PREDICT AND ANALYSE STOCK DATA USING MACHINE LEARNING TECHNIQUES

S Govinda Rao
Associate Professor in CSE
GRIET,Hyderabad
E-Mail: govind.griet@gmail.com

1.ABSTRACT

Algorithmic Trading is an automated trading system in which strategies, techniques and decisions are pre-programmed and executed based on market conditions. This type of trading became a widely valuable topic in the area of financial management and in modern artificial intelligence.

Algorithmic Trading comes under the study of Artificial Intelligence, Data Science and Statistics. Machine learning models are becoming increasingly widespread in the study of algorithmic trading and investment management. Trading is a highly risky and uncertain task, where the price of stocks changes very frequently with time. Analysing the drastic changes of every stock is a nightmare for many traders. Stock market business has grown to such an extent which has never happened in history. This has become an imperative to the success of all the organisations of online trading business worldwide.

Key Words: Algorithmic Trading, Linear Regression, Random Forest Classifier, Repeated K Fold,Momentum Strategy.

2.INTRODUCTION

Trading is a process of buying or selling goods and services mostly in exchange for money. It takes place within the country or between the trading nations.

Algorithmic trading (also known as automated trading or black box trading) is a set of rules which is applied on the input data to generate buying and selling points(signals).

Algorithmic trading systems operate between two fields. From one field Trading system receives market data and, in another field, sends order requests and receives replies.The generated signals by an algorithmic trading system can be executed automated or manually. When the signals are traded automatically then the system is known as an "Automated trading system".

One major reason to use algorithmic trading works without humanintervention.It works more accurately, at speed and at reduced costs. As trading algorithms use more complex strategies to analyse markets which will provide high profits to traders. Algorithmic trading generates profits at a speed and frequency that is impossible for a human trader.

2.1.Strategies used in algorithmic trading: In algorithmic trading different types of strategies are used. Some of them are:
Momentum strategy: Momentum strategy involves different methods which have different accuracy[1].
I. Simple Moving Average Crossover:
In simple moving average crossover strategy, we plot a moving average of the stock price over a certain period of time, for example 10-day moving average, 30-day moving average and 100-day moving average. When the stock price is lower than average it's a buy signal when it moves above the average it's a sell signal.
ii. Dual moving average Crossover:
In dual moving average crossover, two moving averages are plotted in the graph, crossover of the both averages signifies the buy or sell of the stock based on the direction. Moving average crossovers are used to spot trend reversals.
iii. .Naive trading strategy[2] :

Take integer threshhold_val as parameter. Count the number of times the value increases and decreases. If the increase count exceeds the threshold value, then we will buy the stocks and if the decrease count exceeds the threshold value we will sell the stocks.

iv. Turtle strategy[3]:
Here, consider window_size days for which you want to trade. Store the low and high of window_size days and if there is a breakout in the price then buy the stocks and if there is a downfall sell the stocks.

3.LITERATURE SURVAY

i.Moving Average Crossover Strategy[4]:
Consider the average price of the market over a period, say 30 days. This will reduce the noise from the data and smoothen the graph which will help us to predict the future prices easily. If the market price moves from one side of the moving average graph indicates the change in momentum.

In this strategy the algorithm calculates the average price over the last certain days and plots the graph. And if there is a transition of the price from one side of the average price graph to the other side then we will buy/sell stocks.

Here, if the current price of the stock exceeds the average price of the stock then buy the stocks. If the price of the stocks goes below the average price then sell the stocks.

Because there are many momentum changes we use Dual Moving Average Strategy to limit the momentum changes.

ii.Smart Order Routing Model[5]:

Smart order routing is used in online trading. It is an automated process for handling orders, aimed at taking the best available opportunity throughout a range of different trading venues. It follows a set of rules for assessing trading liquidity.

Smart order routing is now a necessity as the popularity of electronic trading has grown. This allows us to execute the buying and selling of the stocks in the best possible way.

iii.Time Weighted Average Price Model[6]:

Investors define a time period in TWAP Strategy over which they want to buy shares. It will slice evenly big orders into smaller ones and execute them over a defined period. TWAP strategy works in a declared period of time.

