

The background of the slide is a light gray surface covered with numerous small, three-dimensional geometric shapes made of paper. These shapes include cubes, pyramids, and other polyhedrons in various shades of teal and light blue. They are scattered across the frame, with a higher concentration on the right side and a few isolated shapes on the left.

Master's Thesis Time Series Forecasting Using ML and DL



Implementing Stock Price Forecasting

Goal - Predict Closing Price of Stock Market

- ARIMA, short for 'Auto Regressive Integrated Moving Average' is actually a class of models that 'explains' a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$$

ARIMA model in words:

Predicted Y_t = Constant + Linear combination Lags of Y (upto p lags) + Linear Combination of Lagged forecast errors (upto q lags)

THE DATA IS OF ALTABA INC. FROM 1996-04-12 till 2017-11-10



STATIONARITY

- First, we need to check if a series is stationary or not because time series analysis only works with stationary data.
- We perform the ADF (Augmented Dickey-Fuller) Test to test for stationarity.
- p-value is greater than 0.05 so we cannot reject the **Null hypothesis**. Also, the test statistics is greater than the critical values. so the data is non-stationary.
- In order to perform a time series analysis, we may need to separate seasonality and trend from our series. The resultant series will become stationary through this process.

THE ARIMA MODEL

- Its time to choose parameters for ARIMA model. We chose the value of the parameters by observing the plots of ACF and PACF but we use the Auto ARIMA package to get the best parameters.
- The Auto ARIMA model provided the value of p,d, and q as 3,1 and 2 respectively.

```
=====
                        ARIMA Model Results
=====
Dep. Variable:          D.Close    No. Observations:          4886
Model:                 ARIMA(3, 1, 2)  Log Likelihood           9204.201
Method:                css-mle      S.D. of innovations        0.037
=====
```

Model Performance



As we see in the plot our model did quite well. Some common accuracy metrics are as follows:

- MSE: 0.03330921053066402
- MAE: 0.13801238336759786
- RMSE: 0.18250811086267923