Stock market technical analysis using Japanese candlesticks and machine learning

Milind Kolambe [1,2], Sandhya Arora [2]

[1] *Smt Kashibai Navale College of Engineering, Pune, Maharashtra, India*
[2] *Cummins College of Engineering for Women, Pune, Maharashtra, India*

Abstract

One of the thorough methods used in the stock market to assess the situation and forecast stock price movements is the charting approach. One of the most common options is the Japanese candlestick. These candles offer fascinating and recurrent shapes when plotted in a time-series graph. Each candle, whether alone or in a group, may emit a powerful or weak trading signal. History is said to repeat itself in the stock market. We are interested in applying the apriori technique to uncover item sets that are frequently found with some pattern in order to establish association rules from the time series data of Japanese candlestick shapes. By doing this, the investor will increase his or her understanding of the upcoming trend and maximize their return. We are only concentrating on candlestick patterns, therefore as the dataset gets smaller, it will take less time to make predictions about the future. We used the NSE India ten-year dataset. As part of data pre-processing, candles are encoded. Following the completion of this experiment, it was discovered that using the apriori method on Japanese candlestick shapes produced several intriguing outcomes and eventually increased forecast accuracy.

Keywords

Stock Market, Machine Learning, Apriori Algorithm, Japanese Candlesticks, Trend Prediction, Candlestick Patterns

1. Introduction

Before making any investment, technical analysis is carried out in the stock market. In every field, the investigation is essential. Suppose you want to have dinner with your friend at an unknown place and you would like to pick the best restaurant to have your dinner, you will have two choices. The first option is to taste food from all possible restaurants every day before you will have your friend with you on the pre-planned day, or another option is to pick the restaurant directly where you will find it most crowded. The second option is the best way to explain carrying out technical analysis in the stock market to make various divisions where we need to find out good opportunities according to the market trend. But actually, we do not have such a thing to find out the best. Every approach has its pros and cons, but obviously, technical analysis saves time. Commonly used types of charts used in technical analysis are line charts, bar charts, Japanese Candles.

The Japanese technique of price charting consist of a shape which looks like a candle when presented in the charts. One of the most common ways for current traders who use technical tools to view charts is using the Japanese Candlestick method. This is how the candlestick operates. Each candlestick displays the price at its open, high, close, and low price points for a specific timeframe. The candles' length displays the price range's high and low price points. The thickness of the area represents the price spread across the price when market opens and the price when market closes. The candlestick rectangle turns white if the price when market closes is higher than the price when market opens. This is indicative

ACM-2022: Algorithms Computing and Mathematics Conference, August 29 – 30, 2022, Chennai, India.
EMAIL: milind.kolambe@cumminscollege.in (Milind Kolambe)
ORCID: 0000-0003-0688-0494 (Milind Kolambe)

of a bullish stock market sentiment. The thick part of the candlestick turns black if the price when market closes is less than the price when market opens. This suggests that the market is experiencing a bad mood.

The candlesticks come in a variety of shapes, sizes, and patterns, and they all have distinctive names. For instance, the "Hammer" candlestick pattern is one of the most basic and well-known shape. In this case, the candle has a long, thicker bottom portion and a shorter, thinner top portion. The Hammer is a sort of bullish reversal candlestick that consists of just one candle. All such types of candles suggest some expected movement in the stock market, though noise and trends components in the stock market always have different degrees of predictability [1].

We may use charts to display this information in the most accessible manner since we are aware that the data recorded about the stock price values i.e. price at which market opens, price at which market closes, the highest price and the lowest price are the finest choices [2, 3] for the trading action for the given period. Each trading day has four data points. We can call it the OHLC. For example, if we prepare a chart for five days, we have to plot 20 points. It gives us an idea to go for some long period calculation.

