Leveraging Traditional ML models & Strategy indices for Retail Traders: A Study in Indian Context

Subhajit Chattopadhyay

Research Scholar
Lovely Professional University, Punjab , India

October 16, 2024

Abstract

Selection of technical indicators, evaluation period matching with forecasting horizon and selection of strategy indices helps to improve trading performance even with simple ML techniques. The return of Buy & Hold (B & H) strategy from Nifty 50 stands at 205.5% . But with customised evaluation period the maximum return generated by Nifty 50 , Nifty Alpha 50 and Nifty 200 Quality 30 are 289.3%, 647% and 349% respectively. Such performance also outperforms the return from B& H strategy of respective indices. Return differential of Nifty 50 between standard and customized evaluation period are 6.6% , 6% and 391% for 1 day , 3 day and 5 day trading period respectively. Similar, trend can be observed for other indices as well. The study identified the important predictors of periodic return of the selected indices and compared the performance of regularized and adaptive regressions.

Keywords: strategy indices, regularised regression, multivariate adaptive regression spline, prediction of market movement, prediction of periodic return

1 Introduction

In the recent past, SEBI has published two reports, Handbook of Statistics (2022) [1] and Analysis of Profit and Loss of Individual Traders dealing in Equity F&O Segment (2022) [2]. These reports underline following patterns and characteristics about Indian stock market and trading behaviour of the retail traders and investors. Following patterns have been observed for NSE (national Stock Exchange) & BSE (Bombay Stock Exchange) -

- At NSE, use of non-algo trade has decreased by 57.48% over the period 2018 - 2023 (in derivatives trading). At BSE, it reduced by 99.98% over the same period.
- Over the same period, mobile trading increased by 2.78 times (approx. 3 times).
- At NSE, use of facilities like Co-location has increased by 43.94% (approx 44%). At BSE, use of Co-location increased by 1.57 times.

Following are the salient observations of the report on Analysis of Profit and Loss of Individual Traders dealing in Equity F&O Segment (2022) -

- active traders (who traded through top 10 brokers) increased 5 times between 2019 and 2022. Top 1% and top 5% of the active traders account for 51% and 75% of the total net profit respectively.
- between 2019 and 2022 ,number of individual traders index options and stock options went up by 8 and 5 respectively. The preference to trade in options has increased compared to that of future. 89% of the individual traders who traded in index options incurred an average loss of Rs.77,000. 89% of the individual traders incurred losses averaging Rs. 110,000 in 2022.
- for active traders, the average loss was over 15 times of the average profit. And traders are preferring stock and index options.

Central Issue: Skewness of profit distribution, low profitability underlines abysmal performance of retail traders which may be the result of inability to use technology and /or sophisticated algorithms, given the cost associated with such options [3], [4]. Poor performance of retail traders can be attributed to their inability to predict the direction of movement of the stock , sector and market along with lack of expertise in predicting future periodic return . Direction and periodic return helps to ascertain entry and exit points.

[1] https : //www.sebi.gov.in/reports − and − statistics/publications/may − 2023/handbook − of − statistics − 2022$_7$4606.html

[2] https : //www.sebi.gov.in/reports − and − statistics/research/jan − 2023/study − analysis − of − profit − and − loss − of − individual − traders − dealing − in − equity − fando − segment$_6$7525.html

[3] https : //www.nseindia.com/trade/platform − services − co − location − facility

[4] https : //nsearchives.nseindia.com/web/sites/default/files/inline − files/Download$_R$eal$_T$ime$_T$ariff$_D$omestic$_0$1042023.pdf

Approach: Therefore, the study aims to explore how choice of strategy indices and traditional machine learning techniques can help retail traders. To understand future movement of market logistic regression has been applied. And for prediction of periodic return regularised and adaptive regressions have been used. Further, traders pursue various trading objectives, some follow trend, some goes against the trend , some chase alpha etc. Kakushadze et al. (2018) in his book outlined 151 trading strategies. Similarly, Colby (2003) has indicated trading rules various technical indicators. Now, strategy indices also helps investors or traders to pursue a trading strategy. In this study, two strategy indices such as Nifty Alpha50 (henceforth Alpha50) and Nifty 200 Quality30 (henceforth Quality30) have been selected [5] . The first index is constituted by the companies having high alpha, good market capitalization and liquidity. Sectoral allocation weights of Financial Services and Capital Goods stand at 29.33% and 17.42% [6]. For Quality30, the sectoral allocation weights are 26.97% for FMCG and 23.71% for IT [7]. Alpha50 represents risk-adjusted return based strategy and constituted by companies having good market capitalization and liquidity. And the second one allows to invest in financially strong companies which can be termed as value based trading.

