**PROJECT DOCUMENTATION**

**-Team 17**

**Stock Market Prediction**

**The Blueprint file structure follows the following pattern:**

Dashboard

Creation

Model

Creation

Feature

Engineering

EDA

Data

Preparation

Data

**Data Set Link:**

<https://github.com/sravanirayapati/Stock_Market_Prediction/tree/main/Datasets>

Collection of Data Sets :

* DataFrame.csv
* MSFT.csv

**Data Preparation:**

* Data Cleaning: Identifying and correcting mistakes or errors in the data.In data sets, Neither Missing nor duplicate rows present.
* Identifying input variables that are more relevant to the task.
* Adding new features and attributes to the data sets(DataFrame.csv, MSFT.csv)

**EDA:**

* Importing the Data Sets.
* See the view and shape of the data set.
* Descriptive statistics of the data set.
* Checking about the correlation between features In a data set.
* Checking about data types and missing values in the data.

**Feature Engineering:**

* Imputation: if missing values are present ,impute them.
* Using Logistic Regression and SVM algorithm we try to analyze the daily stocks and predict for the next day.
* Encoding categorical features and Standardization of data.

**Model Creation:**

* SVM model can be applied to stock price data through the Great Recession and subsequent recovery period.
* Data Collection and Timeframe.
* SVM Model : The specific kernel function we use in this study is the radial kernel.
* Feature Selection(we use four features to predict stock price direction – price volatility, price momentum, sector volatility, and sector momentum)
* Method.

**Dashboard Creation:**

* Identify key Finance formulas and its parameters.
* Create a data source of information .Importing the Data Sets.
* Analyze the data set and create charts of the analyzed data.
* Create a summarized dashboard of charts and info graphics.

**INSIGHTS**

* Importing Required libraries for model building
* Loading data into notebook
* Reading data into a data frame
* Perform some important EDA-

1. Converting date feature into date time format
2. Checking description of data, how data varies
3. Checking last five rows and shape of data frame
4. Predicting Stock prices starts from and ends at

* Visualize the spread andskewness through the distribution plot Use the

Column list

* correlation
* View the matrix in a table to identify the numerical values of strengths
* Outlier Detection and Removal Extensive Analysis on Historical Data to Find Patterns

1. Since the datais time series data, we should be able to predict the future through

forecasting techniques

1. Delete the index column due to reset
2. Decompose the time series year-wise and month-wise to analyze further
3. Firstly, plot the data year-wise to see the duration of when it hiked and  dipped
4. Group the data by year and plot

* Model Building using LSTM Networks
* Checking model summary, how many neurons in network
* Compiling the model with optimizer and loss function
* Model fitting
* Test data preparation for model prediction
* Model prediction using test data
* making into normal scale from standardizing
* Predicting output
* After Prediction evaluating with actual vs predicted
* Plot the graph of Stock Price Prediction over the years using LSTM.