K. K. Wagh Institute of Engineering Education and Research
Hirabai Haridas Vidyanagari, Amrut Dham, Panchavati, Nashik-422003

Project based Learning

Report on

"Stock Price Prediction Model"

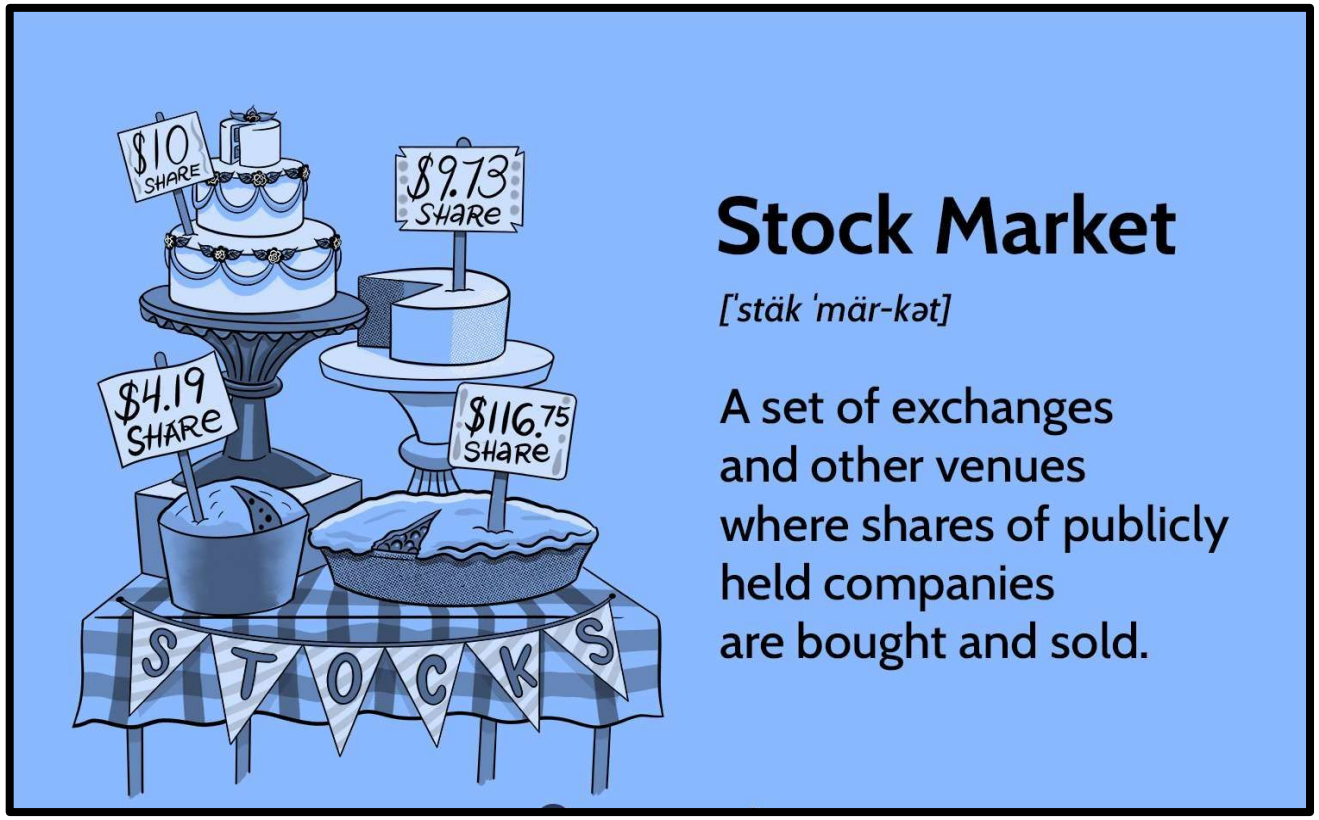
Presented By

Advait Khandalkar
Chetan Killewale
Om Shirsat
Parimal Shinde
Rameshwar Pol
Saish Joshi

Second Year Engineering
Savitribai Phule Pune University
(2022-2023)

