doi: 10.31674/ijrtbt.2021.v05i02.004

IJRTBT

CLASSIFICATION OF NIFTY STOCKS BASED ON PIVOT POINTS USING THE PRINCIPLE OF NEAREST NEIGHBOURHOOD

R. Subathra

Government Arts College (Autonomous), Tamil Nadu, India

**Corresponding Author's Email: rsubathraarumugam@gmail.com*

ABSTRACT

Investing in stocks oils the economic wheels of a country due to its impact on business investing, financial investing, government investing and consumer spending. In the perception of a trader, stock investing is a mind-boggling process mainly due to the availability of too many alternatives and too many indicators. The success in the process of investing depends on the usage of an ideal combination of indicators. Even if an investor has an ideal combination of indicators, the application of the same requires a statistical model which has the capability to sense the prospective Buy and Sell positions. Due to this reason many classification models of Statistics are gaining more and more importance in the field of Stock market investment. This work analyses three different methods of computing pivots namely standard method, DeMark method and Woodie's method. The objective of this work is to identify the most competing method of computing the pivot. Since the usage of any one technical indicator is not considered a good idea, the study identifies a combination of technical indicators to be used with the pivot points based on the statistical tests. Three combinations of the technical indicators are used to classify the stocks based on K-nearest neighbor. The study identifies the DeMark method as the most competing method and the result is also theoretically justified because this method gives more importance to the recent price action. The identification of the most competing method is done based on accuracy of the model, specificity and sensitivity as derived from the confusion matrix.

Keywords: ***Average True Range (ATR), Chaikin's Oscillator, Confusion Matrix, Momentum Indicator, Moving Average Convergence and Divergence (MACD), Relative Strength Index (RSI), Trend Indicator, Volatility Indicator, Woodie's Pivot Point, K-Nearest Neighbor***

INTRODUCTION

Even in the modern age when computers have revolutionized the process of handling big data, stock price forecasting remains a challenge. The reasons for this are the availability of a variety of statistical tools, the difficulty in selecting the predictors and the existence of too many technical indicators. The approach used by most of the traders is the technical analysis which is a bunch of indicators to predict the share price movements. These indicators claim that specific patterns will lead to movement of stock prices in specific directions. But the literature has proved that all those indicators have uncertainty to certain extend. It is this reason which motivates many researchers to use probabilistic models to make decisions based on technical indicators. The technical analysis tools have enhanced the usage of probabilistic models by throwing more light on the movement of prices. Using the traditional statistical models, the stock prices can be predicted, and the trader should combine his trading skills with these predictions to arrive at a trading decision. Instead, if a model can directly indicate the direction of the share price movement, the traders can make quick decisions. This can be successfully facilitated by the classification models. The objective of this work is twofold. It aims at identifying the best method of computing the pivot points by working on a combination of technical indicators to be used along with the pivot points. Consequently, the study also identifies the best combination of technical indicators to be used in conjunction with the pivot point. The identification is done using K-Nearest Neighbor which is a non-Parametric classifier.

LITERATURE REVIEW

The research done on the impact of technical indicators on stock prices is voluminous. Criticism of technical analysis is a topic of concern in academic research which supports the “weak form” efficient market hypothesis as defined by Fama (1970). The validity of technical analysis is often dismissed due to the belief that stock markets follow a random walk.

The studies by Fama (1965), Fama & Blume (1966) and Jensen & Benington (1970) advocated random walk theory.

Metghalchi, Chang & Marcucci (2008), Brock, Lakonishok & LeBaron (1992) suggested that simple moving average techniques have predictive power when examining the Dow Jones Index between 1897 and 1985. Similar results were established by Bessembinder & Chan (1998) and Ellis & Parbery (2005). However, both Bessembinder & Chan (1998) and Ellis & Parbery (2005) suggested that the buy-and-hold strategy is superior.

Kwon & Kish (2002), on the other hand, suggested that technical trading rules had the possibility to be more profitable than a buy-and-hold strategy when examining the NYSE.