Though it is said that time series in stock market is very complicated if we compare it with other types of series because of long duration, cyclical changes, seasonal changes and price change in irregular pattern [4], the task becomes simpler with charting. Sometime its random combination of short and long trend [5, 6]. Charting shows signals. We can see the variation of prices in the entire day as prices are continuously corrected by market participants [7]. In spite of such variations in the price, we can consider it as a signal [8]. We can definitely find same patterns determined by some factors [9]. We can predict the time series components [10]. Sometimes we couldn't obtain all the data to form candles, and this presents big problems for time series trend research [11]. Sometimes other factors affect the prediction accuracy. This is generally the result of some strong factors like some news or events [12, 13, 14, 15]. It would be uncertain to find out which one and that’s why we should never forget that interpreting the characteristic of time series is always difficult [16].

Most frequently used charts cannot be useful for technical analysis but one i.e. the line chart. The reason is we are interested in four data points not a single one. Some popular patterns are as follows.

- Marubozu
- The spinning top
- Doji Candlestick
- The Paper Umbrella
- Hammer
- Hanging Man
- Shooting star

2. Literature Survey

We can find a vast list of authors who have written about the modeling and research of financial sector time series in literature. The existing methods and pattern are not sufficient to give good accuracy [17]. Available systems do not achieve outstanding results for long-term prediction [18] because most of these systems did not consider effect of various factors like economics, politics, and psychology investment [19].

Lee, Kim, koh, kang [20] applied deep Q-Network and netural network to predict global stock market pricing using stock chart images. Recently Naik and Mohan [21] has proposed a stock crisis prediction method. They used hybrid feature selection algorithm. They also used the technical indicators like RSI and moving averages. Brenda and Arthur [22] used three different methods for sentiment analysis to apply on news to find impact on Brazilian stock market. They found Naive Bayes classifier

and the Multilayer Perceptron to be better than the best lexical approach. Zeynep and Ramazan [23] additionally used user comments on twitter data to predict stock market direction. Suman and Gao [24] proposed a stock ranking prediction method to maximize the profit. Sheikh, Alam and Parul [25] predicted indian stock market volatility using ARIMA model.

Let's talk about the methods of using Japanese Candlesticks. It has been used in Japan first in 18th century by a merchant in Japan who name was Homma Munehisa. Japanese candlesticks had been used for long time in Japan but in the rest of the world found to be less effective. The western world traders were clueless about it. Later on, Steve Nison discovered candlesticks, and he explained the method to the entire world. He also published a book "Japanese Candlestick Charting Techniques" which is being used by many traders. Almost all patterns are given Japanese names. In this method, the particular candle can be called as a bullish candle or bearish candle. Candle be shown as white (bullish) or black (bearish).

Very few researcher has used the Japanese candlestick chart for stock market prediction. Lin, Liu, Yang, Wu [26] used these candlestick along with technical indicators and machine learning methods to predict the market trend. Birogul, temur and kose [27] developed a candlestick based prediction system which helps to make a 'buy' or 'sell' decision. Ding and Luo [28] came up with a clustering method. The primary driving force behind this work is to provide a fresh method of prediction for analyzing stock market movement and to evaluate the predictability of the suggested candlestick model using the apriori algorithm's concept of 'confidence'.

We all know that daily 'open price', 'closing price', 'high' value, 'low' values play very important role in next day price computation in stock market. Along with the stock market index value the highest quote, lowest quote, closing quote, volume, and total amount of contracts traded and the day of contract from beginning are the parameter that decides the contract current quote. One more motivation behind this study is to provide inspiring investor for long term investment strategies to maximize the profit.

3. Proposed System

In the case of stock market Japanese candlestick, we are very optimistic to get some frequent item sets as Japanese candlestick pattern. We have already discussed the various patterns found in data across time for the stock market. Some of us can find some patterns only by doing a careful observation. But to find more interesting patterns, we should go for algorithmic way using a large dataset. It could give us more interesting frequent pattern that could give us some promising results.

