Stock Price Prediction

Name:- Suyog Sudhir Pawar

1] Problem Statement

Stock Price Prediction using machine learning helps you discover the future value of company stock and other financial assets traded on an exchange. The stock market appears in the news every day. You hear about it every time it reaches a new high or a new low. The rate of investment and business opportunities in the stock market can increase if an efficient algorithm could be devised to predict the short term price of an individual stock. The proposed solution is comprehensive as it includes pre-processing of the stock market dataset, utilization of multiple feature engineering techniques, combined with a customized deep learning based system for stock market price trend prediction. We conducted comprehensive evaluations on frequently used machine learning models and conclude that our proposed solution outperforms due to the comprehensive feature engineering that we built. The system achieves overall high accuracy for stock market trend prediction. With the detailed design and evaluation of prediction term lengths, feature engineering, and data pre-processing methods, this work contributes to the stock analysis research community both in the financial and technical domains.

2] Market/Customer/business need assessment

The stock market prediction has extra advantages for novice traders as they are the kind of traders who are more prone to making mistakes and facing severe losses in the market compared to experienced traders. You can better analyse and predict the stock market by gaining a complete understanding of the same. One of the main aims of a trader is to predict the stock before its value decline, or buy the stock before the price rises. The efficient market hypothesis states that it is not possible to predict to predict stock prices and that stock behaves in the random walk.

3] Target specification and characterization

Based on the multiple regression model, this study examines the potential predictive effect of customer stock returns to firm stock returns and the moderating effect of diverse customer characteristics on the predictability. Those customer stock returns positively predict firm stock returns in the subsequent month. Additional examinations reveal that the positive predictive effect of customer stock returns on firm stock returns is more intense for firm with high proportion of state-owned customers, customer stability, customer bargaining power and customer concentration than for those with low indicators. Overall, this study contributes to the growing literature on supply chain and predictability of stock returns by shedding light on the forecasting effect of customer stock returns on firm stock returns and the predictive heterogeneity owing to customer characteristics.

4] External search

[1] Masoud, Najeb MH. (2017) “The impact of stock market performance upon economic growth.” International Journal of Economics and Financial Issues 3 (4) : 788–798.

[2] Murkute, Amod, and Tanuja Sarode. (2015) “Forecasting market price of stock using artificial neural network.” International Journal of Computer Applications 124 (12) : 11-15.

[3] Hur, Jung, Manoj Raj, and Yohanes E. Riyanto. (2006) “Finance and trade: A cross-country empirical analysis on the impact of financial development and asset tangibility on international trade.” World Development 34 (10) : 1728-1741.

[4] Li, Lei, Yabin Wu, Yihang Ou, Qi Li, Yanquan Zhou, and Daoxin Chen. (2017) “Research on machine learning algorithms and feature extraction for time series.” IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC): 1-5.

[5] Seber, George AF and Lee, Alan J. (2012) “Linear regression analysis.” John Wiley & Sons 329 [6] Reichek, Nathaniel, and Richard B. Devereux. (1982) “Reliable estimation of peak left ventricular systolic pressure by M-mode echographicdetermined end-diastolic relative wall thickness: identification of severe valvular aortic stenosis in adult patients.” American heart journal 103 (2) : 202-209.

5] Bench marking alternate products

Several intelligent data mining approaches, including neural networks, have been widely employed by academics during the last decade. In today's rapidly evolving economy, stock market data prediction and analysis play a significant role. Several non-linear models like neural network, generalized autoregressive conditional heteroskedasticity (GARCH) and autoregressive conditional heteroscedasticity (ARCH) as well as linear models like Autoregressive Integrated Moving Average (ARIMA), Moving Average (MA) and Auto Regressive (AR) may be used for stock forecasting. The deep learning architectures inclusive of Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), Multilayer Perceptron (MLP) and Support Vector Machine (SVM) are used in this paper for stock price prediction of an organization by using the previously available stock prices. The National Stock Exchange (NSE) of India dataset is used for training the model with day-wise closing price. Data prediction is performed for a few sample companies selected on a random basis.

6] Applicable patents

As we know, patents is not only can protect the companies technologies development and promote the advance of the core technologies but can be used as one of the evaluation indicators to estimate the companies innovation activities. Therefore, patents have an influence on turbulence and fluctuations of companies' stock price. The aim is this study is to adopt patent and financial indicators for stock's price prediction. Thus, we collect patent and financial data as the attributes to predict the stock price by regression analysis. The experimental results show the proposed approach can accurately predict company's stock price.