This model considers the parameters Target Quantity, Step size, Delay, Start Time and Price Limit. From the mentioned Start time or till reaching Target Quantity frequency of strategy positions opening depends on Delay, Price Limit parameter and market price.

$$Typical\ Price = (Close + High + Low + Open)\ /\ 4$$

iv. Bollinger Band Strategy[7]:

Bollinger Band Trading Strategy used three bands namely Upper Band, Lower Band and Middle Band. Bollinger Bands are used to determine the overbought and oversold stocks in the market. When

the stock price touches the upper band, it is known as overbought and in the future it is more likely to fall down. In this we sell our stocks from the portfolio. When the stock price touches the lower band, it is known as oversold and in the future it is more likely to rise in the future. In this case we buy that particular stock.

Calculations:

UPPER BOLLINGER BAND = MOVING_AVERAGE(Price, n) + $m * \sigma$[Price, n]

LOWER BOLLINGER BAND = MOVING_AVERAGE(Price, n) − $m * \sigma$[Price, n]

Where:s

n = Number of days in smoothing period

m = Number of standard deviations

σ[Price, n] = Standard Deviation over last n periods of Price

Assuming normal distribution at the peak and trough of the stock data 92% of the data lies within the 2 standard distributions. Therefore m can be approximated to 2.

UPPER BOLLINGER BAND = MOVING_AVERAGE(Price, n) + 2 $*\sigma$[Price, n]

LOWER BOLLINGER BAND = MOVING_AVERAGE(Price, n) − 2 $*\sigma$[Price, n]

4.METHODOLOGYAND DATASET

4.1.Methodology:

For the analysis and prediction of the market price we are going to create different trading strategies based on the trend being followed and the momentum. Also we are creating a strategy for mean reversion behaviour of the market.

4.2.Momentum and Trend following Strategy[8]:

Here we are predicting the stock price with the help of the trend that is being followed over a specified period of time. If the prices of the stock have increased/decreased , say for the last 15 days, the trend is expected to continue further. So this will help in making the decisions to buy/sell the stocks for more profit.

Using this strategy we can work with the volatility and buy/sell the stock based on short-term trends and make more profits. Using this strategy there is potential for high profits for the trader who can take risk and stick to the strategy.

Using Momentum trading strategy a trader can expect high profit in the bull market but the profit shrinks in the bull market.

Advantages:

The algorithm can be easily implemented and easily understandable.

Disadvantages:

This strategy does not consider the influence of the sentiment and the noises in the market which may decrease the accuracy.

4.3.Types of momentum strategies:

i.DualMoving Average Crossover Strategy[9]:In this strategy instead of considering a single moving average, we will consider two moving averages, short-term moving average and long-term moving average. This strategy gives better accuracy. There are reduced crossovers, so we get better results.

This algorithm calculates the average price over a short period of time as well as a long period of time and plots the graph. If the short-term moving average crosses the long-term moving average in the downward direction, sell the stocks and buy if vice-versa.

But when the price is swinging back and forth it generates multiple signals and may give a wrong prediction. So in this case use another indicator.

ii.Turtle Strategy[10]:

Turtle Trading is one of the most famous trend-following strategies. In the history of trading Turtle Trading was an enormous success which was held by two famous traders Richard Dennis and William Eckhardt.

In this trading strategy we are not going to use the average values but the high and low over window_size days. If there is a breakout of price then we buy the stocks or contracts and sell if there is a retracement or price fall.

In this strategy there is a frequent buy and sell orders placed. Profits are made if the investment amount is in bulk, i.e. buy

more shares as profit per share increases. Pre-plan your entry and exit points, before buying the stocks. Average True Range(ATR) is used as a volatility indicator in technical analysis. Use ATR to calculate the volatility of the stock and use it to vary your position size.

iii.Naive Trading Strategy[11]:

In this strategy we will consider the number of times there is an increase or decrease in the price. If the count of increase or decrease in the price exceeds over a threshold value, then you can buy or sell stocks accordingly.

In this strategy we create a long signal when the price reaches the high for the last window of days. We create a short signal when the price reaches its lowest point in the window.

Entry Rule : Stock Price > Highest value in the window.

Stock Price < Lowest value in the window.

Exit Rule : Highest value in the window < Average value of the window.

Lowest value in the window > Average value in the window.