To perform this experiment, we have fetched 10 years nifty time series data (2006-2015). We intentionally did not use recent data where impact of pandemic situation has been seen. This dataset is numerical data. Each row shows seven numerical values per day. As we have used Japanese candlestick data in time series manner which need only four different values of a stock, we discarded the remaining three columns from the dataset. The size of this dataset having selected four attributes per every row is less than 50 KB.

We will use the most commonly used eleven popular shapes of candles for this experiment. In figure 1, we can see few of these patterns as a general example of Japanese candlestick patterns.

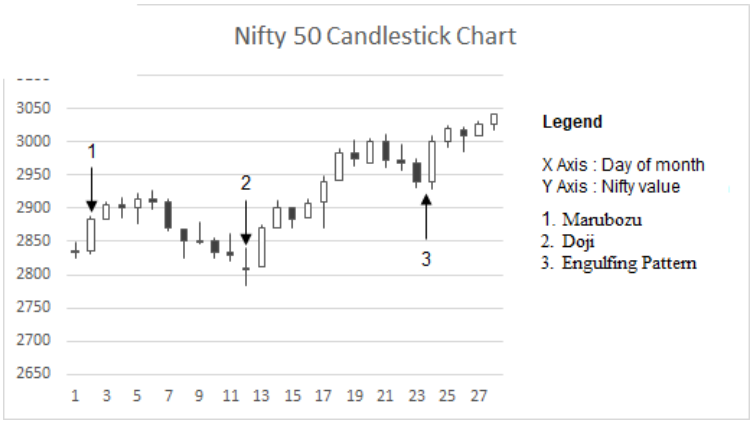

Figure 1: Japanese candle-stick patterns

There are more than 50 such patterns or shapes available to use but we have picked widely used 11 single and multi-candle Japanese candlestick patterns is as shown below.

1. Bullish Marubozu
2. Bearish Marubozu
3. The spinning top
4. Doji
5. Hammer
6. Hanging man
7. Shooting Star
8. Engulfing pattern
9. Harami
10. Morning star
11. Evening star

In NSE INDIA, one of widely used market segment is future and options. We can say that almost every stock market investor invest in this segment after attaining sufficient knowledge of stock market. This segment has some contracts to buy or sell. These contracts have different life time or expiry period. The oldest and commonly used type is monthly contract. It indicates that this contract expires one month after it becomes tradable. Monthly contracts expire on last Thursday of every month. We believe that this expiry date can have influence on support and demand ratio. Considering this assumption we have selected the length of every item set for a month. This item set will be considered up to this date and is used in apriori algorithm. Every row shows the number of patterns discovered in this month. Here we are reducing the size of dataset by encoding them as categorical value as “Y” and “N” at appropriate place of the row when the particular Japanese candlestick shape is found in the given item set. Note that all these entries are added to the new dataset thus reducing the data set size (10 times smaller) as data items. Then we applied apriori algorithm to find frequent item sets as shown in figure 2. We can also add our own patterns to increase the accuracy as shown in the figure 2.

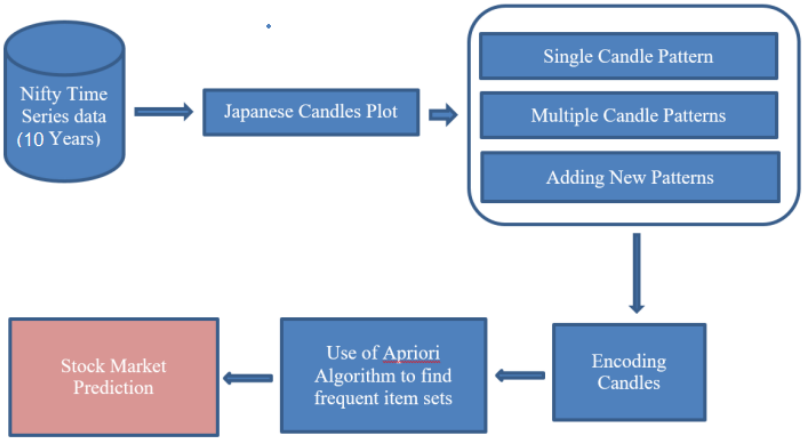

Figure 2: Proposed System

Now we can predict the expected candle pattern to guide stock market action in order to make decision which will be more profitable. For every item set, we tried to find out if the occurrences of some patterns have any association rules. Results shows that we can clearly see some rules as frequent item sets e.g. we can find some frequent item sets which indicates that bullish trend followed by indecision situation and then bearish trend followed by one more bearish indicator confirming the long bearish trend as a market correction.

