How to Cite:
Upreti, A., Agrawal, A., Joshi, J. K., & Seniaray, S. (2022). Quantitative study of candlestick pattern & identifying candlestick patterns using deep learning for the Indian stock market. *International Journal of Health Sciences*, *6*(S3), 5739–5749. https://doi.org/10.53730/ijhs.v6nS3.7230

Quantitative study of candlestick pattern & identifying candlestick patterns using deep learning for the Indian stock market

Animesh Upreti
Undergraduate Student at Delhi Technological University
Email: animeshupreti_2k18mc015@dtu.ac.in

Ansh Agrawal
Undergraduate Student at Delhi Technological University
*Corresponding author email: anshagrawal_2k18mc018@dtu.ac.in

Jai Kumar Joshi
Undergraduate Student at Delhi Technological University
Email: jaikumarjoshi_2k18mc045@dtu.ac.in

Sumedha Seniaray
Assistant Professor at Delhi Technological University
Email: sumedhaseniaray@dtu.ac.in

***Abstract*---**The stock market is an integral aspect of any country's economic infrastructure. Analyzing and attempting to play the markets to maximize profits is an endeavor a large fraction of the population aspires to. Candlestick patterns are the backbone of Technical Analysis, used for trading in the stock market. There are a number of candlestick patterns in the market, each with its own benefits and downsides. Due to this, the task that befalls the hands of analysts is deciding which patterns provide the most effective gauge of the current market situation. Due to the large level of noise and widely recognized semi-strong form of market efficiency, analyzing and forecasting the stock market is infamously difficult. For traders that use Technical Analysis to trade, it's critical to be able to recognize candlestick patterns quickly. We will be attempting to determine their respective effectiveness with respect to the Indian Stock Market via exploratory analysis conducted on real-world market data. Also, we'll use candlestick charts to train neural networks and subsequently find patterns. Deep Learning will be used to recognize candlestick patterns in large-cap Indian equities.

International Journal of Health Sciences ISSN 2550-6978 E-ISSN 2550-696X © 2022.
Manuscript submitted: 9 March 2022, Manuscript revised: 18 April 2022, Accepted for publication: 1 May 2022

Keywords---candlesticks, NIFTY50, stop-loss, bullish reversal patterns, technical analysis.

Introduction

Introduction to Indian Stock Market

The Indian stock market is composed of a number of marketplaces and exchanges where daily transactions such as buying, selling, and issuing publicly held company shares take place. A stock exchange is a marketplace where investors can buy and sell shares of investible assets. The partial ownership of a firm is represented through stocks, also known as equities. India's two largest stock exchanges, the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE), are where the majority of trading takes place (NSE). The National Stock Market of India (NSE), which was founded in 1922 and is based in Mumbai, is India's major stock exchange for buying and selling shares of publicly traded companies. NIFTY50 is the flagship index of the NSE. The top 50 businesses by trading volume and market capitalization are included in the index. This index is frequently regarded as a barometer of the Indian capital markets by investors in India and throughout the world.

Introducing Candlestick Graphs

A candlestick pattern is a juxtaposition of candlesticks tracked over a certain time period (a day in this context); a candlestick, on the other hand, is a graphical representation of the price volatility of a certain asset during the trading day. Candlesticks with a lower opening price than the closing price are known as "white candlesticks," whereas those with a higher opening price than the closing price are known as "black candlesticks." Fig. 1 shows an example of both.

Deep Learning

Deep Learning is a type of Machine Learning that is based on the anatomy of the human brain. Deep learning algorithms use a preset logical framework to examine data and come to similar conclusions as humans. Deep learning does this by employing neural networks, which are a multi-layered structure of algorithms.