4.4.DATASET :

In algorithmic Trading algorithm we require 1 dataset which consists of 6 fields(date, open, close, low, high, volume). Every field is used by the algorithm to find the signals in the trading. Date refers to the data of the stock on that data, Open is a value that refers the stock value at start of date, Close is a value that refers the stock value at the end of the date, Low is a value that refers the lowest cost of the stock in the day and High refers to the highest cost of the stock in the day. All these fields are used to calculate the signals that indicate the trader to invest or break the stock in trading.

date	open	close	low	high	volume
04-01-2010	112.37	113.33	111.51	113.39	1.19E+08
05-01-2010	113.26	113.63	112.85	113.68	1.12E+08
06-01-2010	113.52	113.71	113.43	113.99	1.16E+08
07-01-2010	113.5	114.19	113.18	114.33	1.31E+08
08-01-2010	113.89	114.57	113.66	114.62	1.26E+08
11-01-2010	115.08	114.73	114.24	115.13	1.06E+08
12-01-2010	113.97	113.66	113.22	114.21	1.63E+08
13-01-2010	113.95	114.62	113.37	114.94	1.62E+08
14-01-2010	114.49	114.93	114.42	115.14	1.16E+08
15-01-2010	114.73	113.64	113.2	114.84	2.12E+08
19-01-2010	113.62	115.06	113.59	115.13	1.39E+08
20-01-2010	114.28	113.89	112.98	114.45	2.16E+08
21-01-2010	113.92	111.7	111.56	114.27	3.45E+08
22-01-2010	111.2	109.21	109.09	111.74	3.46E+08

IV. .RESULTS AND DISCUSSIONS

The final experimental result of the test data sets are shown in the figure below. The results are based on the stocks data of nifty 50 companies in the interval of 2010 to 2019 with an initial investment of Rs.1,00,000. The experimental results on the development and test data are consistent and shown the best returns on the investment. The following are the conclusions drawn from the results.
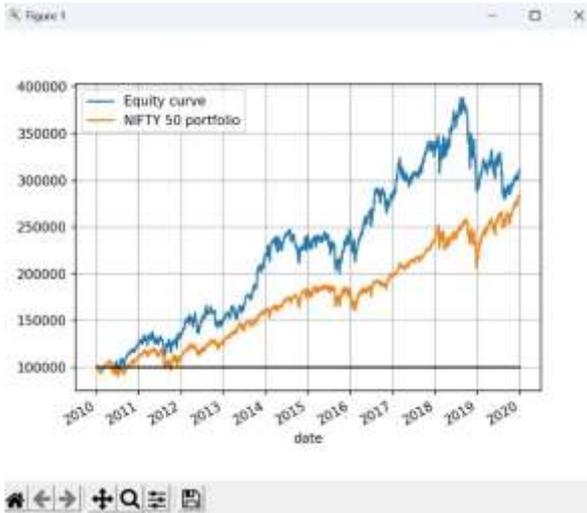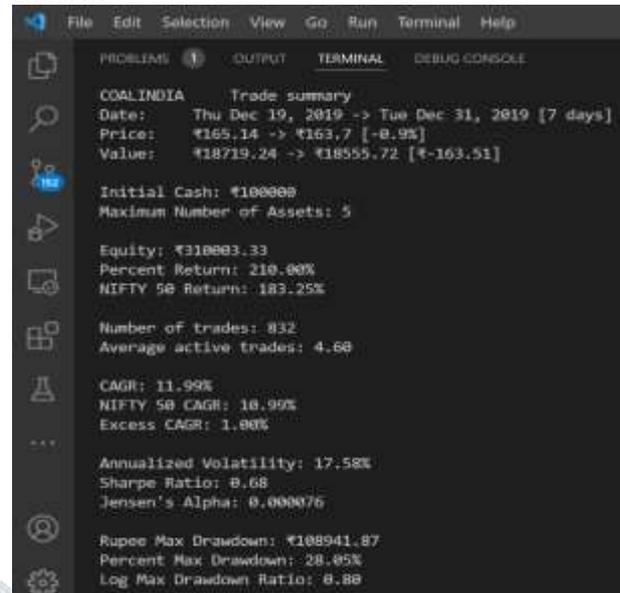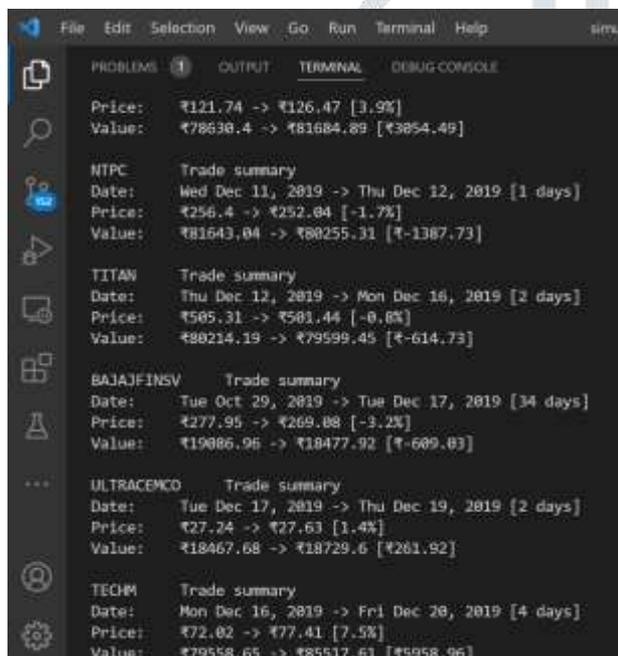
thetraders.