Teixeira & de Oliveira (2010) proposed a method with stop loss, stop gain, and RSI filter in the nearest neighbor classification algorithm.

Imandoust & Bolandraftar (2014) developed three models and compared their performances in predicting stock price movement in Tehran Stock Exchange (TSE) Index. 10 macroeconomic variables were used as input and the analysis was done with decision tree model, Random forest, and Naïve Bayesian Classifier classification techniques. The experiment resulted in Decision tree model with 80.08% accuracy, Random Forest with 78.8% accuracy and Naïve Bayesian Classifier with 73.8% accuracy.

Subathra (2020) examined three different types of pivots using discriminant analysis and identified that the pivot point which assigns maximum weight to close price gives better classification.

Brock, Lakonishok & LeBaron (1992) proved that technical analysis helps to understand stock price changes. Teixeira & de Oliveira (2009) predict stock price trends using the technical indicators. They used K-NN classification model with closing prices and trading volumes. A comparison of this trading system with the purchase, maintenance and sales system reveals that this system is more efficient.

Son & Noh (2012) predicted KOSPI200 index using Binary classification. The inputs of the model included the technical indicators. The results of the classification models are compared with three different sets of data. The results indicate the superiority of the SVM method when no dimensional reduction has been made.

Di (2014) focused on prediction of stock price trend using technical indicators. Using data was related to the stock prices of the Apple, Amazon and Microsoft companies, the study proves the capability of classification techniques.

A Naive SVM-KNN based stock market trend reversal analysis for Indian benchmark indices was done by Nayak, Mishra & Rath (2015). They have used Support Vector Machine and K-nearest neighbor for prediction. They predicted the profit and loss using SVM kernel function and this output is used to identify the best prediction set to be used as an input to the K-nearest neighbor.

RESEARCH METHODOLGY

In general, the most promising points of entry and exit are identified using technical indicators. The trend, volume, momentum, and volatility indicators are the four major classifications of technical indicators. The previous session's high price, low price, close price, and open price are also the potential information to understand the trend of the prices. A scientific combination of all the above indicators is the Pivot point. They enjoy wide applications due to following reasons:

1. The frequency of the time series data to be used is always a subject of discussion in share analysis. The resourceful aspect of pivot point is that it can be computed over different time frames. At any subsequent time period, prices above the pivot point are an indication of uptrend while prices below the pivot point is an indication of downtrend.

2. The moving averages and oscillators are generally dynamic, but the pivot points are static since they remain the same throughout the trading session. Any largest price movement is expected to occur at the pivot point, and it serves as the basis on which the projections of the support and resistance levels are made.

The belief that the interaction at the pivot points causes a reaction is not always true as revealed by the experience of the traders. Thus, the pivot points like any other technical indicator have some randomness in their behavior. Random behavior of any variable deserves an explanation through probabilistic models. Since the objective of this study is to compare various types of pivot points, they are used as dependent variables in this study. No technical indicator has proved to be so perfect in predicting the trend and it is a usual practice to combine a few technical indicators to arrive at a trading decision. Hence to increase the chance of success it is necessary to frame probabilistic models in which pivot point is a dependent variable and other technical indicators are predictors.

Various methods of computing the pivot points are in practice. In this work three different methods of computing the pivot points are compared using multiple indicators. They are Standard pivot point, Woodie's pivot point and DeMark pivot point.

Let C be the previous Close price, O the previous Open price, L the previous Low price and H the previous High price.

- Standard Pivot Point is the most basic Pivot Point. It is the simple average of High, Low and Close from a prior period.
- The Woodie's Method is a weighted average which assigns more weight to the close price. According to this method Pivot point = (H + L + 2C) / 4
- Demark Pivot Points are conditional on the relationship between the close and the open. According to this approach,

If C < O, then X = H + (2 X L) + C

If C > O, then X= (2 X H) + L + C

If C = O, then X= H + L + (2 X C)

Pivot point = X / 4

This work tests the credibility of the three different approaches of computing the pivot points using multi-indicator strategy. A multi-indicator strategy may become redundant when they provide same type of information. To overcome the redundancy, one indicator is selected from each broad category of technical indicators. The following technical indicators are used to analyze the performance of the three types of pivots.