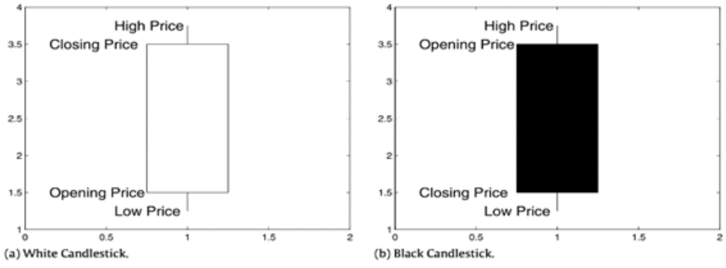
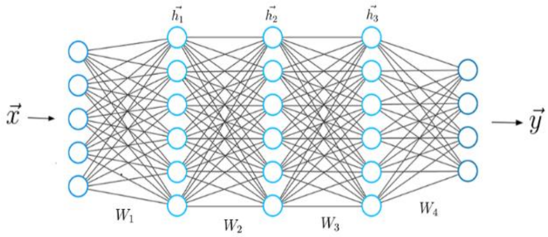
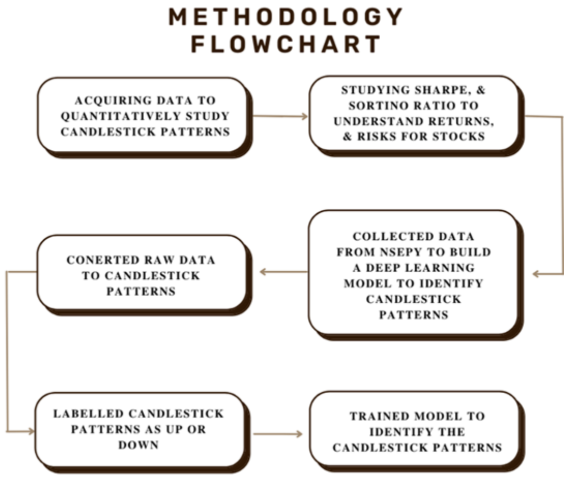


Fig. 1. Example of white candlestick and black candlestick.

Artificial Intelligence is a subset of Machine Learning, while Machine Learning is a subset of Machine Learning. Artificial intelligence (AI) is a general term for techniques that allow computers to mimic human behaviour. Machine learning, which is a set of algorithms taught on data, allows for all of this.





Libraries used in our model

- *Nsepy* - NSEpy is a library for extracting historical and real-time data from the National Stock Exchange's website.
- *Pandas* - Pandas is a Python data analysis package.
- *Numpy* is a Python module that allows you to interact with arrays.
- *Matplotlib* - Matplotlib is a library in python that is used to create static, animated, and interactive visualisations.
- *Fastai* is a computational intelligence module that provides increased components that can rapidly and readily give state-of-the-art results.

Related work

In [1] candlesticks were utilised to predict stock price movement using sentiment analysis. They proposed a novel multichannel network based on the market capitalisation of five high-demand equities (AAPL, TSLA, IBM, AMZN, and GOOG). [10] improved stock tracking decisions by using pattern recognition (PRML) from machine learning. To begin the pattern recognition schedule, 4 prominent machine learning methods and 11 different feature types are applied. In the empirical stage, they were unable to discover more complex candlestick patterns since more complex patterns necessitate larger data sets. To anticipate stock patterns, [12] utilised ensemble machine learning algorithms with candlestick charting. Kusuma et al [13] use Deep Convolutional Networks with candlestick charts to investigate the stock market's predictability. They developed a decision support framework, which traders can use to identify future stock price movement indicators. For the Taiwanese and Indonesian stock markets, their Convolutional Neural Network model produced 92.2 percent and 92.1 percent accuracy, respectively. [14] presented the GAF-CNN technique, a two-step approach to solving the pattern classification problem. The first step uses a Gramian Angular Field (GAF) to encode the time series as different sorts of pictures, and the second stage uses a Convolutional Neural Network (CNN) to learn eight distinct types of candlestick patterns from the GAF images. They were able to attain a 90.7 percent average accuracy in real-world data using this method, beating the LSTM model.

Methodology

Data Acquisition for quantitative study of Candlestick patterns

In our study, we used a large data set that included stocks included in the NIFTY50 stock index calculation. The stock price data used was from January 2000 to December 2015, and was an Expiration (EOD) form. Each data stock has provided more than 3800 data points per day.The NIFTY50 stock index is one of the leading indicators related to the Indian stock market; has 50 companies covering various industries. Since the NIFTY50 index is a dynamic index with stocks added and subtracted as the market moves, during the 16-year period selected for our research purposes, only 17 stocks all listed in the accompanying table) were considered there. to calculate the index.

