**IDEA** : STOCK PRICES PREDICTION

**COLAB:** <https://colab.research.google.com/drive/1r2XDlZhq8mXFL9UpIWts4QLLMOBxT16s?authuser=1> (LR, DT)

<https://colab.research.google.com/drive/1fUZmMQcHqkXlN_JMYvpzNZ6O9TpKfegc?authuser=1> (main)

LR -> <https://colab.research.google.com/drive/1mfnFoiA2bgQP20lj-YU89AXIrSnrTyiL?authuser=1#scrollTo=r-IEVxCzW5iF>

DT -> <https://colab.research.google.com/drive/1KHkBVLWio5gM-dCV4OvgW9FIMM26FHts?authuser=1#scrollTo=WOpwXLHwbvBJ>

RF -> <https://colab.research.google.com/drive/1N41x4epBggtUFk-eKpmx6sC5EaaEOU7k?authuser=1#scrollTo=-dkLDXAaKBQg>

LSTM ->

<https://colab.research.google.com/drive/1gxLkBYeEH0EXHjSVbwDpGo71buQCB218?authuser=1#scrollTo=keiGPHHjk3OM>

ARIMA ->

<https://colab.research.google.com/drive/1-vEFYrUxh_GCnbQCTqlCq7dSKKaB-HEv?usp=sharing>

REPORT:

<https://docs.google.com/document/d/1qlNh4T0tms9aVCILkvVVn3GyM0VnWUjRTcRNZ35WwcA/edit#heading=h.yrktn71x9cz5>

EXCEL:

<https://docs.google.com/spreadsheets/d/1R4yWJXMkgsuBliT2uWvJMaVygb_TNWXyflKqn9XG4vs/edit?usp=sharing>

Video to watch:

<https://youtu.be/Vfx1L2jh2Ng>

Sir suggestion: CART model (Decision tree), ARIMA model, sp500 paper

Report, Code

**ALGO: LSTM, GRU**

**STEPS TO FOLLOW:**

# Starting from data collection,

# Dataset preparation, (train, test)

# Feature engineering,

# Understanding and

# Applying the right algorithms, and (linear, LSTM etc)

# Reporting the evaluation results. (errors/accuracy)

That said, it is better to pick a project that you will be able

to go deep with (regarding trying different methods, error analysis, etc.)

**Resources:**

<https://www.google.com/search?q=Stock+Prices+Predictor+ML+using+python&oq=Stock+Prices+Predictor+ML+using+python&aqs=chrome..69i57j33i160l3j33i22i29i30.8232j0j7&sourceid=chrome&ie=UTF-8> (X)

<https://www.simplilearn.com/tutorials/machine-learning-tutorial/stock-price-prediction-using-machine-learning>

<https://data-flair.training/blogs/stock-price-prediction-machine-learning-project-in-python/>

<https://www.analyticsvidhya.com/blog/2018/10/predicting-stock-price-machine-learningnd-deep-learning-techniques-python/?fbclid=IwAR3pJ76pOX1oKmzx25GogSPl0Z329BW31cdWPUC3m8b9FCWmCcvA4DBpe5E>

YT: <https://youtu.be/KUFmCwCVXWs> concept (X)

<https://www.youtube.com/watch?v=kGdbPnMCdOg> time-series concept(X)

<https://www.youtube.com/watch?v=CbTU92pbDKw> ms stock lstm

<https://www.youtube.com/watch?v=1O_BenficgE> [good video]

Paper With Code: <https://paperswithcode.com/>

Colab: <https://colab.research.google.com/drive/1b3CUJuDOmPmNdZFH3LQDmt5F0K3FZhqD?usp=sharing> (weather forecasting using LSTM multivariate)

<https://colab.research.google.com/drive/1Bk4zPQwAfzoSHZokKUefKL1s6lqmam6S?usp=sharing#scrollTo=PSIHfWL23fBi> (microsoft stock LSTM)

Github:

<https://github.com/harishpuvvada/BitCoin-Value-Predictor> (to get some idea)

<https://github.com/dataquestio/project-walkthroughs/tree/master/sp_500> (good repo)

**About stocks:**

<https://www.investopedia.com/terms/s/stockmarket.asp>

**NOTES:**

**LSTM:**

1. numpy, pandas, matplotlib - numerical computation, data manipulation, data visualization

2. Plotly - build graphs

-> LSTM (Long Short Term Memory) Network (using Google dataset)

(Used for processing & predicting on the basis of time series data)

-> Train model (#Epoch = iterations)

-> Plot graph

-> Test model (using new dataset)

-> Actual stock price vs predicted stock price (Time vs stock price graph)

3. sklearn.preprocessing -> MinMaxScaler

4. Keras

**Summary:**

The target is to discover the future value of company stock by

stock market prediction using machine learning models.

**TIPS**

# Dataset

# Time Series Data

# Model -> ARIMA, LSTM, Transformer (deep learning neural network)

# Data integration

Factors -> Financial, economic, environmental

# Diff models comparison analysis

# Without company specific -> diff companies

# Bangladesh specific -DSE (Dhaka Stock Exchange)

**DATASET LINKS:**

**>>> https://finance.yahoo.com/quote/MSFT/history?period1=511056000&period2=1667692800&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true**

**>>>** [**https://www.kaggle.com/discussions/general/272226**](https://www.kaggle.com/discussions/general/272226) **[list of notebooks for predicting stock price]**

# <https://www.kaggle.com/code/faressayah/stock-market-analysis-prediction-using-lstm> (used)

# <https://www.kaggle.com/datasets/borismarjanovic/price-volume-data-for-all-us-stocks-etfs> (X)

#<https://www.kaggle.com/code/thikhuyenle/stock-price-prediction-using-lstm> (in progress)

# <https://www.kaggle.com/code/qusaybtoush1990/tesla-stock-price> (done)

HOW CAN WE IMPROVE THE DATASET:

Papers:

# https://arxiv.org/pdf/1912.10806v1.pdf

1. By introducing new columns: inflation rate
2. Natural language processing: by seeing different news websites and analyzing different behavior of different companies.that can help more to predict stock price of a company.
3. Stock prices are driven by a number of factors: industry performance, company news and performance, investor confidence, micro and macro economic factors like employment rates, wage rates, etc >> stocktwits
4. This paper explores the relationship between StockTwits and stock market closing prices through URL mining and sentiment analysis to check whether this information can enhance stock price prediction accuracy.
5. As a result of adding smart users into the dataset, the dataset shrank in size but the accuracy of the model got better. One problem, however, was that there were many unpredictable days due to the narrowed dataset. Only companies that were most frequently discussed could be predicted consistently using this model.
6. In this method, a model is required to make a one-week prediction, and the actual data for that week is used in the model for making the forecast for the next week. This is both realistic and practical, as in most of the real-world applications, forecast horizons longer than one week are not used.
7. It also reveals the fact that multivariate analysis is not a good idea in LSTM-based regression, as univariate models are more accurate and faster in their execution.