Trend indicator-Moving average convergence-divergence:

Moving Average Convergence Divergence (MACD) is considered as a leading indicator, but with a bit of lag. It is the difference of 26-period Exponential Moving Average and 12-period Exponential Moving Average. When the MACD line crosses above the signal line or when the MACD line crosses above the Zero line, it is a buy signal. On the other hand, when the MACD line crosses below the signal or crosses down the Zero line, a Sell signal is generated.

Momentum Indicator-Relative Strength Index:

Relative Strength Index (RSI) is a leading indicator that measures the speed and the price movement. RSI oscillates between zero and 100. RSI is a kind of momentum oscillators that is used to calculate the market recent gains against its recent losses and translates that information into a number between 0 and 100. The value of RSI is considered overbought when above 70 and oversold when below 30.

Volatility Indicator-Average True Range:

The Average True Range is considered as an accurate volatility measure. It measures the intensity of movement of an asset in the past.

Volume Indicator-Chaikin Oscillator:

The Chaikin Indicator is used to analyze the strength of a price trend based on trading volume. It anticipates directional changes like momentum indicators because it measures the momentum behind the movements in the Accumulated Distribution Line.

The main objective of this work is to compare the categorical trend variables generated using three different types of pivots. The ability of the pivot points to classify the trend can be analyzed using classification algorithms. The techniques which perform this classification include Bayes Classifier, K-Nearest Neighbour (KNN), Decision trees, Discriminant Function and Logistic Regression. Among these methods the K-Nearest Neighbour is a classifier which works on a similarity measure called distance function. It assumes that similar things exist in proximity. The main advantage of this method is that it is a non-parametric method and hence it is not necessary to pre-process the data to make it suitable for the tool. The tool has the integrity to store all the extracted information and hence it classifies a new data based on its similarity. In this method the historical stock data and the test data are mapped into a set of vectors. Then a similarity metric such as Euclidean distance is computed to decide. All other classification methods build a model, or a function first and then classifies the testing data using the classifier function. But KNN does not build a model previously and hence it is termed as a lazy algorithm. The algorithm first determines the value of K which is the number of nearest neighbors. Then it computes the distance between training and test data. These calculated distances are then arranged according to the magnitude. It uses a maximum vote assigned for the class labels of K nearest neighbour and assigns it as a prediction value for the test data. The success in this method is determined by the value of K which is the number of nearest neighbours. Low values of K lead to more errors but high values are difficult to operate. The most preferred value of K is 5.

Evaluation Measures used for the classification model:

In general, the performance of the classifier is evaluated based on confusion matrix. Among various evaluation measures furnished by the Confusion matrix, the following measures are used in this study:

i. The ratio of number of correct predictions to the total number of predictions is called Accuracy of the fitted model. The accuracy achievable by always predicting the majority class label is called No Information Rate (NIR). Probability that the accuracy (ACC) is greater than NIR is given by the *p*-value.

ii. Sensitivity gives the proportion of the positive class which are correctly predicted. This shows the proportion of defaulters correctly predicted.

iii. Specificity gives the proportion of the negative class which are correctly predicted. This shows the proportion of those who paid that were correctly predicted.

iv. Positive Predictive Value gives the number of the positive class correctly predicted as a proportion of the total positive class predictions made.

v. Negative Predictive Value gives the number of the negative class correctly predicted as a proportion of the total negative class predictions made.

vi. Prevalence gives how often the positive class occurs in our sample.

vii. Detection Rate shows the number of correct positive class predictions made as a proportion of all the predictions made.