Table 1
The 17 Sample Stocks

S.No	Stock Name	Industry
1.	ACC	Cement
2.	AMBUJACEM	Cement
3.	BHEL	Industrial Manufacturing
4.	CIPLA	Pharmaceuticals
5.	HDFC	Financial Services
6.	HDFCBANK	Financial Services
7.	HEROMOTOCO	Automobiles
8.	HINDALCO	Metals
9.	HINDLEVER	FMCG
10.	INFY	IT Services
11.	ITC	FMCG
12.	M&M	Automobiles
13.	RELIANCE	Energy
14.	SBIN	Financial Services
15.	TATAMOTORS	Automobiles
16.	TATA POWER	Power
17.	TATASTEEL	Metals

The daily history data of seventeen stock samples identified in a previous step is analysed in this study. On the NSE India website and CMIE Prowess, daily stock data includes OPEN, HIGH, LOW, LOW, CLOSE, and total volume data for a 16-year period. For sixteen years, each stock contains 3983 data points. Various economic events occurred throughout the sample period, including the dotcom boom and the global financial crisis of 2007-2009. Sample time also passed for national and international events. The candle scanner software collected 17 shares based on EOD data. Candles canner is a technical analysis software programme that finds, evaluates, and analyses candlestick patterns in financial market data (Table 2). Within 10 days of their occurrence, the effectiveness of candle designs is studied. Table 3 indicates how the observed patterns were classified.

Table 2 List of Bullish Reversal and Bullish continuation Patterns

Sl.No	Candlestick Pattern	No. of Candles	Forecast
1	Abandoned Baby	3	Bullish Reversal
2	Belt Hold	1	Bullish Reversal
3	Doji Star	2	Bullish Reversal
4	Engulfing	2	Bullish Reversal
5	Hammer	1	Bullish Reversal
6	Harami Cross	2	Bullish Reversal
7	Harami	2	Bullish Reversal
8	Homing Pigeon	2	Bullish Reversal
9	Inverted Hammer	2	Bullish Reversal
10	Kicking Up	2	Bullish Reversal
11	Last Engulfing Bottom	2	Bullish Reversal
12	Matching Low	2	Bullish Reversal
13	Meeting Lines	2	Bullish Reversal
14	Morning Doji Star	3	Bullish Reversal
15	Morning Star	3	Bullish Reversal
16	Piercing	2	Bullish Reversal
17	Southern Doji	1	Bullish Reversal
18	Takuri Line	1	Bullish Reversal
19	Tasuki Line	2	Bullish Reversal
20	Three Inside Up	3	Bullish Reversal
21	Three Outside Up	3	Bullish Reversal
22	Three stars in south	3	Bullish Reversal
23	Three White Soldiers	3	Bullish Reversal
24	Tri star	3	Bullish Reversal
25	Turn Up	2	Bullish Reversal
26	Tweezer Bottom	2	Bullish Reversal
27	Unique three-river bottom	3	Bullish Reversal
28	Gapping Up Doji	1	Bullish continuation
29	Rising Window	2	Bullish continuation
30	Separating lines	2	Bullish continuation
31	Side-by-side white lines	3	Bullish continuation
32	Strong Line	1	Bullish continuation
33	Upside three Gap methods	3	Bullish continuation
34	Upside Tasuki Gap	3	Bullish continuation

It's probable that some of the false positives will have disastrous consequences. Low risk is always safeguarded in some way from the perspective of traders, and this is a method or instrument known as a stop loss strategy. The danger of decrease was reduced to -3.5 percent in this study. This figure is based on the NIFTY 50 index return's substantial weekly loss over the 16-year study period. Adding to the study's stability, an indication of malpractice is also regarded a sign of loss, and those events are scored at -3.50 percent. Table 4 shows 10 candlestick patterns, while Table 5 shows comprehensive statistics. All stock samples are returned, and the four highest probable candle pattern returns are generated based on weight for False, Low, Medium, and Effective returns. In this inquiry, the Sharpe and Sortino ratings were employed in the model modelling. The Sharpe ratio is a popular risk-adjusted investment return metric. The average rate of return is more than the non-risk value per unit of flexibility. The non-hazardous rate used by Sharpe and Sortino is based on the 7.97 percent average yield on GOI dates over the previous 16 years.