OUTLINE OF PRESENTATION

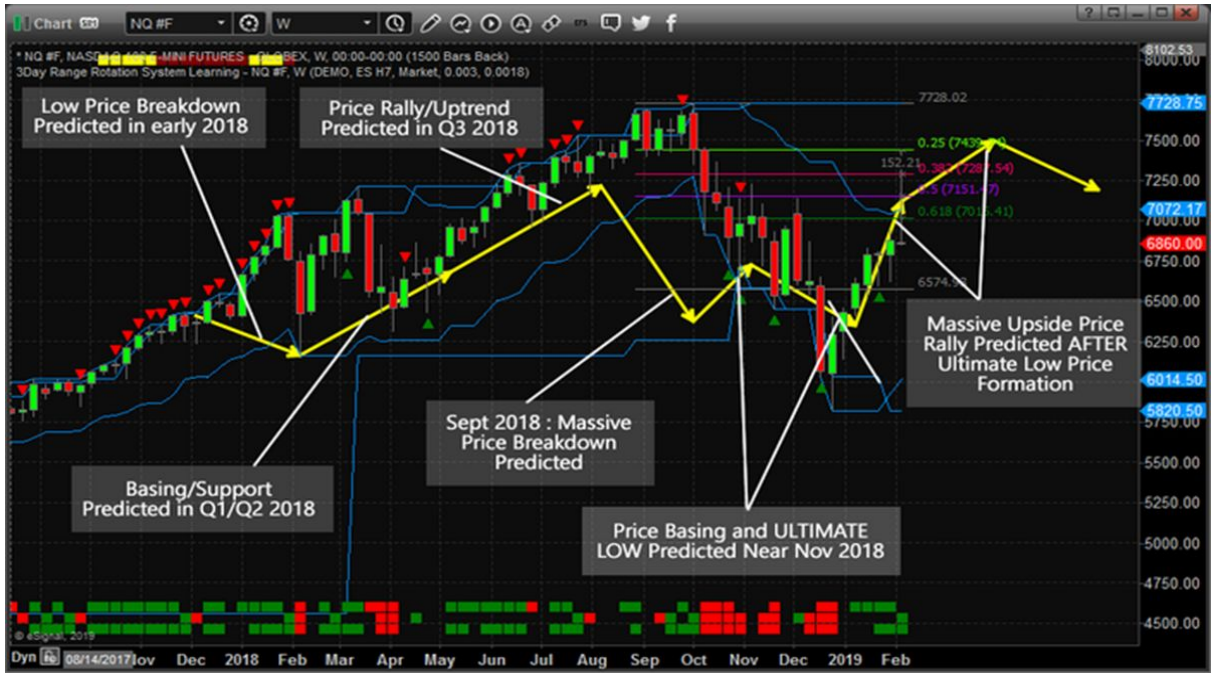
- Introduction
- Problem Statement
- Purpose of Study
- Idea
- Adams Model
- Methodology
- Data Collection and Analysis
- Results and Discussion
- Conclusion
- References



INTRODUCTION

- **Introduction :**

- In today's volatile and dynamic financial markets, accurate predictions of stock prices are crucial for making informed investment decisions.
- Our model aims to leverage advanced data analysis techniques and machine learning algorithms to forecast future stock prices, providing investors with a competitive edge.



Problem Statement :

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.

PURPOSE OF THE STUDY

- **Need of Stock Price Prediction?**

1. **Investment Decision Making:** Predictions help investors identify opportunities and risks, determine when to buy or sell, and optimize their portfolio.
2. **Financial Planning:** Predictions aid individuals in planning their financial goals, estimating returns, and guiding asset allocation decisions.
3. **Algorithmic Trading:** Predictions drive algorithmic trading strategies for rapid decision-making in the market.
4. **Market Analysis and Research:** Predictions provide insights into market conditions, trends, and dynamics.

Idea/Concept

Idea/Concept

- The primary goal is to analyze historical stock price data and advanced data analysis techniques to forecast future stock prices.
- We employ machine learning algorithms, such as regression models or neural networks, to capture patterns and relationships in the data.
- Our model provides a systematic approach to predicting stock prices, enhancing investment decision-making processes.
- The concept revolves around utilizing data-driven insights and machine learning algorithms to deliver actionable predictions for improved investment outcomes.

Benefits of Using Stock Price Prediction Model

- **Market Analysis:** Models aid market analysis by analyzing historical patterns and indicators, offering insights into trends and correlations.
- **Financial Planning:** Models assist in financial planning by estimating future stock prices, setting investment goals, and making informed decisions.
- **Improved Market Understanding:** Stock price prediction models deepen market understanding by analyzing vast amounts of data and identifying patterns that may not be easily recognizable by human analysts.
- **Competitive Advantage:** Utilizing stock price prediction models can give investors a competitive advantage by identifying potential investment opportunities before they are widely recognized in the market.
- **Reduced Emotional Bias:** Models help reduce emotional bias in decision-making, as they rely on data-driven analysis rather than subjective opinions or emotions.
- **Backtesting and Performance Evaluation:** Stock price prediction models can be backtested and evaluated for performance, allowing investors to assess the accuracy and effectiveness of the model before making real-time investment decisions.