RESULT

The daily open price, close price, High price, and Low price of the NSE indices NIFTY50 from January 2015 to July 2019 collected from the official website of National Stock Exchange is used in this study. The pivot points are computed using the Standard approach, Woodie's approach and DeMark's approach. These pivot points are used to generate the dependent variable. If the close price is greater than pivot point, the trend value is 1 and otherwise it is 0. The trend variables strend (generated using Standard pivot), dtrend (generated using DeMark pivot) and wtrend (generated using Woodie's pivot) are the dependent variables in this study.

Introducing the predictors:

Two important issues are considered in this study. The first issue is to find a suitable method of computing the pivot among the three methods considered in this study. The second issue is to find the best combination of indicators to be used in conjunction with the pivot point. To enable the same the three different types of pivot points are used in conjunction with the technical indicators as follows:

Models with the Standard pivot in conjunction with other technical indicators (refer to table 1):

MODEL I: The standard pivot used in conjunction with MACD, RSI, ATR, and Chaikin Oscillator.

MODEL II: The standard pivot used in conjunction with MACD and RSI.

MODEL III: The standard pivot used in conjunction with RSI and ATR.

MODEL IV: The standard pivot used in conjunction with RSI and Chaikin's Oscillator.

Table 1: Results of the models constructed with Standard Pivot Point

Measure	Model I	Model II	Model III	Model IV
Accuracy (ACC)	0.4925	0.4835	0.7147	0.997
No Information Rate (NIR)	0.5345	0.5345	0.5345	0.5345
p-Value for ACC> NIR	0.944	0.9726	1.311e-11	2e-16
Sensitivity	0.3226	0.2839	0.6194	0.9935
Specificity	0.6404	0.6573	0.7978	1
Positive Predictive Value	0.4386	0.4190	0.7273	1
Negative Predictive value	0.5205	0.5132	0.7065	0.9944
Prevalence	0.4655	0.4655	0.4655	0.4655
Detection Rate	0.1502	0.1321	0.2883	0.4625

In Model I where Standard pivot is used in conjunction with MACD, RSI, ATR and Chaikin Oscillator *p*-value is not significant. The same is the case of Model II in which Standard pivot is used in conjunction with MACD and RSI. The *p*-value for model III in which Standard pivot is used in conjunction with RSI and ATR is significant. Also, in Model IV where the Standard pivot is used in conjunction with RSI and Chaikin Oscillator, the *p*-value is significant. Among the two models MODEL III and Model IV, Model IV is identified as the best based on accuracy, Sensitivity, Specificity, Positive Predictive value, and Negative Predictive value.

Models with the Demark's pivot in conjunction with other technical indicators (see table 2 below):

MODEL V: The Demark's pivot used in conjunction with MACD, RSI, ATR, and Chaikin Oscillator.

MODEL VI: The Demark's pivot used in conjunction with MACD and RSI.

MODEL VII: The Demark's pivot used in conjunction with RSI and ATR.

MODEL VIII: The Demark's pivot used in conjunction with RSI and Chaikin's Oscillator.

Table 2: Results of the models constructed with Demark's Pivot Point

Measure	Model V	Model VI	Model VII	Model VIII
Accuracy (ACC)	0.5105	0.5435	0.7147	1
No Information Rate (NIR)	0.5435	0.5435	0.5435	0.5435
p-Value for ACC> NIR	0.8970	0.5226	1.127e-10	2.2e-16
Sensitivity	0.3224	0.3947	0.6184	1
Specificity	0.6685	0.6685	0.7956	1
Positive Predictive Value	0.4495	0.5000	0.7176	1
Negative Predictive value	0.5402	0.5681	0.7129	1
Prevalence	0.4565	0.4565	0.4565	0.4565
Detection Rate	0.1471	0.1802	0.2823	0.4725

In Model V where Standard pivot is used in conjunction with MACD, RSI, ATR and Chaikin Oscillator p-value is not significant. The same is the case of Model VI in which standard pivot is used in conjunction with MACD and RSI. The *p*-value for model VII in which standard pivot is used in conjunction with RSI and ATR is significant. Also, in Model VIII where the standard pivot is used in conjunction with RSI and Chaikin Oscillator, the *p*-value is significant. Among the two models MODELVII and Model VIII, Model VIII is identified as the best based on accuracy, Sensitivity, Specificity, Positive Predictive value, and Negative Predictive value.