Sharpe Ratio -> $(R_s - R_f) / \sigma_s$

The Sortino ratio is a metric for determining the likelihood of returns dropping below a given threshold. It is estimated with the help of

Table 3 Candle efficiency classification

Signal Type	Returns
False	- 3.5% to 0 %
Low	0 to 2.0%
Medium	2 to 3.5 %
High	Above 3.5%

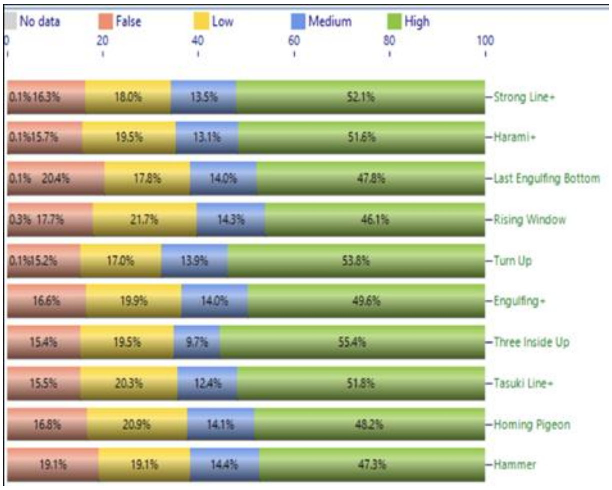


Table 5 Candle Pattern Occurrence statistics (Full sample Data)

Sl.no.	Pattern Name	Code	Candles	Pattern Type	No. of occurrences	% of the total occurrences	Average frequency (days)
1	Strong Line	SLN	1	bullish continuation	1568	18.39%	46
2	Harami	HMI	2	bullish reversal	1035	12.14%	69
3	Last Engulfing Bottom	LEB	2	bullish reversal	909	10.66%	79
4	Rising Window	RSW	2	bullish continuation	729	8.55%	98
5	Turn Up	TUP	2	bullish reversal	699	8.20%	103
6	Engulfing	ENG	2	bullish reversal	674	7.90%	106
7	Three Inside Up	TNP	3	bullish reversal	487	5.71%	147
8	Tasuki Line	TSL	2	bullish reversal	394	4.62%	182
9	Homing Pigeon	HPN	2	bullish reversal	369	4.33%	194
10	Hammer	HMR	1	bullish reversal	277	3.25%	259

Sortino Ratio -> $(R_s - R_f)/\sigma d$
Table 6 shows the Sharpe and Sortino values for each of the 17 stocks' top four(4) candle patterns (* High ranking Sharpe and Sortino measure values).

Table 6 Sharpe and Sortino values (Stock specific for top 4 candle pattern occurrences)