7] applicable regulations

Government regulation of the organized security markets dates from the enactment of the Securities Exchanges act in June, 1934. This extremely complicated statute, in addition to requiring the registration of all national securities exchanges and of each issue of stock and bonds listed on such markets, granted to two fedral agencies extensive control over stock-exchange trading. The government's ability to change its policy has a substantial effect on stock prices. We compare the model-implied stock prices with their counterparts in a hypothetical scenario in which policy changes are precluded. We find that the government's ability to change its policy amplifies the stock price declines around policy changes. In addition, this ability can imply a higher or lower level of stock prices compared to the hypothetical scenario.

8] Applicable constraints

While the environment is rich with stock price, and fundamental data that is both accessible and free, indiscriminate application of pre-processing techniques and machine learning algorithms will produce indiscriminate results. Financial time series data are incredibly nuanced with the signal to noise ratio systemically low, practitioners spend their careers trying to achieve the elusive aim of generating consistent out performance, with only a few succeeding. Thus the need for a more intimate understanding of the data is pertinent to achieving some semblance of success.

9] Business opportunity

The market is entering the new year with a consolidation. There are rising fears of a recession in the US and concerns about the overvaluation of Indian equities compared to emerging markets.
But still, I have a positive outlook for the Indian market as India still has the highest projections globally in terms of growth which justifies the high valuation we demand.

10] Concept generation

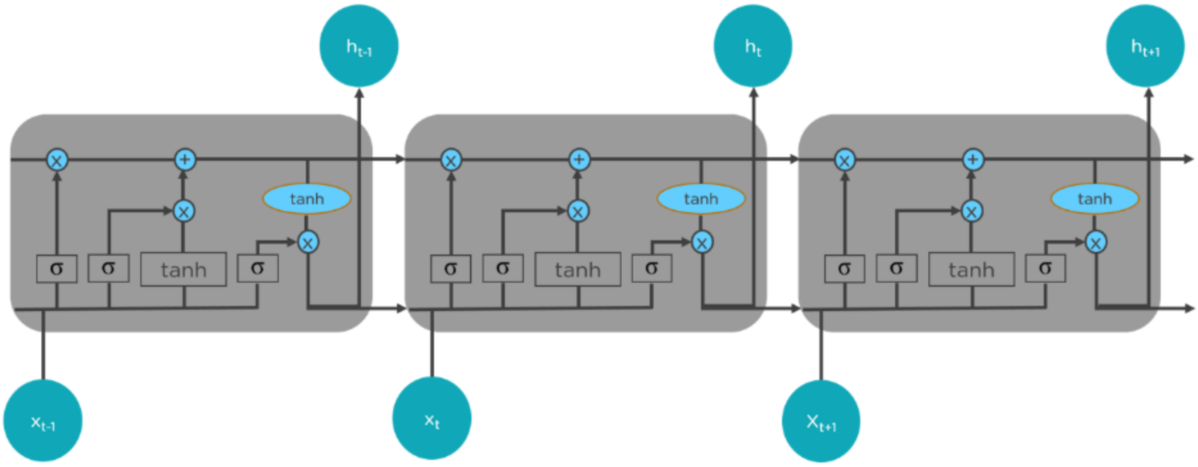
Stock Price Prediction using machine learning helps you discover the future value of company stock and other financial assets traded on an exchange. The entire idea of predicting stock prices is to gain significant profits. Predicting how the stock market will perform is a hard task to do. There are other factors involved in the prediction, such as physical and psychological factors, rational and irrational behaviour, and so on. All these factors combine to make share prices dynamic and volatile.

11] Concept development

There are two prices that are critical for any investor to know: the current price of the investment they own or plan to own and its future selling price. Despite this, investors are constantly reviewing past pricing history and using it to influence their future investment decisions. Some investors won't buy a stock or index that has risen too sharply, because they assume it's due for a correction, while other investors avoid a falling stock because they fear it will continue to deteriorate.

Does academic evidence support these types of predictions, based on recent pricing? In this article, we'll look at four different views of the market and learn more about the associated academic research that supports each view.

