INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Predicting a Small Cap Company Stock Price using Python with Best Accuracy Rate: How the Data Science Working for Predictions and Accuracy Rate

G Shankarlingam[1], K Thirupathi Reddy[1]*

1 Chaitanya Deemed to be University, Kishanpura, Warangal, 506009, Telangana, India

OPEN ACCESS

Received: 04-11-2023

Accepted: 13-11-2023

Published: 30-12-2023

Citation: Shankarlingam G, Reddy KT (2023) Predicting a Small Cap Company Stock Price using Python with Best Accuracy Rate: How the Data Science Working for Predictions and Accuracy Rate. Indian Journal of Science and Technology 16(48): 4620-4623. https://doi.org/10.17485/IJST/v16i48.2793

* **Corresponding author**.

kandadithirupathi@chaitanya.edu.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objective: Predicting a smallcap company stock price with best accuracy using data science and machine learning techniques. **Method:** A Dataset of one year data collected from different sources like paytm money and yahoo finance. The prediction if a stock price involves like data collection, data pre processing, testing, training, fitting an algorithm and then prediction of a stock price finally find the best accuracy using machine learning techniques. **Findings:** This model get the accuracy of 98%. Lstm algorithm is giving the high accuracy rate than all other algorithms. We can predict the stock price using some other algorithms also. But till now the best accuracy rate given by LSTM algorithm only. **Novelty:** This study contributes to finding out the best algorithm to predict the stock price in real time market conditions. Sometimes we may not predict the stock price correctly due to international market conditions like war situations. In this case total index will show in negative only. This is the real challenge in future predictions of stock price. Ultimately predicting a stock price in all conditions and market situations is a great challenge till now. To make it success we have to introduce a new algorithm by making some changes in previous algorithms according to the recent trends markets.

Keywords: Numpy; Matplotlib; RNN; Machine learning; Small cap; Stock price

1 Introduction

Predicting stock prices is one of the most interesting and difficult problems in finance. Precisely predicting the future values of stocks can offer traders, investors, and financial analysts insightful information that will help them manage risk and make wise decisions. The methods and instruments available for predicting stock prices have changed dramatically since the development of data science and machine learning [(1)]. The LSTM algorithm has evolved into a useful tool for stock market categorization tasks, allowing researchers to extract relevant information from historical data. The goal of this work is to explore the application of ARIMA in stock price prediction in order to develop an accurate and efficient automated system for categorizing stock price prediction based on their prior data. This study will assist in gathering information from

several sources and performing performance analysis to evaluate the model's capacity to categorize different stock price forecasts in the stock market. The results could be useful for researching stock price analysis, making forecasts about stock prices, and developing intelligent systems that can make stock price predictions automatically [2]. The accuracy of the LSTM (Long short term memory) and ARIMA (autoregressive integrated moving average) models is 98% higher than that of their competitors, indicating the potential of LSTM's stock market analysis and forecasts.

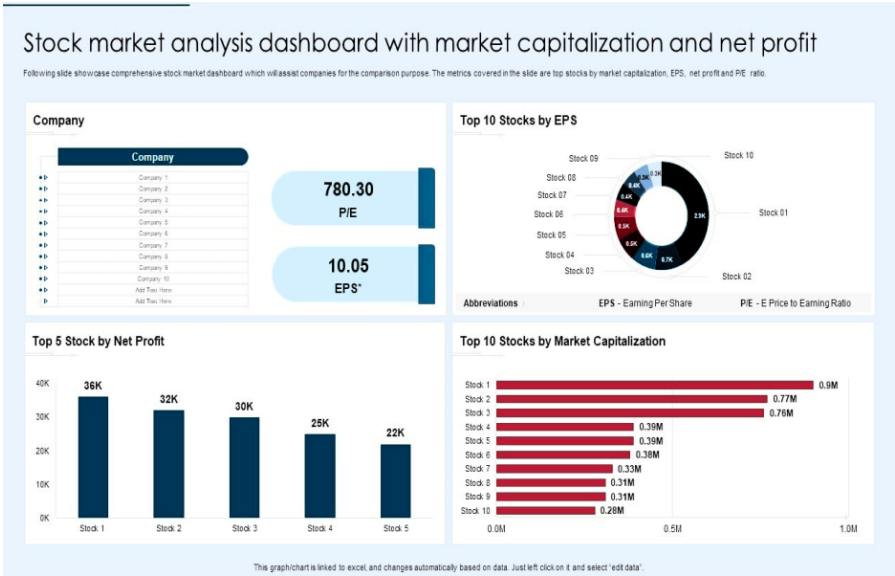

Fig 1. Stock market analysis graph

2 Methodology

2.1 Importing Libraries

You'll need to import a number of libraries to aid in data manipulation, analysis, visualization, and machine learning when working with stock price prediction in Python [3]. The following is a list of popular libraries for predicting stock prices.