ACC	ENG	HMI	LEB	SLN
Sharpe Ratio	3.3886	4.3867*	1.7239	4.0138
Sortino Ratio	6.5009	8.4643*	3.7548	7.6630
AMBUJACEM	HMI	ENG	LEB	SLN
Sharpe Ratio	2.5648	4.7407	2.7788	6.3740*
Sortino Ratio	5.0852	9.7135	5.2406	11.3496*
BHEL	SLN	HMI	ENG	LEB
Sharpe Ratio	6.6962	7.9553*	2.1248	5.9954
Sortino Ratio	12.9343	14.7031*	4.5164	11.7584
CIPLA	HMI	LEB	SLN	TUP
Sharpe Ratio	1.7889	2.1485	4.7934*	1.8625
Sortino Ratio	3.8421	4.4537	8.5628*	3.9349
HDFC	ENG	HMI	LEB	SLN
Sharpe Ratio	5.9614*	2.9819	2.6704	2.4296
Sortino Ratio	11.1157*	5.6710	5.3248	4.7286
HDFCBANK	TUP	ENG	LEB	SLN
Sharpe Ratio	3.4193*	2.2171	2.8684	2.4871
Sortino Ratio	6.0894*	4.1174	5.4908	4.5965
HEROMOTOCO	ENG	HMI	LEB	SLN
Sharpe Ratio	2.9882	5.4799*	2.6143	3.7539
Sortino Ratio	6.2738	10.3783*	5.0969	7.3187
HINDALCO	HMI	LEB	RSW	SLN
Sharpe Ratio	5.8384*	3.2758	3.3993	4.1848
Sortino Ratio	11.2874*	7.2591	6.5072	7.8585
HINDLEVER	HMI	LEB	SLN	TUP
Sharpe Ratio	3.2572	4.3205	3.8758	6.0209*
Sortino Ratio	6.1274	8.5961	7.2491	9.9983*
INFY	RSW	SLN	ENG	HMI
Sharpe Ratio	3.8684	2.7156	2.3172	4.3281*
Sortino Ratio	7.9543*	5.3554	4.6394	7.1693
ITC	SLN	HMI	LEB	ENG
Sharpe Ratio	2.9513	3.3898	5.4755*	3.1934
Sortino Ratio	5.3468	5.8739	9.0120*	6.0524
M&M	HMI	LEB	SLN	TUP
Sharpe Ratio	17.2834*	16.9653	18.0790	12.3158
Sortino Ratio	45.2576*	36.7180	40.7386	29.3046
RELIANCE	ENG	HMI	LEB	SLN
Sharpe Ratio	2.8424	4.2273*	3.7425	3.8617
Sortino Ratio	5.3991	7.5125*	7.4796	7.0637
SBIN	HMI	LEB	RSW	SLN
Sharpe Ratio	4.6496	4.2080	7.0336*	7.2714
Sortino Ratio	9.0185	7.9378	11.8908	12.3392*
TATAMOTORS	HMI	RSW	SLN	TUP
Sharpe Ratio	2.4752	0.9029	7.3122*	5.7795
Sortino Ratio	4.9045	2.4988	16.6763*	13.3769
TATASTEEL	HMI	LEB	RSW	SLN
Sharpe Ratio	5.0427	3.6605	7.1665*	5.8548
Sortino Ratio	9.8269	7.7287	12.7552*	11.0395

Deep Learning Methodology
Broad approach

- First, we'll use the nsepy library to collect daily historical stock data for multiple Indian stocks.
- The data will then be converted into relevant candlestick charts.
- Labeling the Candlestick charts is vital to determine upswing or decline. We shall designate them as “Up”, or “Down” based on the percentage movement of close price in a given direction.
- Following that, the data will be utilized to build a deep learning model.
- We'll get the areas of interest (candlestick patterns) based on the value of activations and show them using heat maps.
- In this research paper, we use a deep learning model to identify different candlestick patterns in several Indian large-cap stocks. Our data set was the historical stock data which we gathered using the python library NSEpy.

We considered the following stocks:

- BAJFINANCE
- RELIANCE
- INFY
- HDFHDFCBANK
- HDFCLIFE
- ZEEL
- DELTACORP
- ITC
- ASIAN PAINT

We converted the data into candlestick charts and labeled each chart either uptrend or downtrend. On our dataset, we trained a deep learning model to predict candlestick patterns.

Code functionality
Obtain data function

It takes the ticker symbol as well as the start and end dates as input and returns a Pandas Dataframe with stock data.

Plot candles function

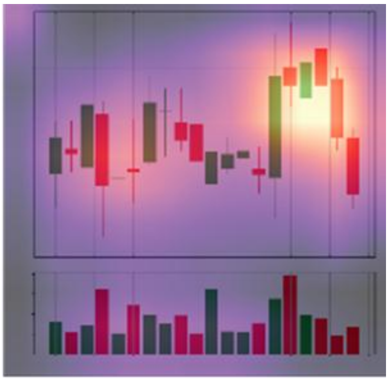
This accepts a Pandas Dataframe and produces a Candlestick chart. For the data gathering, we obtain historic data using "obtain_data" function (last 2 years) and classify the particular candlestick in "uptrend" or "downtrend" based on the movement of the next five days. We then set up the data and use it to train our deep learning model by setting a particular batch size and getting different datasets. Using transfer learning and a suitable learning rate, we trained a pre-trained neural network on our charts. Using transfer learning and a suitable learning rate, we retrain the existing network on the new larger images. We then made a list of all the charts to be examined and used activations to produce a heat map using matplotlib.