12] Final product prototype



Stock market plays a pivotal role in financial aspect of the nation's growth, but stock market is highly volatile and complex in nature. It is affected by significant political issues, analyst calls, news articles, company's future plans of expansions and growth and many more. Hence, any investor would be interested in understanding the stock market overtime and how the factors mentioned above affect the behaviour of the stock market. On Every business day, millions of traders invest in stock market. Most of these investors lose money and others gain. However, considering any trading day, loss or gain is absolutely inconsistent. The demand to predict stock prices are extremely high hence is the need for stock market analysis. This project is focused on analysing a stock of any given company based on statistical technical indicators. Some of these indicators are deterministic in nature and the remaining are probabilistic. The objective of this project is to minimize the risk of loss in every trade thereby maximizing the profit.

```python
import pandas as pd
stock_data = pd.read_csv('./NFLX.csv',index_col='Date')
stock_data.head()
```

	Open	High	Low	Close	Adj Close	Volume
Date						
2019-03-04	359.720001	362.250000	348.040009	351.040009	351.040009	7487000
2019-03-05	351.459991	356.170013	348.250000	354.299988	354.299988	5937800
2019-03-06	353.600006	359.880005	351.700012	359.609985	359.609985	6211900
2019-03-07	360.160004	362.859985	350.500000	352.600006	352.600006	6151300
2019-03-08	345.750000	349.920013	342.470001	349.600006	349.600006	6898800

```python
import matplotlib.dates as mdates
import matplotlib.pyplot as plt
import datetime as dt

plt.figure(figsize=(15,10))
plt.gca().xaxis.set_major_formatter(mdates.DateFormatter('%Y-%m-%d'))
plt.gca().xaxis.set_major_locator(mdates.DayLocator(interval=60))
x_dates = [dt.datetime.strptime(d,'%Y-%m-%d').date() for d in stock_data.index.values]

plt.plot(x_dates, stock_data['High'], label='High')
plt.plot(x_dates, stock_data['Low'], label='Low')
plt.xlabel('Time Scale')
plt.ylabel('Scaled USD')
plt.legend()
plt.gcf().autofmt_xdate()
plt.show()
```

```python
target_y = stock_data['Close']
X_feat = stock_data.iloc[:,0:3]
```

```python
#Feature Scaling
sc = StandardScaler()
X_ft = sc.fit_transform(X_feat.values)
X_ft = pd.DataFrame(columns=X_feat.columns,
                    data=X_ft,
                    index=X_feat.index)
```

```python
def lstm_split(data, n_steps):
    X, y = [], []
    for i in range(len(data)-n_steps+1):
        X.append(data[i:i + n_steps, :-1])
        y.append(data[i + n_steps-1, -1])

    return np.array(X), np.array(y)
```

```python
X1, y1 = lstm_split(stock_data_ft.values, n_steps=2)

train_split=0.8
split_idx = int(np.ceil(len(X1)*train_split))
date_index = stock_data_ft.index

X_train, X_test = X1[:split_idx], X1[split_idx:]
y_train, y_test = y1[:split_idx], y1[split_idx:]
X_train_date, X_test_date = date_index[:split_idx], date_index[split_idx:]

print(X1.shape, X_train.shape, X_test.shape, y_test.shape)
```

```
(755, 2, 3) (604, 2, 3) (151, 2, 3) (151,)
```

```python
lstm = Sequential()
lstm.add(LSTM(32, input_shape=(X_train.shape[1], X_train.shape[2]),
                    activation='relu', return_sequences=True))
lstm.add(Dense(1))
lstm.compile(loss='mean_squared_error', optimizer='adam')
lstm.summary()
```

```python
history=lstm.fit(X_train, y_train,
                 epochs=100, batch_size=4,
                 verbose=2, shuffle=False)
```

```
Epoch 1/100
152/152 - 2s - loss: 0.5510 - 2s/epoch - 16ms/step
Epoch 2/100
152/152 - 1s - loss: 0.1016 - 829ms/epoch - 5ms/step
Epoch 3/100
152/152 - 0s - loss: 0.0124 - 447ms/epoch - 3ms/step
Epoch 4/100
152/152 - 0s - loss: 0.0122 - 313ms/epoch - 2ms/step
```

13] Conclusion

Predicting stock market returns is a challenging task due to consistently changing stock values which are dependent on multiple parameters which form complex patterns. The historical dataset available on company's website consists of only few features like high, low, open, close, adjacent close value of stock prices, volume of shares traded etc., which are not sufficient enough. To obtain higher accuracy in the predicted price value new variables have been created using the existing variables. ANN is used for predicting the next day closing price of the stock and for a comparative analysis, RF is also implemented. The comparative analysis based on RMSE, MAPE and MBE values clearly indicate that ANN gives better prediction of stock prices as compared to RF. For future work, deep learning models could be developed which consider financial news articles along with financial parameters such as a closing price, traded volume, profit and loss statements etc., for possibly better results.