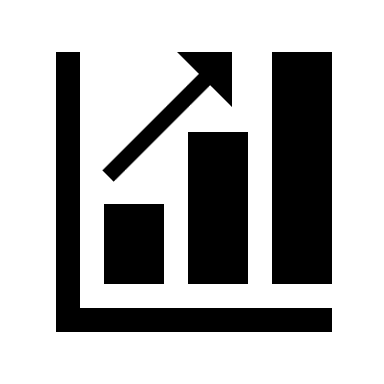
**STOCK PRICE PREDICTION**

**(PHASE-2)**

**Introduction** 

**Stock Price Prediction** is the task of forecasting future stock prices based on historical data and various market indicators. It involves using statistical models and machine learning algorithms to analyse financial data and make predictions about the future performance of a stock. The goal of stock price prediction is to help investors make informed investment decisions by providing a forecast of future stock prices.

**Prediction methods**

Prediction methodologies fall into three broad categories which can (and often do) overlap. They are fundamental analysis, technical analysis (charting) and technological methods.

**Fundamental analysis:**

Fundamental analysis is a method of evaluating the intrinsic value of a stock by analysing various financial and economic factors related to the company. It is used to make predictions about a stock's future price based on its underlying fundamentals.

**Technical analysis:**

Technical analysis is a method of predicting future stock prices by analysing historical price and volume data and identifying patterns, trends, and indicators in stock charts. Unlike fundamental analysis, which focuses on a company's financial health and performance, technical analysis primarily relies on past market data.

**Technological methods:**

Technological methods in stock price prediction leverage advanced technologies and data analysis techniques to make more accurate and data-driven predictions.

**Processes:**

Predicting stock prices involves various

processes like

▪ Data collection

▪ Data preprocessing

▪ Feature engineering

▪ Model selection

▪ Model training

▪ Evaluation

**Data collection:**

Data collection for stock price prediction typically involves gathering various types of financial and market data. Here we are going to use historical stock market data. The link for the historical stock market dataset is given below:<https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset>

**Data preprocessing:**

* Proper data preprocessing enhances the quality of input data and contributes to the effectiveness of your stock price prediction model.
* The specific steps and techniques employed may vary depending on the modeling approach and the characteristics of the data.
* Handling missing data and converting categorical features into numerical features (for example: each stock can be categorized in terms of its investment style as a growth stock or a value stock.
* In CNN-LSTM model , we normalise the data to ensure all features are on the same scale. Common techniques include min-max scaling or z-score standardization.
* Divide the data into training , validation , and testing sets to evaluate the model’s performance.

**Feature engineering:**

Creating additional features to enhance the

predictive power of the model.

* Converting a series of asset prices into percent change values is a simple form of feature engineering.
* Charting prices vs. a moving average is an implicit form of feature engineering.
* Any technical indicator (RSI, MACD, etc...) are also forms of feature engineering.

**Model selection:**

Selecting the appropriate model for stock price prediction is a crucial step in the process, as different models have different strengths and weakness. Here are some common models and methods used in stock price prediction:

**Time Series Models**:

**Autoregressive Integrated Moving Average (ARIMA):** Suitable for univariate time series data with stationary patterns.

**Machine Learning Models:**

**Linear Regression:** Simple models that can work well with basic time series features. May need additional data preprocessing and feature engineering.

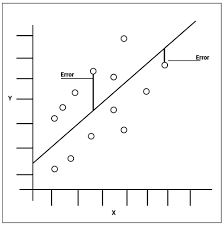
**Support Vector Machines (SVM):** Effective for classification tasks related to price movements.

**Neural Networks (e.g., LSTM, CNN-LSTM):** Suitable for capturing complex patterns and dependencies in time series data. Here for this stock prediction we are going to select CNN-LSTM model.

**Model training:**

Train the chosen model with the pre processed  data.

* Feed sequences of data into the model to make predictions for the target values (e.g., future stock prices).
* Use an optimizer (e.g., Adam) to minimize the loss and update the model's weights
* Compute the loss, using a regression loss function such as mean squared error (MSE).

The Mean Squared Error measures how close a regression line is to a set of data points. It is a risk function corresponding to the expected value of the squared error loss. It’s calculated by taking the average, the mean, of errors squared from data as it relates to a function. 

**Evaluation:**

Evaluating the model's performance on the testing dataset, using the same metrics as used during validation. Assess metrics like MAE, MSE, Root Mean Squared Error (RMSE), and others.

**Mean Absolute Error (MAE)**:MAE is the average of the absolute differences between the predicted and actual values. It measures the average prediction error. A lower MAE indicates better performance.

**Mean Squared Error (MSE):**MSE is the average of the squared differences between the predicted and actual values. It gives more weight to larger errors. A lower MSE means the model's predictions are closer to the actual values.