Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/15/20230846

The stock price forecast under the failure of silicon valley bank based on the ARIMA model

Dengyi Gu

The department of Statistics, The University of California, Santa Barbara, 93117, The United State

dengyi@ucsb.edu

Abstract. The collapse of Silicon Valley Bank on March 10, 2023, had a profound impact on the stock prices of many companies in the United States. This study aims to examine the response of other banks in the US to this event by utilizing the Autoregressive Integrated Moving Average (ARIMA) model to forecast their stock prices. The research demonstrates that the ARIMA model effectively predicts the general trend of these banks' stock prices, with Root Mean Squared Error (RMSE) values below 1 for four out of six major US banks. These findings indicate that the proposed method is a promising tool for managing sudden fluctuations in stock prices, outperforming traditional linear regression models. Consequently, this research provides valuable insights for investors and financial institutions in managing and mitigating risks associated with abrupt market changes. Additionally, the study contributes to a greater understanding of the effects of bank collapses on the stock market. Overall, the research highlights the significance of incorporating advanced forecasting methods, such as ARIMA, in analyzing and predicting stock price movements in volatile market conditions.

Keywords: machine learning, ARIMA, stock price forecast.

1. Introduction

Silicon Valley Bank's long-term Treasury bond portfolio suffered losses as the Federal Reserve raised interest rates to combat inflation. Customers withdrew large amounts of their funds, prompting the bank to sell $21 billion in securities, borrow $15 billion, and sell Treasury stock. However, due to warnings from prominent investors, customers withdrew $42 billion the next day, leading to a bank run. On March 10, 2023, the 16th largest bank in the U.S., Silicon Valley Bank, collapsed, making it the second-largest bank failure in U.S. history [1-3]. As a result of the Silicon Valley Bank crisis, the financial sector was hit hard. Large technology stocks, such as Nvidia, Apple, Microsoft, and Amazon shares have suffered varying degrees of decline. In addition, the stocks of the six major U.S. banks also suffered significant declines. Goldman Sachs Bank shares even fell 4.22% [2]. In this case, the price volatility of the stock is remarkebly high. If market participants can accurately predict the direction of stock prices, this will allow them to consistently earn higher risk-adjusted returns than the market [4]. Therefore, it is crucial to make accurate predictions for such highly volatile stocks.

The term "machine learning" was first introduced by Arthur Samuel, a prominent figure in the field of computer games and artificial intelligence, who was an employee of IBM. [5, 6]. It is a field within computer science and artificial intelligence that aims to imitate human learning processes and progressively improve accuracy through the use of data and algorithms [7]. Machine learning encompasses a variety of disciplines, including statistics, probability theory, convex analysis, computational complexity theory and approximation theory. Machine learning algorithms are designed to automatically analyze data, detect patterns, and make predictions about previously unknown data. It has found widespread applications in diverse areas, such as data mining, natural language processing, and stock price prediction [8]. Many different machine learning algorithms have been proposed and utilized in the field of stock price prediction, such as Linear Regression Model (LR), Random Forest Model (RF), and Neural Network Model (NN) [4]. For instance, Kim et al. (2020) conducted a study to evaluate the effectiveness of incorporating effective transfer entropy (ETE) with popular machine learning techniques like LR, Multilayer Perceptron (MLP), RF, Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM) for predicting the direction of the S&P 500 index. [4]. The majority of previous studies have relied on statistical time series methods based on historical data to predict stock prices and returns. Some of the widely used techniques in stock price prediction include the Autoregressive Conditional Heteroskedasticity (ARCH) model, Autoregressive Integrated Moving Average (ARIMA) models, Moving Average (MA) models, Autoregressive Moving Average (ARMA) models, Kalman filter, and Exponential Smoothing methods [4]. However, there are no models in previous studies that predict the stock prices of other banks for the sudden event of Silicon Valley Bank's collapse.

To address this research gap, the present study employs models to predict the stock prices of various banks during the Silicon Valley Bank collapse and conducts a thorough analysis of the experimental results. To more accurately predict sudden stock price fluctuations in a short period, this paper will use the ARIMA model to predict the stock prices of other banks. In summary, the findings and recommendations provided in this paper can be useful for investors and financial analysts in making informed decisions and managing risks. Additionally, this paper also provides further recommendations based on the research findings to better prepare for any future occurrences of similar situations.