LSTM :

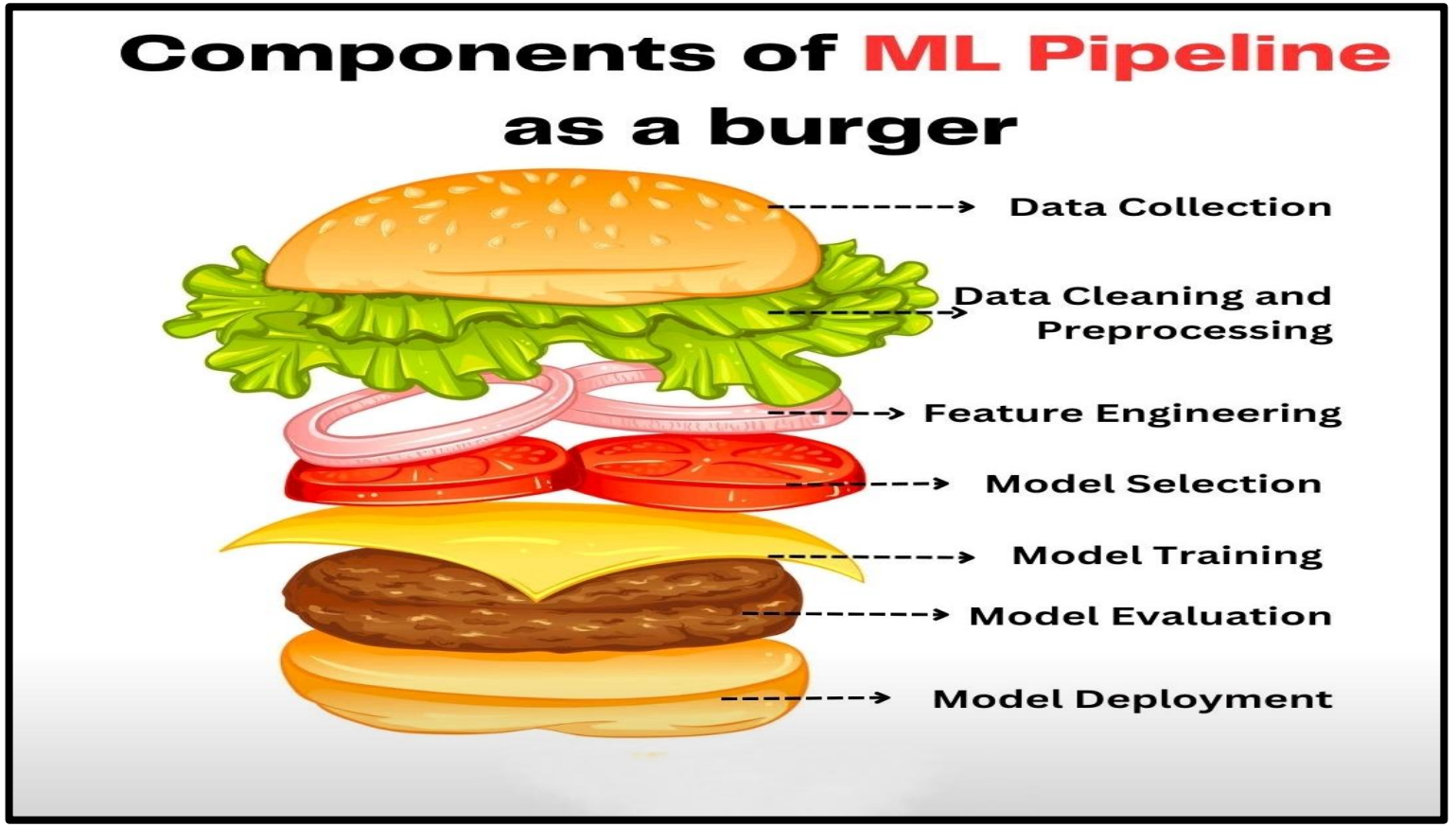
LSTM networks are commonly used for stock market prediction. They capture temporal dependencies in sequential data. The process involves preparing historical stock price data, designing the LSTM model architecture, training the model using historical data, and using the trained model for prediction. It's important to note that stock prediction is challenging, and LSTM models don't guarantee accurate results. Other techniques and domain knowledge should be considered for better predictions.

ADAMS MODEL :

- Data Collection: Gather historical stock price data and relevant features.
- Data Preprocessing: Clean and transform the data for training.
- Feature Engineering: Select and create meaningful features.
- Model Training: Train a machine learning algorithm on the preprocessed data.
- Model Evaluation: Assess the model's accuracy using evaluation metrics.
- Prediction and Deployment: Use the trained model to make predictions on new data.

The iterative nature of the process allows for parameter tuning and continuous monitoring to adapt to market conditions.

How Machine Learning Model Works ?



1. Data Collection	Gather historical stock price data and relevant features.
2. Data Preprocessing	Clean the data and prepare it for analysis.
3. Feature Engineering	Select and create meaningful features from the data.
4. Splitting the Data	Divide the data into training and testing sets.
5. Model Training	Train the machine learning model using the training data.
6. Model Evaluation	Evaluate the model's performance using appropriate metrics.
7. Prediction	Use the trained model to make predictions on new data.