Models with the Woodie's pivot in conjunction with other technical indicators (refer to table 3):

MODEL IX: The Woodie's pivot used in conjunction with MACD, RSI, ATR, and Chaikin Oscillator.

MODEL X: The Woodie's pivot used in conjunction with MACD and RSI.

MODEL XI: The Woodie's pivot used in conjunction with RSI and ATR.

MODEL XII: The Woodie's pivot used in conjunction with RSI and Chaikin's Oscillator.

Table 3: Results of the models constructed with Woodie's Pivot Point

Measure	Model IX	Model X	Model XI	Model XII
Accuracy (ACC)	0.997	1	0.7177	1
No Information Rate (NIR)	0.5345	0.5435	0.5405	0.5405
p-Value for ACC> NIR	2e-16	2.2e-16	2.569e-11	2.2e-16
Sensitivity	0.9935	1	0.7582	1
Specificity	1	1	0.6833	1
Positive Predictive Value	1	1	0.6705	1
Negative Predictive value	0.9944	1	0.7687	1
Prevalence	0.4655	0.4565	0.4595	0.4595
Detection Rate	0.4625	0.4565	0.3483	0.4595

The *p*-values for all the four models are significant. Among these models, Model XII is identified as the best based on accuracy, Sensitivity, Specificity, Positive Predictive value, and Negative Predictive value.

The model evaluation measures suggest Model IV, Model VIII and Model XII as the best in their respective groups. Thus, the pivot points when used in conjunction with RSI and Chaikin's oscillator provides the best classification. The results of these three models are now compared to identify the best approach of computing the pivot. Table 4 is a summary of the models IV, VIII and XII.

Table 4: Results of the models constructed in conjunction with RSI and Chaikin's oscillator

Measure	Model IV	Model VIII	Model XII
Accuracy (ACC)	0.997	1	1
No Information Rate (NIR)	0.5345	0.5405	0.5405
p-Value for ACC> NIR	2e-16	2.2e-16	2.2e-16
Sensitivity	0.9935	1	1
Specificity	1	1	1
Positive Predictive Value	1	1	1
Negative Predictive value	0.9944	1	1
Prevalence	0.4655	0.4565	0.4595
Detection Rate	0.4625	0.4725	0.4595

For all the three models, the evaluation measures are approximately equal, but a careful analysis of the evaluation measures gives the following conclusion.

The accuracy, Sensitivity and Negative predictive value for the standard pivot are less compared to the other two approaches. Hence among the three approaches, the standard method has less capability to classify the stocks. Also, among the other two approaches namely DeMark's approach and Woodie's approach, DeMark's procedure is preferred based on the detection rate.

DISCUSSION

The availability of too many alternatives for trading and too many technical indicators make the trading decision a complex process. It is a normal practice to use a combination of technical indicators as a guide to make a trading decision (Brock, Lakonishok & LeBaron, 1992; Bessembinder & Chan, 1998). The use of too many indicators leads to inefficient decisions. To optimize the usage of indicators, redundancy should be avoided. An improper combination leads to the multiple counting of the same information. A good combination should contain indicators that complement each other. An ideal combination is the one which throws light on trend, volume, momentum, and volatility. With this realization this work considers the Moving Average Convergence and Divergence (MACD) for studying trend. The Relative Strength Index for studying momentum and

Average True Range for studying the Volatility and Chaikin's oscillator for studying the volume. The aim of this work is to compare the categorical trend variables generated using three different types of pivots namely Standard pivot, DeMark's pivot and Woodie's pivot (Subathra, 2020). Among the widely used classification methods like Naïve Bayes classifier, K-nearest neighbor, Decision trees, discriminant function and Logistic Regression, the K-nearest Neighbor is a non-Parametric method. Since the inclusion of some technical indicators becomes impossible due to the strong assumptions of parametric theory, this study considers the Non-Parametric method for classification. The suitability of the fitted models is analyzed using Accuracy, No Information Rate, *p*-value, sensitivity, specificity, Positive Predictive value, Negative Predictive value, and detection rate which are the outcomes of Confusion matrix. Six models had *p*-values less than 0.05.

