IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

STOCK MARKET PREDICTION

[1]Dr.Nitin Janwe, [2] Pragati Chandekar, [3] Ashwini Chandalwar, [4] Khushi Dorlikar

Assistant Professor

[1, 2, 3, 4] Department of Computer Science Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India

Abstract

In Stock Market Prediction, the aim is to predict the future value of the financial stocks of a company. The recent trend in stock market prediction technologies is the use of machine learning which makes predictions based on the values of current stock market indices by training on their previous values. Machine learning itself employs different models to make prediction easier and authentic. The paper focuses on the use of Regression and LSTM based Machine learning to predict stock values. Factors considered are open, close, low, high and volume.

Keywords :- Predictive Modeling, Machine Learning, Time Series Analysis, SARIMA Model

Introduction

In recent times stock market predictions is gaining more attention, maybe due to the fact that if the trend of the market is successfully predicted the investors may be better guided. The profits gained by investing and trading in the stock market greatly depends on the predictability. If there is a system that can consistently predict the direction of the dynamic stock market will enable the users of the system to make informed decisions. More over the predicted trends of the market will help the regulators of the market in taking corrective measures.

The aim of the project is to examine a number of different forecasting techniques to predict future stock returns based on past returns and numerical news indicators to construct a portfolio of multiple stocks in order to diversify the risk. We do this by applying supervised learning methods for stock price forecasting by interpreting the seemingly chaotic market data.

A stock market, equity market or share market is the aggregation of buyers and sellers (a loose network of economic transactions, not a physical facility or discrete entity) of stocks (also called shares), which represent ownership claims on businesses; these may include securities listed on a public stock exchange as well as those only traded privately. Examples of the latter include shares of private companies which are sold to investors through equity crowd funding platforms. Stock exchanges list shares of common equity as well as other security types, e.g. corporate bonds and convertible bonds.

Stock price prediction is one of the most widely studied problem, attracting researchers from many fields. The volatile nature of the stock market makes it really difficult to apply simple time-series or regression techniques. Financial institutions and active traders have created various proprietary models to beat the market for themselves or their clients, but rarely did anyone achieve consistently higher than the average returns on investment. The challenge of stock market price forecasting is so appealing because an improvement of just a few points of percentage can increase the profit by millions of dollars. This paper discusses the application of Support Vector Machines and Linear Regression in detail along with the pros and cons of the given methods.

The paper introduces the parameters and variables which can be used to recognize the patterns in stock prices which can be helpful in future stock prediction and how boosting can be integrated with various other machine learning algorithms to improve the accuracy of our prediction systems.

Literature Survey

Since the introduction of the Stock Market so many predictors are constantly trying to predict stock values using different Machine Learning algorithms such as Support Vector Regressor (SVR), Linear Regression (LR), Support Vector Machine (SVM), Neural Networks Genetic Algorithms, and many more on stocks of various companies. There is a diversity in many papers based on different parameters. Many different ML algorithms are used by different authors based on different parameters. Some authors believe that Neural Networks have given better performance as compared to other approaches . Like, in paper Hiransha M and GopalKrishnan E. A has trained four models Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) and it was observed that CNN has performed better than the other three networks. On the other hand, many authors believe that Support Vector Regression which is known to solve regression and prediction problems gives better performance as seen in paper by Haiqin Yang, Laiwan Chan, and Irwin King. In paper Paul d. Yoo has trained 3 models Support Vector Machine, CaseBased Reasoning classifier (CBR), and Neural Networks (NN) from which Neural has given the most appropriate prediction. Sumeet et al has done an approach where they have combined two distinct fields for stock exchange analysis. It merges price prediction based on real time data as well as historical data with news analysis. In this paper LSTM(Long Short-Term Memory) is used for prediction. The datasets are collected from large sets of business news in which relevant and live data information is present. Then the results of both analyses are combined to form a response which helps visualize recommendation for future increases. So, in many papers, it has been seen that neural networks give the expected prediction value.

METHODOLOGY

Define Objective: Clearly outline the goal of your project. Are you aiming to predict short-term or long-term stock movements? Are you interested in specific stocks, sectors, or the market as a whole?

Data Collection: Gather historical stock market data including prices, trading volumes, financial statements, economic indicators, news sentiment, and any other relevant data sources. You can use APIs, web scraping, or financial databases for this purpose.

Data Preprocessing: Clean the data to handle missing values, outliers, and inconsistencies. Perform feature engineering to create meaningful predictors for your model. This might involve calculating moving averages, technical indicators, or sentiment scores from news articles.

Exploratory Data Analysis(EDA): Explore the data to gain insights and identify patterns. Visualize the data using plots and charts to understand relationships between variables and how they correlate with stock prices.

Feature Selection: Determine which features are most relevant for predicting stock prices. You can use techniques like correlation analysis, feature importance from machine learning models, or domain knowledge to select the most informative features.

Model Selection: Choose appropriate models for prediction. This could range from traditional statistical models like linear regression and ARIMA to more complex machine learning algorithms like random forests, gradient boosting, or neural networks.

Model Training: Split the data into training and testing sets. Train your models on the training data and tune hyperparameters using techniques like cross-validation to optimize performance.

Model Evaluation: Evaluate the performance of your models using appropriate metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), or accuracy scores. Compare the performance of different models to select the best one.