Included libraries :

- `math`: A standard Python library for mathematical operations.
- `threading`: A library for creating and managing threads in Python.
- `pandas`: A powerful data manipulation library for data analysis.
- `numpy`: A library for numerical operations in Python.
- `datetime`: A module for working with dates and times in Python.
- `pandas_datareader`: A library for fetching financial data from various sources, including stock market data.
- `MinMaxScaler`: A class from the scikit-learn library used for scaling or normalizing data.
- `matplotlib.pyplot`: A plotting library for creating visualizations in Python.

Data collection :

1. Data Sources: Identify reliable sources such as financial data providers or APIs for real-time and historical stock market data.
2. Historical Data: Collect stock price data (open, high, low, close) and trading volume for the desired time period. Consider additional features like technical indicators or fundamental data.
3. Data Cleaning: Handle missing values, outliers, and adjust for stock splits or dividends.
4. Data Storage: Set up a system (database, file system, or cloud storage) to organize and store the collected data.

Data Preprocessing :

1. Data Cleaning: Handle missing values and outliers in the collected data.
2. Feature Selection and Engineering: Select relevant features and create new ones that capture important patterns or relationships.
3. Normalization and Scaling: Standardize numerical features to bring them to a common scale.
4. Handling Categorical Variables: Encode categorical variables appropriately for modeling purposes.

Splitting the Data :

1. Training Set: Allocate the majority of the data (around 70-80%) to the training set. This subset is used to train the machine learning model and capture underlying patterns.
2. Validation Set: Set aside a portion of the data (around 10-15%) for the validation set. It is used to fine-tune the model, select optimal hyperparameters, and assess its performance during training.
3. Testing Set: Reserve a separate portion of the data (around 10-20%) for the testing set. This unbiased dataset is used to evaluate the final performance of the trained model on unseen data.
4. Randomization and Stratification: Randomly shuffle the data before splitting to remove any order or biases.

Training Dataset :

Model Training :

1. Model Selection: Choose an appropriate machine learning algorithm or model architecture specifically designed for stock market prediction tasks, considering factors like regression models, time series models, or ensemble models.
2. Choose an appropriate sequence length and batch size for training the LSTM model. Additionally, utilize the Adam optimizer for efficient parameter updates during training. Experiment with different values for sequence length and batch size to find the optimal configuration while leveraging the benefits of the Adam optimizer.
3. Training Data: Use the allocated training dataset to train the model, ensuring it captures the underlying patterns and relationships in the stock market data.

Model Evaluation :

Model evaluation is crucial in stock market prediction to assess the performance of trained models. By employing evaluation metrics, cross-validation, out-of-sample testing, and benchmarking, stakeholders can measure accuracy, generalization, and compare against baselines for informed decision-making. This section highlights key points for effective model evaluation in stock market prediction.

1. Evaluation Metric: Use appropriate evaluation metrics like RMSE to measure the predictive accuracy of the model.
2. Cross-Validation: Perform cross-validation to estimate the model's generalization ability and ensure robust evaluation.
3. Out-of-Sample Testing: Evaluate the model on independent data to validate its performance on unseen stock prices.
4. Benchmarking: Compare the model's performance against relevant benchmarks or baseline models to assess its predictive capabilities.

RMSE (ROOT MEAN SQUARE ERROR) VALUE :

The Root Mean Squared Error (RMSE) is one of the two main performance indicators for a regression model. It **measures the average difference between values predicted by a model and the actual values**. It provides an estimation of how well the model is able to predict the target value (accuracy).

```
[ ]  #get the rmse
     rmse=np.sqrt(np.mean(((predictions- y_test)**2)))
     rmse

     0.7243791381194871
```

Our RMSE value (Infosys Stock) : **0.7243791381194871**

Predicted Data Values On Testing Data:

Date	Close	predictions
2021-05-28	19.340000	18.489100
2021-06-01	19.270000	18.625912
2021-06-02	19.290001	18.726149
2021-06-03	19.139999	18.798601
2021-06-04	19.290001	18.826643
...	...	...
2022-12-23	17.920000	17.822947
2022-12-27	18.030001	17.775915
2022-12-28	17.959999	17.753649
2022-12-29	18.240000	17.735090
2022-12-30	18.010000	17.759699

402 rows × 2 columns

Graphical Representation of Model :

References :

1. Yahoo Finance https : finance.yahoo.com

2. Google Colab https : colab.research.google.com

3.Our ML model link :

https://colab.research.google.com/drive/1bOvm7xQ7M7qRMGxaYZW9untjCa-zLd1K?usp=sharing

4.Information about ADAMS model :

https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/

THANK YOU