4. Results and Discussion

The rules discovered on the application of Apriori algorithm on the historical data of ten years are as follows.

- Bearish Marubozu, Hammer >> Bullish Marubozu (Confidence = 0.81)

In Jan 2006 we can see Bearish Marubozu with Hammer. Bullish Marubozu in the next week as shown in figure 3. This pattern is repeated frequently with confidence= 0.81

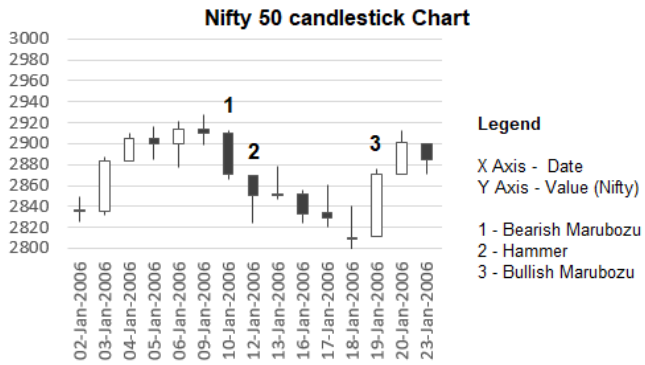

Figure 3: Prediction of bullish Marubozu

- Bullish Engulfing, doji >> Hammer (Confidence = 0.8)

In May 2007, we can see this pattern as shown in figure 4 which is repeated frequently with the confidence = 0.8

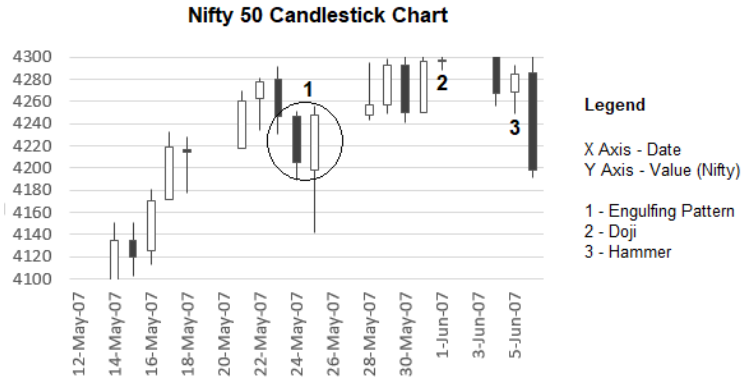

Figure 4: Prediction of Hammer

- Morning Star, Harami >> Marubozu (Confidence = 0.68)

In February 2009 we can see this pattern as shown in figure 5. It is again found to be repeated frequently with confidence = 0.68

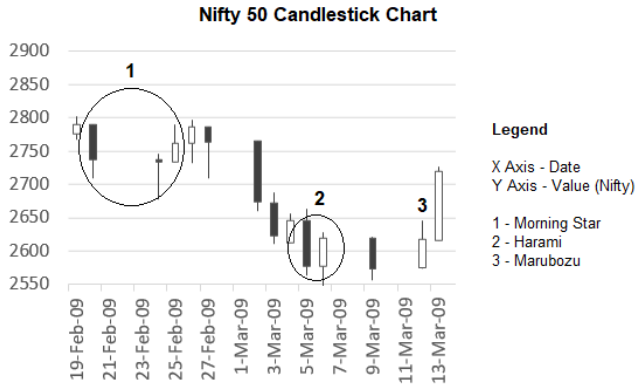

Figure 5: Prediction of bullish Marubozu

Application of this algorithm on large dataset has given us more interesting frequent pattern that could give us some promising results. The discovery of these rules will be helpful for the investors in different ways.