CONCLUSION

A comparative analysis of the results highlighted that the pivot points used in conjunction with RSI and Chaikin's oscillator have more accuracy. The Accuracy, Sensitivity and Negative predictive value for the standard pivot are less compared to the other two approaches. Hence among the three approaches, the standard method has less capability to classify the stocks. Also, among the other two approaches namely DeMark's approach and Woodie's approach, DeMark's procedure is preferred based on the detection rate. This is theoretically justified because the DeMark approach gives more importance to recent price action. This work can be further extended by incorporating more technical indicators in the model to increase the accuracy of prediction.

Conflict of Interests

The author declares that he has no conflict of interest.

ACKNOWLEDGEMENT

The author is thankful to the institutional authority for completion of the work.

REFERENCES

Bessembinder, H. & Chan, K. (1998). The profitability of technical trading rules in the Asian stock markets. *Pacific-Basin Finance Journal*, 3(2-3), pp 257-284.

Brock, W., Lakonishok, J. & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance*, 47(5), pp 1731-1764.

Di, X. (2014). Stock Trend Prediction with Technical Indicators using SVM. SCPD student from Apple Inc. Retrieved From: http://cs229.stanford.edu/proj2014/Xinjie%20Di,%20Stock%20Trend%20Prediction%20with%20Technical%20Indicators%20using%20SVM.pdf

Ellis, C. & Parbery, S.A. (2005). Is smarter better? A comparison of adaptive, and simple moving average trading strategies. *Research in International Business and Finance,* 19(3), pp 399-411.

Fama, E.F. & Blume, M.E. (1966). Filter Rules and Stock Market Trading Profits. *Journal of Business*, 39(1), pp 226-341.

Fama, E.F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1), pp 34-105.

Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), pp 383-417.

Imandoust, S.B. & Bolandraftar, M. (2014). Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange, *International Journal of Engineering Research and Applications*, 4(6), pp 106-117.

Jensen, M & Benington, G. (1970). Random Walks and Technical Theories: Some Additional Evidence. *The Journal of Finance*, 25(2), pp 469-482.

Kwon, K.Y. & Kish, R.J. (2002). Technical trading strategies and return predictability: NYSE. *Applied Financial Economics*, 12(9), pp 639-653.

Metghalchi, M, Chang, Y.H. & Marcucci, J. (2008). Is the Swedish stock market efficient? Evidence from some simple trading rules. *International Review of Financial Analysis*, 17(3), pp 475-490.

Nayak, R.K., Mishra, D, & Rath, A.K. (2015). A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. *Applied Soft Computing*, 35(1), pp 670-680.

Son, Y., Noh, Dj. & Lee, J. (2012). Forecasting trends of high-frequency KOSPI200 index data using learning classifiers. *Expert Systems with Applications*, 39(14), pp 11607-11615.

Subathra, R. (2020). Discriminating the Pivot Points using Discriminant Function. *Data Mining and Knowledge Engineering*, 12(6), pp 109-114.

Teixeira, L.A. & de Oliveira, A.L.I. (2009). Predicting stock trends through technical analysis and nearest neighbor classification. IEEE International Conference on Systems, Man and Cybernetics, USA. Retrieved From: http://vigir.missouri.edu/~gdesouza/Research/Conference_CDs/IEEE_SMC_2009/PDFs/997.pdf

Teixeira, L.A. & de Oliveira, A.L.I. (2010). A method for automatic stock trading combining technical analysis and nearest neighbour classification. *Expert Systems with Applications*, 37(10), pp 6885–6890.