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import yfinance as yf
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error
import talib
import nltk
import statsmodels.api as sm
import talib
import tensorflow as tf
from tensorflow import keras
```

Fig 2. Importing libraries in python language

2.2 Preparing dataset

A number of crucial procedures must be followed in order to prepare a dataset for stock price prediction, including data collection, cleaning, feature engineering, and division of the data into training and testing sets. This is a detailed how-to for getting your dataset ready for stock price prediction. To deal with outliers and missing values, clean up the dataset [4], addressing missing values by interpolating, backfilling, or forward-filling them with suitable values, eliminating redundant data entries if required and managing data outliers.

2.3 Fitting an Algorithm

Choosing a suitable machine learning or deep learning model, training it on your preprocessed dataset, and utilizing it for predictions are the steps involved in fitting an algorithm to predict stock prices. This is a general guide to stock price prediction algorithm fitting.

- **Choosing a Predictive Model:** Select a deep learning or machine learning model appropriate for forecasting time series. Typical models used to forecast stock prices are as follows. An established time series model that is capable of capturing linear dependencies in the data is the autoregressive integrated moving average (ARIMA).
- **Develop the Model:** Divide your data into training and validation sets, then use the training data to train the chosen
- **Test the Model:** Following successful model training and validation, evaluate the model's output using a test set of unidentified data. To evaluate how well the model generalizes to new data, this step is crucial.

3 Results and Discussions

LSTM and ARIMA based prediction model technologies in stock markets have ushered in a new era of sustainable and profitable investing methods. Thanks to these tools, investors may now detect losses early on and take proactive measures to protect their portfolios and make well-informed decisions. With the ability to increase earnings, decrease losses, and enhance financial viability, the LSTM and ARIMA based stock prediction model offers a ground-breaking outcome that holds great promise for the future of stock markets. It is critical to recognize the constraints on our research. The dataset's restriction to a single industry-specific business may have hampered the model's capacity to generalize. Future research should enlarge the dataset to include a wider range of stock values from various industries in order to get around this. Here, Figures 3 and 4 show, respectively, the analysis and accuracy for our ARIMA model and LSTM model.

0.0	0.95	1.00	0.98	93534
1.0	0.59	0.01	0.03	4658
accuracy			0.95	98192
macro avg	0.77	0.51	0.50	98192
weighted avg	0.94	0.95	0.93	98192

Fig 3. Accuracy rate for predicted stock price

Natural disasters and changes in the state of the global market could also have an impact on the model's performance. The robustness of the model could be improved by adding augmented stock price data to the dataset and investigating methods like ensemble methods and ARIMA learning. With an accuracy rate of 98%, our study shows how well the LSTM and ARIMA model work for plant stock market analysis [5]. The outcomes outperform earlier methods and demonstrate the potential of automated systems for identifying stock market momentum. Subsequent investigations have to concentrate on mitigating the detected constraints and augmenting the efficacy of the model. Figure 3 lists a summary of some stock price classes based on various models, datasets, and accuracy [6]. A suitable model that can more accurately and more thoroughly classify stock price predictions is required in light of the short selected survey.

Fig 4. Stock price analysis report and prediction report with colors

4 Conclusion

With the use of LSTM and ARIMA models, intelligent prediction has demonstrated significant promise as a means of overcoming the challenges associated with traditional techniques to stock price analysis prediction. These technologies have the ability to completely transform stock market prediction by providing accurate and timely detection of stock price forecast. The recommended automated stock market forecast is feasible with a high degree of accuracy and minimal use of technology. By leveraging artificial intelligence, this system offers a high degree of accuracy and helps minimize loss in stock value identification [7]. Investors can slow the spread of losses in their portfolio by taking timely preventative action when losses are identified early.

The commercial viability of these LSTM and ARIMA model prediction systems is very advantageous to investors. By employing these technologies, investors can increase yield while consuming less resources. For example, companies like as Motilal Oswal in India use networked sensors and machine learning to forecast stock prices for investors. Studies have shown that prediction model systems based on LSTM and ARIMA have above 98% accuracy rates when it comes to stock price predictions. This degree of foresight can significantly improve decision-making processes, guaranteeing that the right actions are taken to reduce losses and maximise revenues. Profit increases of up to 20% have been observed by investors who have employed prediction model systems based on LSTM and ARIMA. A more successful and long-lasting stock market sector is the outcome of these cost savings and higher returns.

The application of LSTM and ARIMA based prediction model technologies in stock markets has ushered in a new era of profitable and sustainable investing strategies [8]. With the use of these tools, investors may now detect losses at an early stage, protecting their portfolios and helping them to make informed decisions. The innovative outcome of the LSTM and ARIMA based stock prediction model holds great potential for the stock markets' future, offering opportunities to increase earnings, decrease losses, and enhance financial viability.

References

1) Allaire JJ, Xie Y, Mcpherson J, Luraschi J, Ushey K, Atkins A, et al. rmarkdown: Dynamic Documents for R. 2023. Available from: https://CRAN.R-project.org/package=rmarkdown.
2) Baumer BS, Kaplan DT, Horton NJ. Modern Data Science with R. 2nd ed. CRC Press. 2021. Available from: https://mdsr-book.github.io/mdsr2e/.
3) Chattopadhyay S, Prasad I, Henley AZ, Sarma A, Barik T. What's Wrong with Computational Notebooks? Pain Points, Needs, and Design Opportunities. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020;p. 1–12. Available from: https://doi.org/10.1145/3313831.3376729.
4) Depratti R. Jupyter Notebooks versus a Textbook in a Big Data Course. *Journal of Computing Sciences in Colleges*. 2020;35(8):208–220. Available from: https://dl.acm.org/doi/abs/10.5555/3417639.3417658.
5) Koenzen A, Ernst N, Storey MA. Code Duplication and Reuse in Jupyter Notebooks. 2020. Available from: https://doi.org/10.48550/arXiv.2005.13709.
6) Xiao D, Su J. Research on Stock Price Time Series Prediction Based on Deep Learning and Autoregressive Integrated Moving Average. *Scientific Programming*. 2022;2022:1–12. Available from: https://doi.org/10.1155/2022/4758698.
7) Shen J, Shafiq MO. Short-term stock market price trend prediction using a comprehensive deep learning system. *Journal of Big Data*. 2020;7(1):1–33. Available from: https://doi.org/10.1186/s40537-020-00333-6.
8) Sonkavde G, Dharrao DS, Bongale AM, Deokate ST, Doreswamy D, Bhat SK. Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications. *International Journal of Financial Studies*. 2023;11(3):1–22. Available from: https://doi.org/10.3390/ijfs11030094.