5. Conclusion and Future Scope

In this study, a method is proposed to identify frequently found patterns of Japanese candlestick forms using time series data of Japanese candlestick charts and the apriori algorithm. Results unmistakably demonstrate that there are certain useful association rules between different candlestick patterns. The investor will be able to foresee the trend's direction as well as have the chance to increase their profit margin as a result. Additionally, the dataset size is far smaller than the algorithms that use the raw time series data, and the prediction algorithm will operate more quickly as a result. Although the apriori algorithm has produced acceptable results for identifying frequent item sets, it has some drawbacks. As a result, future research will focus on using alternative learning algorithms, such as neural networks [29] and support vector machines [30], to predict stock market trends using candlestick patterns. Additionally, the choice of item set would be optimized in accordance with the numerous

elements influencing the index value. The application of this methodology to other stock market segments, such as stocks, mutual funds, etc., is another potential future application.

6. References

[1] E. Maiorino, F. M. Bianchi, L. Livi, A. Rizzi, and A. Sadeghian, ‘‘Data driven detrending of nonstationary fractal time series with echo state networks,’’ Information Sciences, vols. 382–383, pp. 359–373, Mar. 2017.

[2] W. Bao, J. Yue, and Y. Rao, ‘‘A deep learning framework for financial time series using stacked autoencoders and long-short term memory,’’ PLoS ONE, vol. 12, no. 7, 2017,

[3] F. Ye, Z. Liming, Z. Defu, F. Hamido, and G. Zhiguo, ‘‘A novel forecasting method based on multi-order fuzzy time series and technical analysis,’’ Information Sciences, vols. 367–368, pp. 41–57, Nov. 2016.

[4] A. K. Rout, ‘‘Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach,’’ Journal of King Saud University – Computer and Information Sciences, vol. 29, no. 4, pp. 536–552, 2017.

[5] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, ‘‘Modeling long-and short- term temporal patterns with deep neural networks,’’ in Proc. 41st Int. ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 95–104, 2018

[6] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. J. Amaratunga, ‘‘Empirical Mode Decomposition based ensemble deep learning for load demand time series forecasting,’’ Applied Soft Computing, vol. 54, pp. 246–255, May 2017.

[7] N. Nava, T. Di Matteo, and T. Aste, ‘‘Financial time series forecasting using empirical mode decomposition and support vector regression,’’ Risks, vol. 6, no. 1, p. 7, 2018.

[8] L. B. Godfrey and M. S. Gashler, ‘‘Neural decomposition of time-series data for effective generalization,’’ IEEE Transactions on Neural Network Learning Systems, vol. 29, no. 7, pp. 2973–2985, Jul. 2018.

[9] S. Jeon, B. Hong, and V. Chang, ‘‘Pattern graph tracking-based stock price prediction using big data,’’ Future Generation Computer Systems, vol. 80, pp. 171–187, Mar. 2018.

[10] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control. Hoboken, NJ, USA: Wiley, 2015.

[11] X. Dai and M. Bikdash, ‘‘Trend analysis of fragmented time series for mHealth apps: Hypothesis testing based adaptive spline filtering method with importance weighting,’’ IEEE Access, vol. 5, pp. 27767–27776, 2017.

[12] X. Ding, Y. Zhang, T. Liu, and J. Duan, ‘‘Deep learning for event-driven stock prediction,’’ in Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.

[13] X. Ding, Y. Zhang, T. Liu, and J. Duan, ‘‘Knowledge-driven event embed- ding for stock prediction,’’ in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics, pp. 2133–2142, 2016

[14] L. Zhang, C. Aggarwal, and G.-J. Qi, ‘‘Stock price prediction via discovering multi-frequency trading patterns,’’ in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2141–2149, 2017

[15] T. Fischer and C. Krauss, ‘‘Deep learning with long short-term memory networks for financial market predictions,’’ European Journal of Operational Research, vol. 120, no. 2, pp. 654–669, 2017.