figure 1 shows the graph of nifty 50 investment and the equity of the amount invested by our algorithm. The amount invested in the nifty 50 portfolio increased by 183% but the same amount invested by the algorithm increased by 210%. There is a huge difference in the returns of the trader. So we can conclude that the algorithm trade works better and accurately.

The following figure shows the overall trade summary of transactions by the algorithm from 2010 to 2020. The initial investment in the trades is Rs.1,00,000 with maximum no of trades per day as 5. Final return on the investment is Rs.3,10,000 (210% return) whereas the nifty is 180% only. There are 832 transactions made in this period with 4 active trades on average.

6.CONCLUSION

We have used different strategies based on momentum and trend following to predict the prices of the stocks. Each of the strategies have varying accuracy and are suitable to certain market scenarios. We have also analysed the pros and cons of these strategies.

Sometimes the market is highly volatile and is not feasible to predict so we use algorithmic trading which takes in consideration various parameters and gives the results which are more likely to happen. Over the years we have shifted from traditional and naive strategies for prediction of market scenarios to more analytical ways by using data and algorithms.

The following figure shows some of the trade executions made by the algorithm. The figure contains only some of the transactions in order if breaking the orders. For example, the NTPC order is placed on Dec 11, 2019 and the order is broken on Dec 12,2019. In this transaction the order is placed for 1 day only and there is a loss of Rs.1387. In the other order like TECHM the order is placed on Dec 16,2019 and broken on Dec 20, 2019. The order is active for 4 days and there is a gain of Rs.5958. In the algo Trading there are some transactions of gain and there are some transactions with loss but overall trading transactions are making a gain to

Also in future we may see various institutions involved in creating new algorithms for market predictions and increase profitability by overcoming drawbacks of current algos. This may create new jobs and business opportunities at various levels.

7.REFERENCES

[1] Chan, Louis KC, Narasimhan Jegadeesh, and Josef Lakonishok. "Momentum strategies." *The Journal of Finance* 51.5 (1996): 1681-1713.

[2]Chen, Jie, Huidong Jin, and Shu-heng Chen. "Stock trend analysis and trading strategy." (2006).

[3] Junior, Leonidas Sandoval, Asher Mullokandov, and Dror Y. Kenett. "Dependency relations among international stock market indices." *Journal of Risk and Financial Management* 8.2 (2015): 227-265..

[4]Pätäri, Eero, and Mika Vilska. "Performance of moving average trading strategies over varying stock market conditions: the Finnish evidence." *Applied Economics* 46.24 (2014): 2851-2872.

[5] Pujol, Gregor, and Alexander Brueckner. "Smart Order Routing and Best Execution." *AMCIS 2009 Proceedings* (2009): 155.

[6]Zhao, Zhiyong, et al. "Time-weighted LSTM model with redefined labeling for stock trend prediction." *2017 IEEE 29th*

international conference on tools with artificial intelligence (ICTAI). IEEE, 2017.

[7]Kabasinskas, Audrius, and Ugnius Macys. "Calibration of Bollinger Bands parameters for trading strategy development in the Baltic stock market." *Engineering Economics* 21.3 (2010).

[8]Berghorn, Wilhelm. "Trend momentum." *Quantitative Finance* 15.2 (2015): 261-284.

[9] Christoffersen, Peter, and Sílvia Gonçalves. *Estimation risk in financial risk management*. CIRANO, 2004.

[10]Khodayari, Mahdi. "STATISTICAL ANALYSIS OF THE TECHNICAL STRATEGY OF TURTLE TRADERS IN THE FOREIGN EXCHANGE MARKET (FOREX MARKET) Masoumeh Akandeh1." *THE CONTEMPORARY SOCIO-ECONOMIC IRANIAN THOUGHTS*: 9.

[11]Chen, Jie, Huidong Jin, and Shu-heng Chen. "Stock trend analysis and trading strategy." (2006).