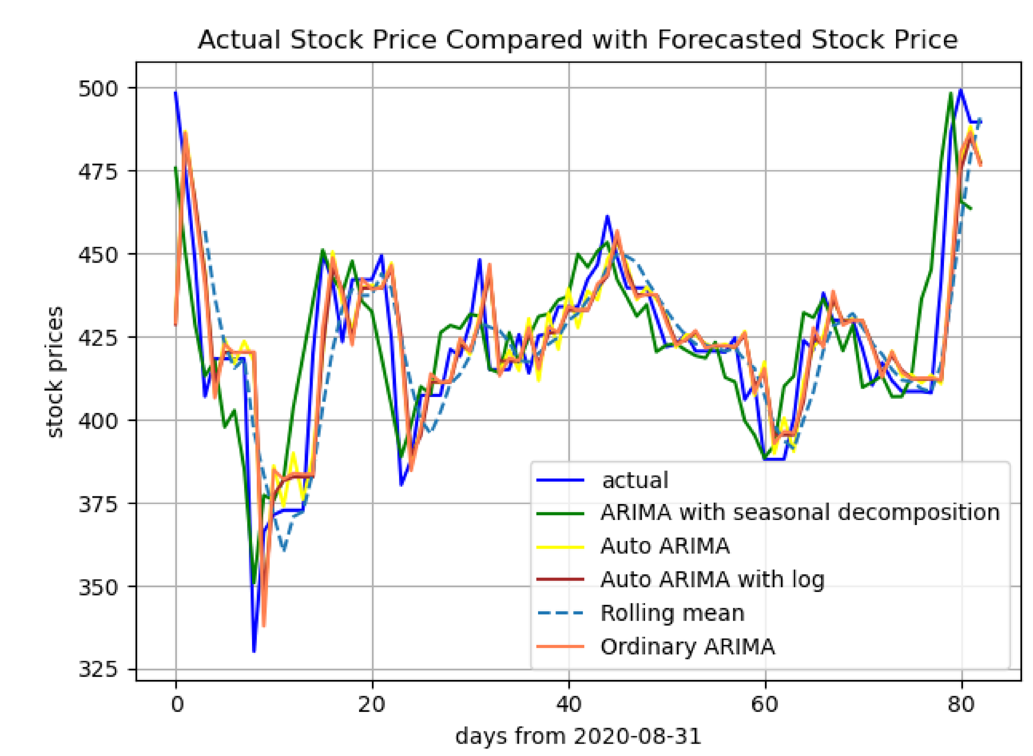

Figure 4:ARIMA fit result

Garch model(2,0) is selected by using both AIC and BIC metrics.

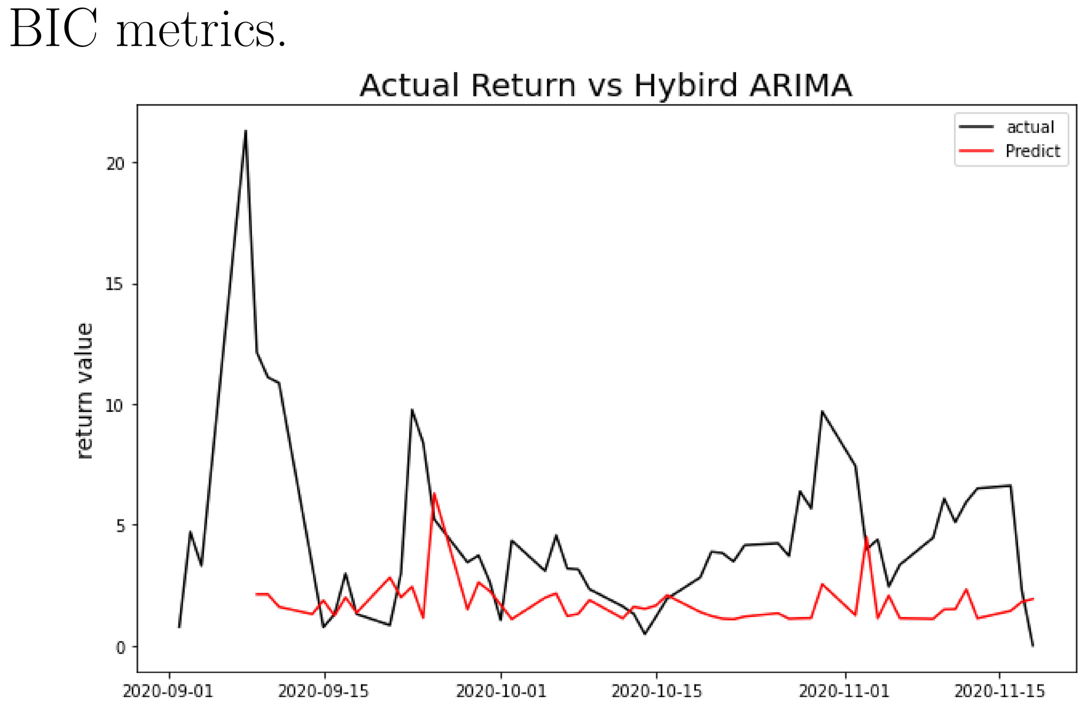

Figure 5:GARCH fit result

Experimental Result

The performance of GARCH fitting Arima model's residuals is not satisfactory.This is one of our model's limitations since it can only explain partial time dependent conditional variance.

Conclusion

After a comparison study among different models, based on *RMSE* result, we choose to use *Seasonal ARIMA-GARCH* model which gives the smallest RMSE = 97.3, indicating our model is valid and good for prediction.

Future Work

There are two aspects we want to focus on later, the first is linearity assumption which may not always suit the realistic case[3].

Another issue is, simply using close stock price to predict is not so meaningful as using multiple related regressors. We collect our data, including Dow Jones Index Average, Brent Crude Oil Price, Revenue and then applies PCA to reduce dimension.Later, more work will be done to fit a more explainable model with these data.

References

[1] G Peter Zhang.
Time series forecasting using a hybrid arima and neural network model.
Neurocomputing, 50:159–175, 2003.

[2] Luc Bauwens, Sébastien Laurent, and Jeroen VK Rombouts.
Multivariate garch models: a survey.
Journal of applied econometrics, 21(1):79–109, 2006.

[3] S. Ertekin U. C. Buyuksahin.
Improving forecasting accuracy of time series data using a new arima-ann hybrid method and empirical mode decomposition.
Neurocomputing, (361):151–163, July 2019.

JOINT INSTITUTE
交大密西根学院