Risk Management: Understand and manage the risks associated with stock market prediction. No model can predict the market with 100% accuracy, so it's important to have risk mitigation strategies in place.

SYSTEM DESIGN

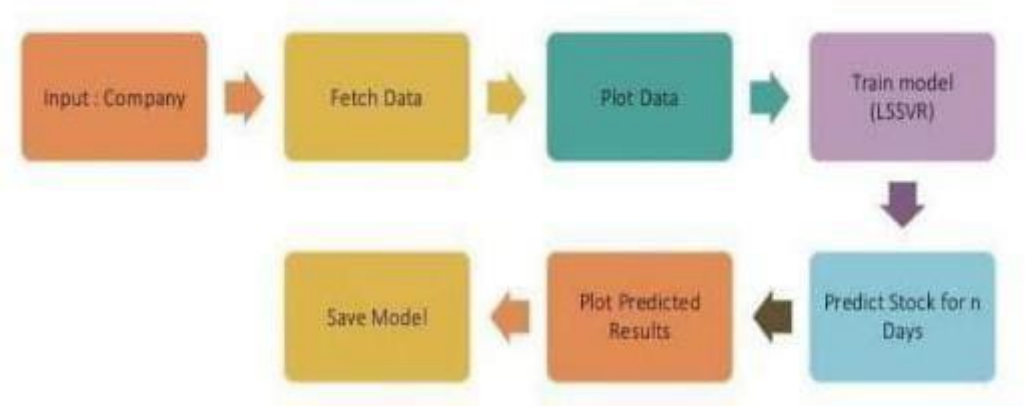

Fig no.1 Data Flow Diagram

SOFTWARE REQUIREMENT

Python: Install Python, preferably the latest version available at the time of your development.

Streamlit: Streamlit is a popular Python library for building interactive web applications.

Data Collection and Processing libraries:

Pandas: For data manipulation and analysis.

NumPy: For numerical computing.

Requests: For making HTTP requests to fetch data from APIs.

yfinance: For fetching historical stock data from Yahoo Finance.

Data visualization libraries:

Matplotlib: For basic plotting.

Plotly: For interactive plots.

Seaborn: For statistical data visualization.

Development Environment:

Choose an IDE such as PyCharm, Visual Studio Code, or Jupyter Notebook for coding.

Web brower : Make sure you have a modern web browser installed to view and interact with the Streamlit web application.

WORKING

Data Collection: Utilize financial data APIs like Alpha Vantage or Yahoo Finance to fetch historical stock market data based on the selected company's ticker symbol.

Prompt the user to input the ticker symbol, start date, and end date to specify the desired data range.

Data Preprocessing: Clean the retrieved data, handle missing values, and ensure the data is in a suitable format for modeling.

Calculate daily returns or other relevant features that might influence stock prices.

Model Selection: Choose the SARIMA model for time series forecasting, considering its ability to capture both autoregressive and seasonal components in the data.

Parameter Selection: Allow the user to select SARIMA model parameters such as order (p, d, q) and seasonal order (P, D, Q, s).

Provide default values or suggest parameter ranges based on best practices or historical analysis.

Model Training: Train the SARIMA model using the selected parameters and the historical stock market data.

Optionally, implement cross-validation techniques to validate the model's performance and fine-tune parameters if necessary.

Prediction: Allow the user to input additional parameters for forecasting, such as the number of periods ahead for prediction.

Use the trained SARIMA model to predict future stock prices based on the specified parameters.

Visualization: Visualize the predicted stock prices alongside the actual prices and historical trends.

Provide interactive charts or graphs to allow users to explore the data and model predictions easily.

User Interface: Design a user-friendly interface where users can input the ticker symbol, date range, model parameters, and prediction settings.

Provide feedback and guidance to users throughout the process, including input validation and error handling.

FUTURE SCOPE

Advanced Machine Learning Techniques: Continued advancements in machine learning algorithms, including deep learning models like recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformers, could lead to more accurate and robust predictions. These models can capture complex patterns and dependencies in financial data more effectively.

Risk Management and Portfolio Optimization: Enhanced prediction models could enable better risk management strategies by identifying potential market risks and optimizing portfolio allocations to minimize losses and maximize returns.

CONCLUSION

By measuring the accuracy of the Linear Regression algorithms, we found that the most suitable algorithm for predicting the market price of a stock based on various data points from the historical data. The algorithm will be a great asset for brokers and investors for investing money in the stock market since it is trained on a huge collection of historical data and has been chosen after being tested on a sample data. The project demonstrates the machine learning model to predict the stock value with more accuracy as compared to previously implemented machine learning models.

REFERENCE

1. Survey of stock market prediction using machine learning approach, Ashish Sharma ; Dinesh Bhuriya ; Upendra Singh 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)

2. Short-term prediction for opening price of stock market based on self-adapting variant PSO-Elman neural network,Ze Zhang ; Yongjun Shen ; Guidong Zhang ; Yongqiang Song ; Yan Zhu, 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)

3. Combining of random forest estimates using LSboost for stock market index prediction,Nonita Sharma ; Akanksha Juneja,2017 2nd International Conference for Convergence in Technology (I2CT) 4. Using social media mining technology to assist in price prediction of stock market,Yaojun Wang ; Yaoqing Wang,2016 IEEE International Conference on Big Data Analysis (ICBDA)