[16] Z.-K. Gao, S. Li, W.-D. Dang, Y.-X. Yang, Y. Do, and C. Grebogi, ‘‘Wavelet multiresolution complex network for analyzing multivariate nonlinear time series,’’ International Journal of Bifurcation and Chaos, vol. 27, no. 8, Jul. 2017, Art. no. 1750123.

[17] Z.-K. Gao, M. Small, and J. Kurths, ‘‘Complex network analysis of time series,’’ IEEE Access, vol. 7, Pages 40220 - 40229, 2019

[18] Jui-Sheng Chou,Thi-Kha Nguyen, “Forward Forecast of Stock Price Using Sliding-Window Metaheuristic-Optimized Machine-Learning Regression,” IEEE transactions on industrial informatics, vol. 14, no. 7, july 2018, pp. 3132 – 3142

[19] Dawei cheng, Ye liu, Zhibin niu, Liqing zhang, "Modeling Similarities Among Multi-Dimensional Financial Time Series," IEEE Access, vol. 6, pp. 43404 – 43413, 2018
[20] Jinho lee, Raehyun kim, Yookyung koh, and Jaewoo kang, "Global Stock Market Prediction Based on Stock Chart Images Using Deep Q-Network", IEEE Access, Vol.7, pp.167260 – 167277, 2019.
[21] Nagaraj naik, (Graduate student member, IEEE), and Biju r. mohan, "Novel Stock Crisis Prediction Technique—A Study on Indian Stock Market", IEEE Access, Vol.9, pp. 86230 – 86242, 2021.
[22] Brenda a. januário; Arthur emanuel de o. carosia, Ana estela a. da silva, guilherme p. Coelho, "Sentiment Analysis Applied to News from the Brazilian Stock Market", IEEE Latin America Transactions, Vol. 20, pp. 512 – 518, 2022.
[23] Zeynep hilal kilimci and Ramazan duvar, "An Efficient Word Embedding and Deep Learning Based Model to Forecast the Direction of Stock Exchange Market Using Twitter and Financial News Sites: A Case of Istanbul Stock Exchange", IEEE Access, vol. 8, pp. 2169-3536, 2020.
[24] Suman saha (graduate student member, ieee), Junbin gao and Richard gerlach, ''Stock Ranking Prediction Using List-Wise Approach and Node Embedding Technique'', IEEE Access, vol. 9, pp. 88981 - 88996, 2021.
[25] Sheikh mohammad idrees, M. afshar alam and Parul agarwal, ''A Prediction Approach for Stock Market Volatility Based on Time Series Data'', IEEE Access, vol. 7, pp. 2169-3536, 2019.
[26] Yaohu lin, Shancun liu, Haijun yang (member, ieee), and Harris wu, ''Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques With a Novelty Feature Engineering Scheme'', IEEE Access, vol. 9, pp. 2169-3536, 2021
[27] Serdar birogul, Günay temür and Utku kose, ''YOLO Object Recognition Algorithm and ''Buy-Sell Decision'' Model Over 2D Candlestick Charts'', IEEE Access, vol. 8, pp. 2169-3536, 2020
[28] Ding fengqian and Luo chao, "An Adaptive Financial Trading System Using Deep Reinforcement Learning With Candlestick Decomposing Features", IEEE Access, vol. 8, pp. 2169-3536, 2020
[29] Min Wen, Ping Li, Lingfei Zhang, Yan Chen, "Stock Market Trend Prediction Using High-Order Information of Time Series," IEEE Access, vol. 7, pp. 28299 – 28308, 2019
[30] J.-S. Chou and A.-D. Pham, "Nature-inspired meta heuristic optimization in least squares support vector regression for obtaining bridge scour information," Information Sciences, vol. 399, pp. 64–80, 2017