Code

Our code for the above research paper in which we've performed the execution of the aforementioned training and building of a deep learning model can be found at the Github repository below: Candlestick Identifier Repository

Results

The use of candlestick technical analysis and its effectiveness in 17 NIFTY50 Indian stocks were discussed. Technical analysis is a complicated approach of assessing price data that does not involve changing the underlying time series. According to Sharpe and Sortino, significant performance discrepancies have been observed. During the Harami research period, the top four candle patterns in 17 equities (7 stocks-ACC, BHEL, HEROMOTOCO, INFY, M&M, and RELIANCE), a Strong Line (3 stocks-AMBUJACEM, CIPLA, and TATA MOTORS), and a Growth Window (3 stocks-SBIN, TATA POWER, and TATA STEEL) demonstrated to be very trustworthy. Candlestick technical analysis can be an effective trading tool if used with the proper stop loss approach. The illuminated portion in the image below indicates a candlestick pattern.



Through our validation dataset, we were able to achieve an initial accuracy of 65%. Tweaking the code to account for a few outliers that were negatively affecting the accuracy helped us achieve a higher accuracy rate of 76.42%.

References

1. Trang-Thi Ho, Yennun Huang “Stock Price Movement Prediction Using Sentiment Analysis and CandleStick Chart Representation”
2. Nison, S, (2001) Japanese Candlestick charting techniques: a contemporary guide to the ancient investment techniques of the Far East. (*Penguin*)
3. Jensen, Michael C. "Some anomalous evidence regarding market efficiency." Journal of financial economics 6.2/3 (1978): 95-101.
4. Lo, A. W., Mamaysky, H., & Wang, J. (2000) Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*, 55(4), 1705-1765.
5. Li, H., Ng, W. W., Lee, J. W., Sun, B., & Yeung, D. S, (2008) Quantitative study on candlestick pattern for Shenzhen Stock Market. In *2008 IEEE International Conference on Systems, Man and Cybernetic*, October (pp. 54-59)
6. Duvinage, M., Mazza, P., & Petitjean, M. (2013) The intra-day performance of market timing strategies and trading systems based on Japanese candlesticks. *Quantitative Finance,* 13(7), 1059-1070.
7. Sharpe, W. F., (1994) The Sharpe Ratio, the Journal of Portfolio Management, Stanford University, Fall.
8. Gaurav Gupta, Geeta Gupta, (2014) An Analysis of Individual Investors Towards Investing in Stock Market, International Journal of Advanced Research in Management, 5(5), pp. 10-18.
9. Manoharan M and Dr Rajesh Mamilla, The Profitability of Bullish Reversal Candlestick Patterns – A Study on Select Indian Nifty50 Index Stocks, *International Journal of Management, 11(6)*, 2020, pp. 1623-1631.
10. Yaohu Lin, Shancun Liu, Haijun Yang, Harris Wu, Bingbing Jiang “Improving stock trading decisions based on pattern recognition using machine learning technology”

11. Chen, Shi, Si Bao, and Yu Zhou. "The predictive power of Japanese candlestick charting in Chinese stock market." *Physica A: Statistical Mechanics and its Applications 457 (2016)*: 148-165.
12. Yaohu Lin, Shancun Liu, Haijun Yang, Harris Wu "Stock Trend Prediction Using Candlestick Char ediction Using Candlestick Charting and Ensemble Machine Learning Techniques with a Novelty Feature Engineering scheme"
13. Rosdyana Mangir Irawan Kusuma, Trang-Thi Ho, Wei-Chun Kao, Yu-Yen Ou, Kai-Lung Hua "Using Deep Learning Neural Networks and Candlestick Chart Representation to Predict Stock Market"
14. Jun-Hao Chen, Yun-Cheng Tsai "Encoding candlesticks as images for pattern classification using convolutional neural networks"
15. Nti, Isaac Kofi, Adebayo Felix Adekoya, and Benjamin Asubam Weyori. "A systematic review of fundamental and technical analysis of stock market predictions." *Artificial Intelligence Review 53.4* (2020): 3007-3057