2. Method

2.1. Dataset description and preprocessing

This study used the closing prices of the six major US banks collected from the historical data section of Yahoo Finance, as a proxy for all U.S. banks to conduct the experiment. These were Citigroup Inc., Wells Fargo & Company, JPMorgan Chase & Co., Morgan Stanley, The Goldman Sachs Group, Inc., and Bank of America Corporation, all of which play significant roles in the US financial market [9-14]. The closing price data for each bank has been recorded from February 21, 2023 through April 3, 2023, including March 10, 2023, when Silicon Valley Bank went bankrupt. Notably, the data collection process excluded weekends, when the stock market is closed weekly. Therefore, there are exactly 30 days of data available. The valid data were recorded in an Excel table and stored in CSV format, with "Date" and "Close" representing the date and closing price of the day, respectively. To ensure accuracy and consistency, six different CSV files were created for the six banks. The recorded data served as the basis for subsequent analysis and modeling aimed at predicting sudden stock price fluctuations.

2.2. Proposed approach

This study used the ARIMA model to predict the stock prices of other banks in the United States during the special period of the Silicon Valley Bank's collapse. The ARIMA model, an acronym for Autoregressive Integrated Moving Average Model, is a time series-based prediction model utilized for analyzing and forecasting data [15]. The ARIMA model assumes that future time series values are composed of past time series values and random error terms, and predicts future time series values based on this historical data. The ARIMA model is comprised of three main components: autoregression (AR), differencing (I), and moving average (MA). $AR(p)$ is a regression model that

DOI: 10.54254/2755-2721/15/20230846

represents the linear relationship between current values and past values, where p is the number of lagged observations in the model. $I(d)$ represents the differencing process applied 'd' times to the original time series data in order to achieve stationarity. $MA(q)$ is used to capture the effect of random errors on the prediction results, where "q" refers to the linear combination of the current residual value and the residual values of the past "q" time points in the ARIMA model. Combining all three types of models yields the $ARIMA(p, d, q)$ model, where "p", "d", and "q" are the three parameters of the ARIMA model, which can be combined to construct different ARIMA models. Therefore, the best ARIMA model was obtained by selecting the parameters.

The data was analyzed using Python's pmdarima library, specifically the auto_arima function, which automatically selects the optimal ARIMA model based on criteria like the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) [16]. The model was then trained and tested using the previous data, with 90% of the data used as training data and the other 10% as testing data. The Root Mean Square Error (RMSE) method was used to calculate and help determine the magnitude of the error between the predicted and true values in the test data. A linear regression model, which is also suitable for short-term prediction but weak for non-linear data, was also used as a control model, and the root mean square error method was used to calculate the error and compare it with this model. The ARIMA model is also used to forecast the short-term stock price for the next 5 days in order to provide some references and help for future investment.

3. Result and discussion

After training and testing the data using the ARIMA model, the results were obtained for the stock price forecasts of the six largest U.S. banks. Among them, Bank of America, Goldman Sachs and Morgan Stanley chose to use ARIMA(0,1,0) after selecting the parameters of the model using the auto_arima function. The other three banks have different parameters, but all of them choose "1" for parameter "d". Figures 1 through 6 show the model's predictions for the stock prices of the six largest U.S. banks, respectively. The orange line is the test price, while the green line is the model's predicted price, and the error is calculated using the RMSE method. The red line shows the subsequent 5 days' forecasts based on these data. As can be seen from the graph, the ARIMA model can predict the general direction of the subsequent stock prices very well. The RMSEs of Citibank, Wells Fargo, JP Morgan, and Bank of America are all below 1. For comparison, the RMSE of the linear regression model for Citibank's stock price prediction is above 1.5, as shown in Figure 7.

The experimental results reveal that the ARIMA model is effective in predicting the volatile stock prices that exhibit non-stationary behavior, and it can provide a rough trend direction. Conversely, linear regression models are unsuitable for predicting stock prices during periods of abrupt fluctuations, such as the instance of Silicon Valley Bank's collapse. This is because such times are typically characterized by non-linear patterns in stock prices, which are not accommodated by the linear regression's assumption of a linear relationship between data points. The ARIMA model, in contrast, can capture the inherent non-linearity in the data, rendering it a more appropriate method for predicting sudden and significant changes in stock prices. Hence, the research suggests that the ARIMA model outperforms linear regression models in these circumstances.

