Chronos Trade: Multimodal Stock exchange report

Bibaswan Biswas

December 2025

Abstract

In this project, we cover ML concepts like ARIMA,LSTM etc to make time series forecast like stock price prediction.

Contents

1 Week-1:ARIMA Model

In the first week, we studied about the ARIMA Model which is a statistical model used to make time series prediction that accounts:

- Autoregression(AR)
- Differencing(I)
- Moving Average(MA)

1.1 Autoregression

Autoregression refers to something like a moving best fit, that is, it is the fitting of data based on previous data points at past time. The model is therefore of the form:

$$x_t = \sum_{j=1}^{p} \phi_j x_{t-j} + w_t$$

where x_i is the observation made at time i. The ϕ_j are the autoregressive parameters and w_t is a noise term for time t. The value p is called the order of the AR.
Thus, AR involves finding the relation among the previous data points to fit the new data point.

1.2 Differencing

Stationarity: For ARIMA, we need to ensure the data does not have too much local variations. Hence, we need the data to be stationary, that is the mean and variance should be close to zero and the covariance between two data points should not depend on the individual time of the observation even though it may depend on the difference between the times.
The need for this is basically that the ARIMA is linear in no of past observations considered and hence it certainly cannot model data with high fluctuations.
Hence, in order to correct the data for stationarity, we use the technique of **differencing**, which is basically taking the difference between consecutive data points, i.e.,

$$x'_t = x_t - x_{t-1}$$

Thus, it is the change in the observed values over time. The number of times we need to use differencing to make the data stationary is denoted by d and is called the order of differencing.

1.3 Moving average

The Moving average(MA) part of ARIMA describes the dependency of the current observation on the previous forcast errors. The model is of the form:

$$x_t = \sum_{j=1}^{q} \phi_j w_{t-j} + w_t$$

where w_i is the noise in the observation at time i. The number of terms in this model, denoted by q is called the order of the MA.

1.4 ACF/PACF

ACF is the autocorrelating function,which is mathematically given by

$$\rho(k) = \frac{Cov(x_t, x_{t-k})}{Var(x_t)}$$

which is basically a measure of how much the time series resembles itself after k time steps. The PACF, is just the partial ACF, which is the ACF due to a lag,say k, with the effects of all shorter lags removed.

1.5 About the code

In the code, I used the weekly closing stock price of Tesla over 4 years for the ARIMA analysis. In the code, f is the fraction used as training data, currently set at 0.8 . At first the data is differenced until it becomes stationary. The number of times differenced gives us the order of differencing, d. Once it is stationary, we find the ACF and PACF plots. After that we find the optimum orders of AR and MA, p and q respectively such that with the pre obtained d, we get minimum AIC(Akaike Information Criterion), i.e., 2(no of paramenters - ln(maximum value of Likelihood function)) . Once, we get the optimal p,q,d we train the model over the training dataset. Finally we forecast the model for the remaining time stamps and calculate the error from actual and forecasted data.

1.6 Analysis and Limitation

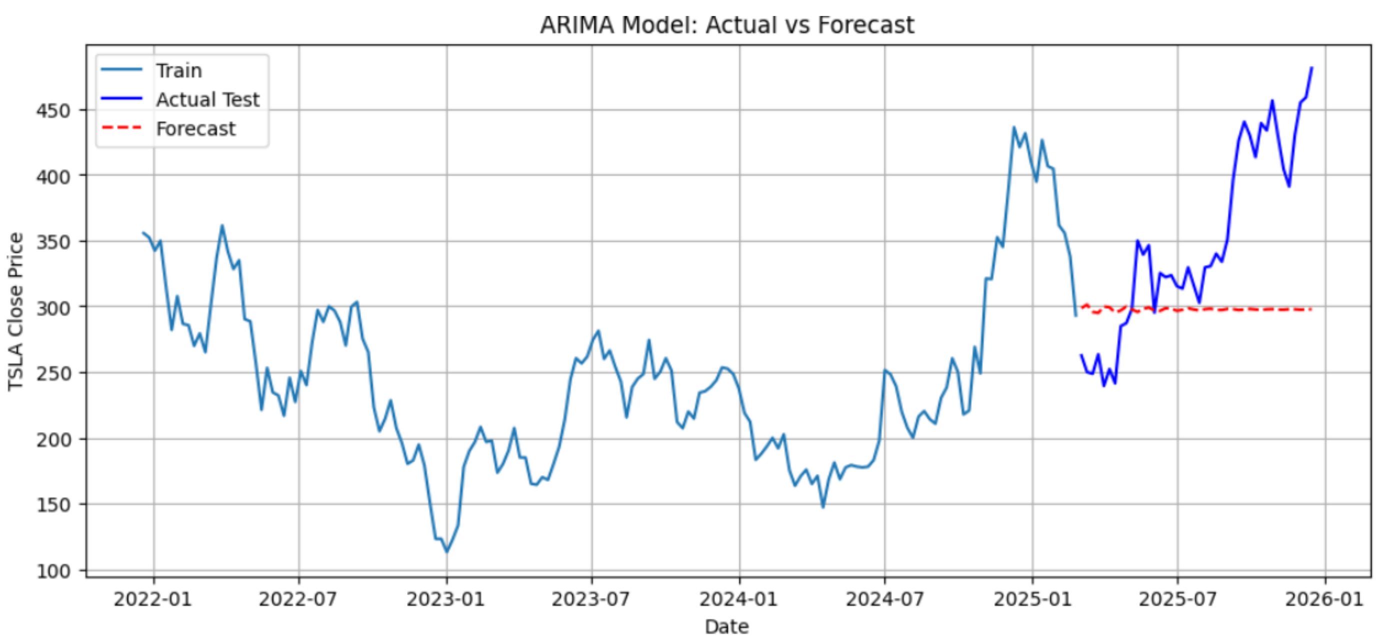

Figure 1: Analysis of Tesla weekly stock closing price

From the final model, we found that the prediction is almost horizontal, however the actual data is very fluctuating. The ARIMA model is not quite accurate as it ignores local variations and needs stationary data. Upon some research, found that LSTM would overcome the shortcomings of ARIMA.