

MINOR PROJECT '23 - SEMESTER 5

Exploratory Data Analysis (EDA) and Financial Indicator Development for NSE Nifty 50 Stock Market Data Using Machine Learning



TEAM:

Paakhi Maheshwari 21103149

Ayush Gupta 21103016

Arpit Verma 21103254

Faculty Supervisor:

Mr. Amit Mishra

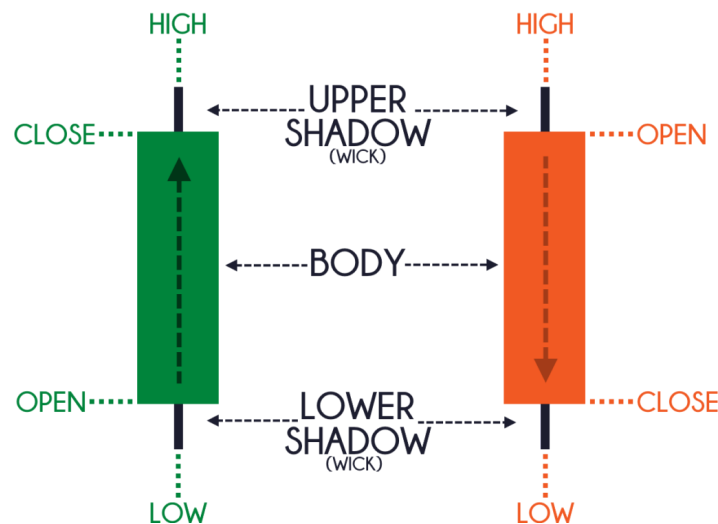
Introduction

The stock market is a dynamic and complex environment where investors strive to make informed decisions. This project aims to leverage data science and machine learning techniques to analyze previous NSE Nifty 50 stock market data, gain insights through exploratory data analysis (EDA), and develop a financial indicator to predict future stock price patterns using machine learning.

Financial indicators are crucial tools used in the field of finance and investment to assess the performance of financial assets – usually, stocks. These indicators provide valuable insights into various aspects of financial markets and are used by investors to make informed investment decisions. Uncalculated decisions in the stock market can lead to financial loss, hence, we analyse historical trends and predict future patterns using Machine Learning techniques to make efficient decisions.

For the analysis of stocks, we look at the OHLC (Open **H**igh **L**ow **C**lose) candles of all the records:

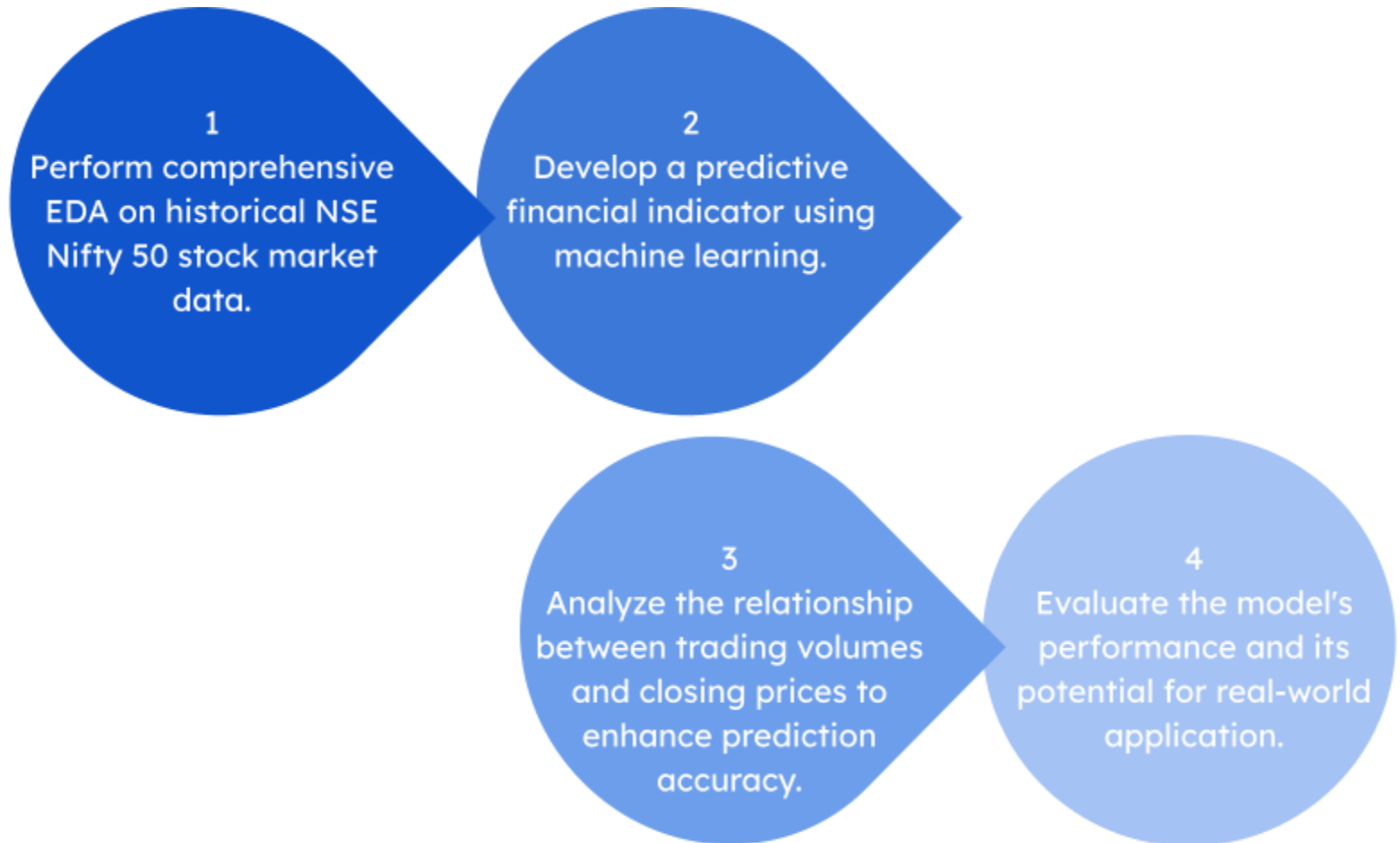
- **Open** is the price of a stock when a time resolution started
- **High** is the highest price reached
- **Low** is the lowest price reached
- **Close** is the price of a stock when a time resolution finishes.



The green represents the stocks which are profitable (buy), while the orange indicates those in loss (sell).

For this project, we will employ the values of Simple and Exponential Moving Averages to build our financial indicators.

Our Objective



This project is not just stock predication, but it will predict trends of a particular timeframe and help investors make calculated decisions.

Our Approach to the Project

Data Collection and Preprocessing:

We will collect historical NSE Nifty 50 stock market data, including daily closing prices and trading volumes, from reliable sources. Data preprocessing will involve cleaning, handling missing values, and ensuring data consistency.

Exploratory Data Analysis (EDA):

- Visualize historical price trends, including daily, weekly, and monthly patterns.
- Calculate descriptive statistics, such as mean, median, and standard deviation.
- Identify outliers, if any, and assess their impact on analysis.
- Conduct correlation analysis to understand the relationship between trading volumes and closing prices.

Machine Learning Model Development:

We will employ various machine learning algorithms, including regression and time series forecasting techniques, to develop a predictive financial indicator. This indicator will use historical data to forecast future stock price patterns.

Model Evaluation and Validation:

We will split the dataset into training and testing subsets to evaluate the model's performance. We will use Simple and Exponential Moving Averages for the same.

Technologies and Resources Required

Python Programming:

We will utilize Python as the primary programming language for data analysis and machine learning. Libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Scikit-Learn will be used for data manipulation, visualization, and modeling.

Data:

- NSE Nifty Data has been provided to us for analysis and model development.

Machine Learning Frameworks:

- Scikit-Learn for implementing machine learning algorithms.
- On further evaluation, we may use TensorFlow or Keras for implementation if it is needed.

Visualization:

- Matplotlib and Seaborn libraries of Python for creating visualizations.

Documentation:

- Jupyter Notebooks for code documentation and reports..

Conclusion and Acknowledgments

Conclusively, The successful completion of this project has the potential to enhance decision-making for investors and traders in the stock market. By leveraging data science and machine learning, we aim to provide valuable insights and tools for financial analysis and prediction.

We would like to thank our faculty supervisor, Mr. Amit Mishra, for his continued support and guidance in this project.