DOI: 10.54254/2755-2721/15/20230846

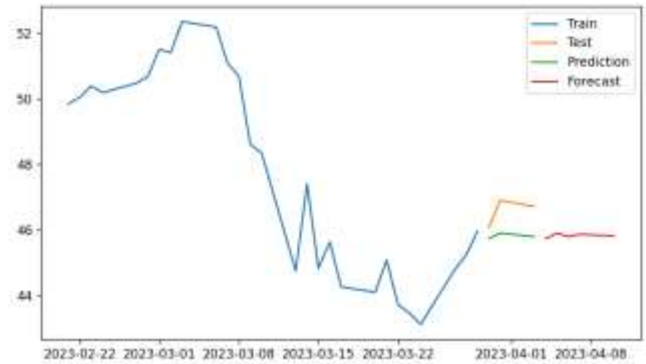

Figure 1. Figure with The Citigroup's Stock Price Forecast Chart in the ARIMA model.

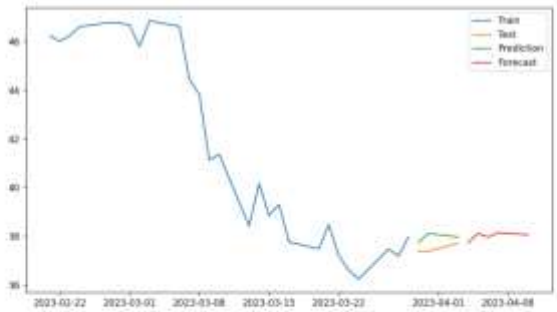

Figure 2. Figure with The Wells Fargo & Company’s Stock Price Forecast Chart in the ARIMA model.

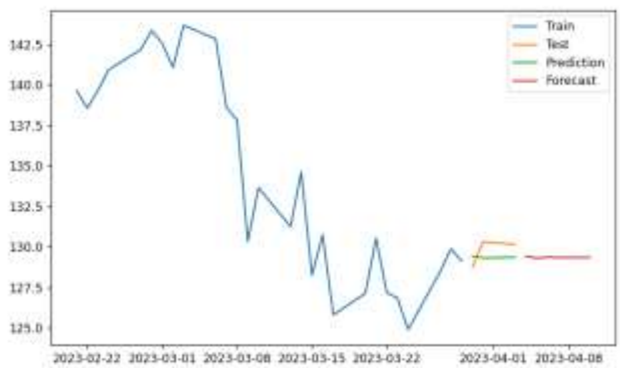

Figure 3. Figure with The JPMorgan Chase's Stock Price Forecast Chart in the ARIMA model.

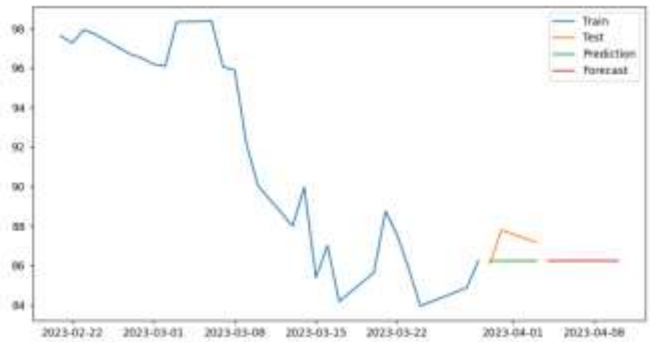

Figure 4. Figure with The Morgan Stanley's Stock Price Forecast Chart in the ARIMA model.

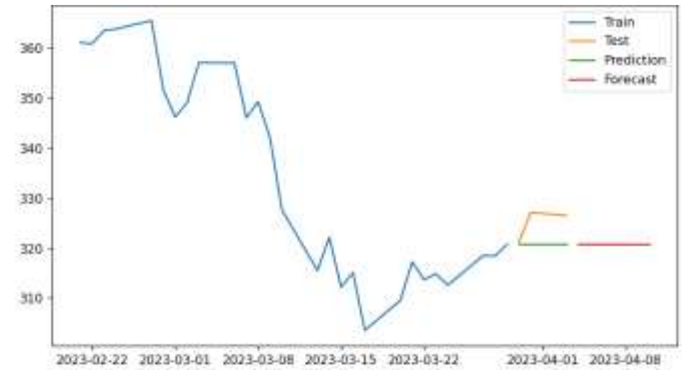

Figure 5. Figure with The Goldman Sachs Group's Stock Price Forecast Chart in the ARIMA model.

