

**Pro\_220064\_Team\_1\_phase\_3**

STOCK PRICE PREDICTION

* Introduction
* Stock market is an aggregation stockbrokers and traders who can buy and sell shares of stocks.
* Stock data is non-stationary, chaotic, random and depends on several technical parameters.
* Since statistical approaches are linear in nature, it hampere prediction performances in case of sudden rise or fall of prices of or fall of stock .
* Combine multiple models and sources of information to improve prediction accuracy.
* In modern days of artificial intelligence, machine learning plays a important role in time series predictions.
* Objectives
* To predict stock prices using both statistical and machine learning approaches.
* To make comparison between statistical and machine learning approach predictions.
* To find a better approach, which predicts prices of the stocks more accurately.
* Linear Regression
* Linear regression is a type of supervised machine learning algorithm that computes the linear relationship between a dependent variable and one or more independent features.
* The goal of the algorithm is to find the best linear equation that can predict the value of the dependent variable based on the independent variable.

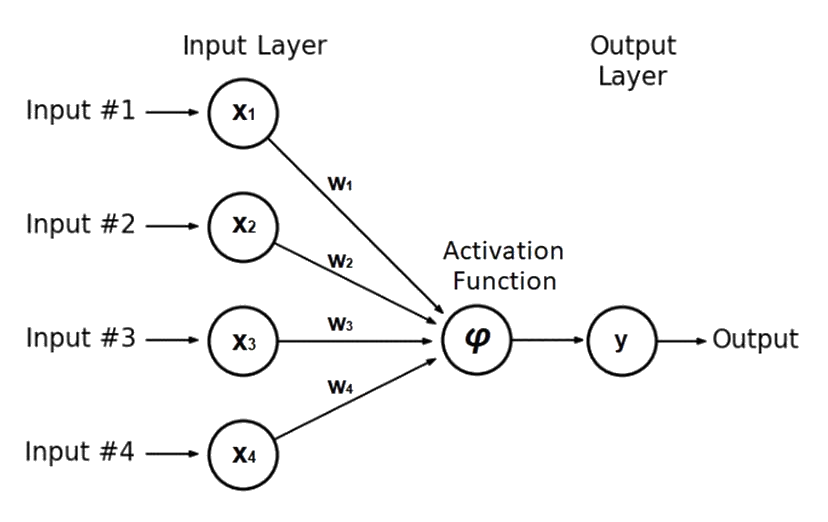
# Machine Leaning MethodS

# **Random Forest**

# **Support Vector Machine (Support Vector Regression)**

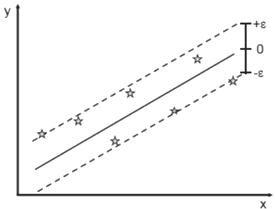
# Random Forest

* ***Random Forest algorithm is one of the ensemble learning algorithms.***



# Support Vector Machine (Support Vector Regression)

* **SVM can be used as regression method.**
* **Support Vector Regression uses the same principles as the SVM for classification, with only a few minor differences.**
* **A margin of tolerance (ϵ) is set in approximation to the SVM.**



# PROGRAM:

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

sns.set\_style('whitegrid')

plt.style.use("fivethirtyeight")

%matplotlib inline

# For reading stock data from yahoo

from pandas\_datareader.data import DataReader

import yfinance as yf

from pandas\_datareader import data as pdr

yf.pdr\_override()

# For time stamps

from datetime import datetime

# The tech stocks we'll use for this analysis

tech\_list = ['AAPL', 'GOOG', 'MSFT', 'AMZN']

# Set up End and Start times for data grab

tech\_list = ['AAPL', 'GOOG', 'MSFT', 'AMZN']

end = datetime.now()

start = datetime(end.year - 1, end.month, end.day)

for stock in tech\_list:

globals()[stock] = yf.download(stock, start, end)

company\_list = [AAPL, GOOG, MSFT, AMZN]

company\_name = ["APPLE", "GOOGLE", "MICROSOFT", "AMAZON"]

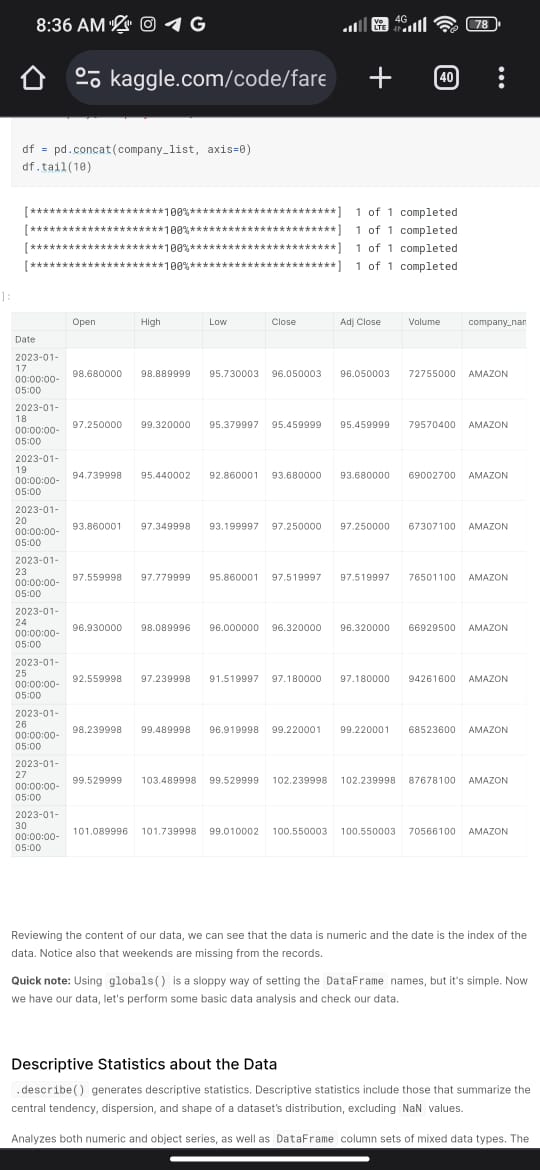
for company, com\_name in zip(company\_list, company\_name):

company["company\_name"] = com\_name

df = pd.concat(company\_list, axis=0)

df.tail(10)

**OUTPUT;**

****

# Prediction

* As statistical methods, predictions have been performed using 10-day, 15-day, 30-day Simple Moving Average and Weighted Moving Average, Exponential Smoothing with α = 0.3, 0.5, 0.75 and naïve approach.
* As Machine Learning methods, predictions have been performed using Simple Linear Regression, Lasso Regression and Ridge Regression, K-Nearest Neighbor, Random Forest with different number of estimators, Support Vector Machine and Neural Network Models like SLP, MLP and LSTM.

# Current stock price prediction

****

* Feature Selection
* **For time series prediction, selection of features is an important task.**
* **Because selection of worst features can direct the prediction to a wrong way.**
* **In this system, three features have been selected.**

**i. The opening price**

**ii. The highest price**

**iii. The lowest price.**

# Conclusion

* A comparative study between statistical approaches and machine learning approaches has been done in terms of prediction performances.
* Machine learning methods, especially, MLP and LSTM are found to be the most accurate to predict stock prices.

THANK YOU