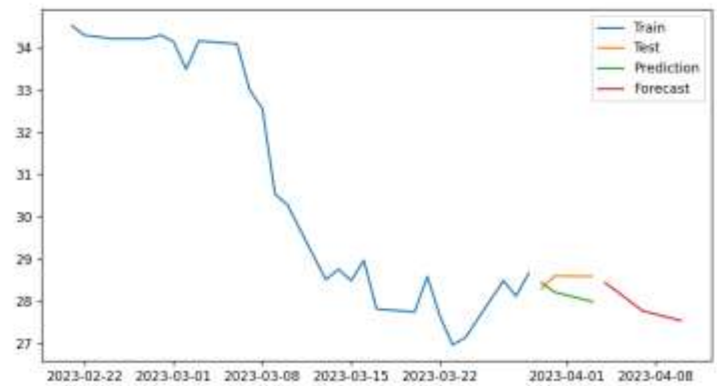

Figure 6. Figure with The Bank of America Corporation's Stock Price Forecast Chart in the ARIMA model.

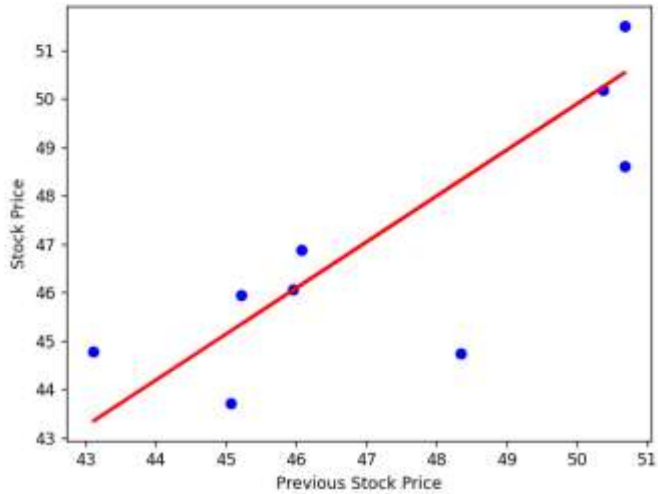

Figure 7. Figure with The Citigroup's Stock Price Forecast Chart in the Linear regression model.

4. Conclusion

This paper presents a methodology for predicting the stock prices of banks in the United States in the event of a potential failure of a bank in Silicon Valley. The study utilizes an ARIMA model that has been trained and tested using historical stock price data to make predictions. Additionally, a linear regression model is used as a comparison to demonstrate the advantages of the ARIMA model in this specific scenario. The experimental results indicate that the ARIMA model performs well in predicting the general direction of stock prices with various parameter settings. In the future, further adjustments will be made to the proposed method to reduce model errors and expand its applicability to other cases of sudden stock price changes beyond just the Silicon Valley Bank failure. Overall, this research provides insights into the potential of ARIMA models for stock price prediction in the banking sector and highlights areas for future improvement and expansion of the methodology.

References

[1] The guardian 2023 Why did Silicon Valley Bank fail? https://www.theguardian.com/us-news/2023/mar/10/silicon-valley-bank-collapse-explainer
[2] Finance Sina 2023 https://finance.sina.com.cn/wm/2023-03-11/doc-imyknara9466032.shtml
[3] Finance Sina 2023 https://finance.yahoo.com/news/silicon-valley-bank-committed-one-135059554.html
[4] Kumbure M M Lohrmann C Luukka P et al. 2022 Machine learning techniques and data for stock market forecasting: A literature review Expert Systems with Applications 116659
[5] Kohavi R and Provost F 1998 Glossary of terms Machine Learning3 vol. 30 no. 2–3 pp. 271–274
[6] Samuel A L 2000 Some studies in machine learning using the game of checkers IBM Journal of research and development 44(1.2): 206-226
[7] Github 2018 https://kangcai.github.io/2018/10/24/ml-overall-1/
[8] El Naqa I Murphy M J 2015 What is machine learning? Springer International Publishing
[9] Finance Yahoo 2023 ohttps://finance.yahoo.com/quote/C/history?p=C
[10] Finance Yahoo 2023 https://finance.yahoo.com/quote/WFC/history?p=WFC
[11] Finance Yahoo 2023 https://finance.yahoo.com/quote/JPM/history?p=JPM
[12] Finance Yahoo 2023 https://finance.yahoo.com/quote/MS/history?p=MS
[13] Finance Yahoo 2023 https://finance.yahoo.com/quote/GS/history?p=GS
[14] Finance Yahoo 2023 https://finance.yahoo.com/quote/BAC/history?p=BAC

DOI: 10.54254/2755-2721/15/20230846

[15] Capitalone 2021 https://www.capitalone.com/tech/machine-learning/understanding-arima-models/
[16] Pmdarima 2.0.3 2023 https://pypi.org/project/pmdarima/