The strategy indices can be expected to behave in a consistent manner in the context of movement of broader market. Thus, one can take positions with respect to strategy indices as well as the constituent sectors and companies based on broader market prediction. Further, such indices help traders a particular trading strategy irrespective of movement of broader market. The following figure shows the movements of Alpha50 and Quality30 along with movement of Nifty50 -

[5] $https://www.niftyindices.com/reports/historical-data$

[6] $https://www.niftyindices.com/Factsheet/Factsheet_Nifty_Alpha50.pdf$

[7] $https://www.niftyindices.com/Factsheet/Factsheet_NIFTY200_Quality30.pdf$

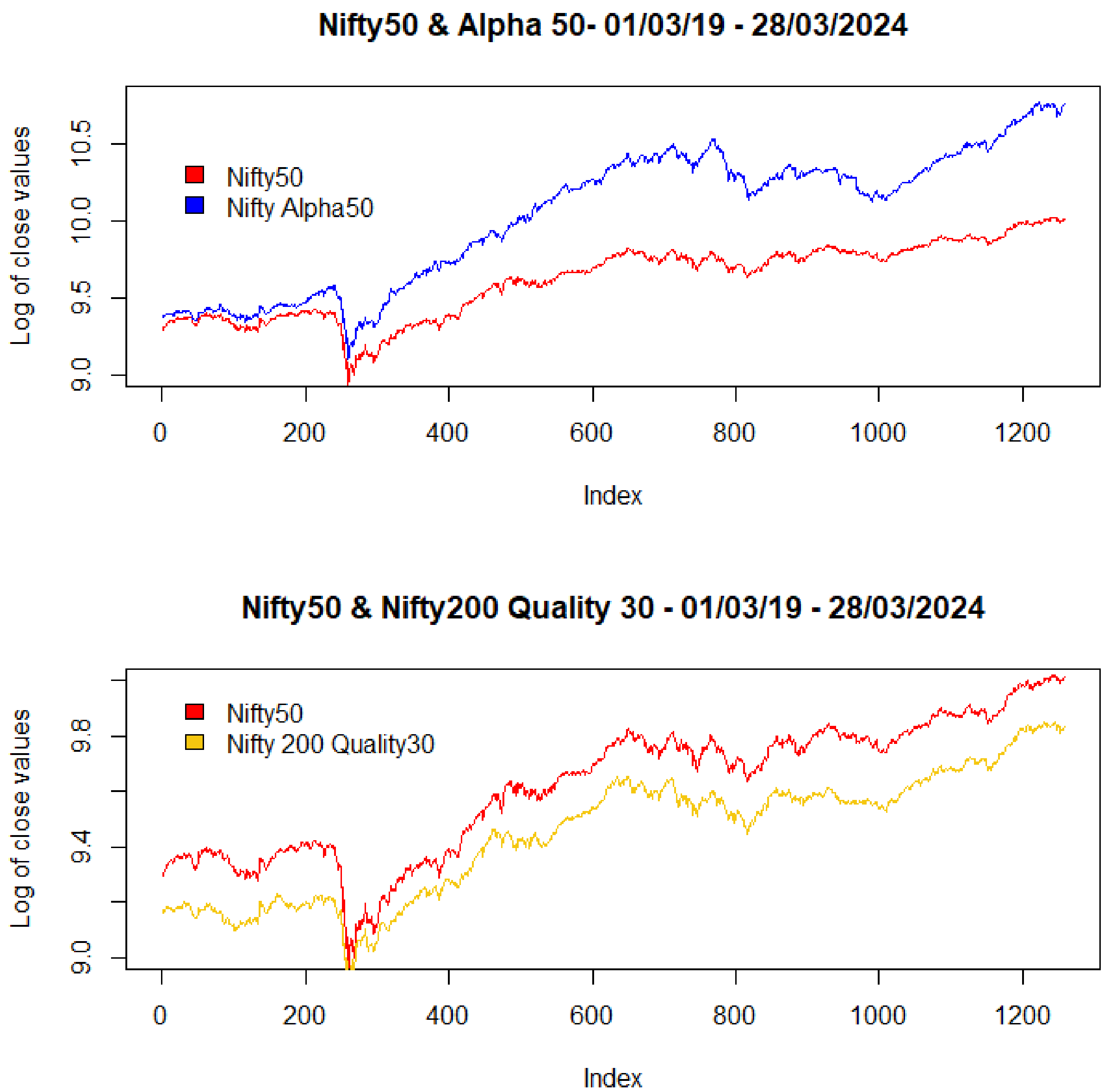

Figure 1: Movement of Alpha50, Quality30 and Nifty50 indices, red line - Nifty50 , blue line - Quality30 or Alpha50

Its apparent from the above figure that the movements of strategy indices are consistent with Nifty50. Alpha50 offers more return and momentum , whereas Quality30 offers less return but volatility is also less.

Given the consistency of movements of indices, from broad market movement one can predict the movement of strategy indices and thereby the movements of constituent sectors and stocks which are traded in the market. Thus, strategy indices helps to make systematic choice of sectors and stocks and also director of trade which is in line with trading objective.

So, issues related with prediction of direction and selection of stocks are addressed to an extent.

The following diagramme presents a logical framework. It presents two paths to predict direction or return of strategy or sectoral indices or selected stocks. In the context of prediction of direction for strategy indices can be done independent of broad market direction . Alternatively, the way is to factor in the broad market movement.

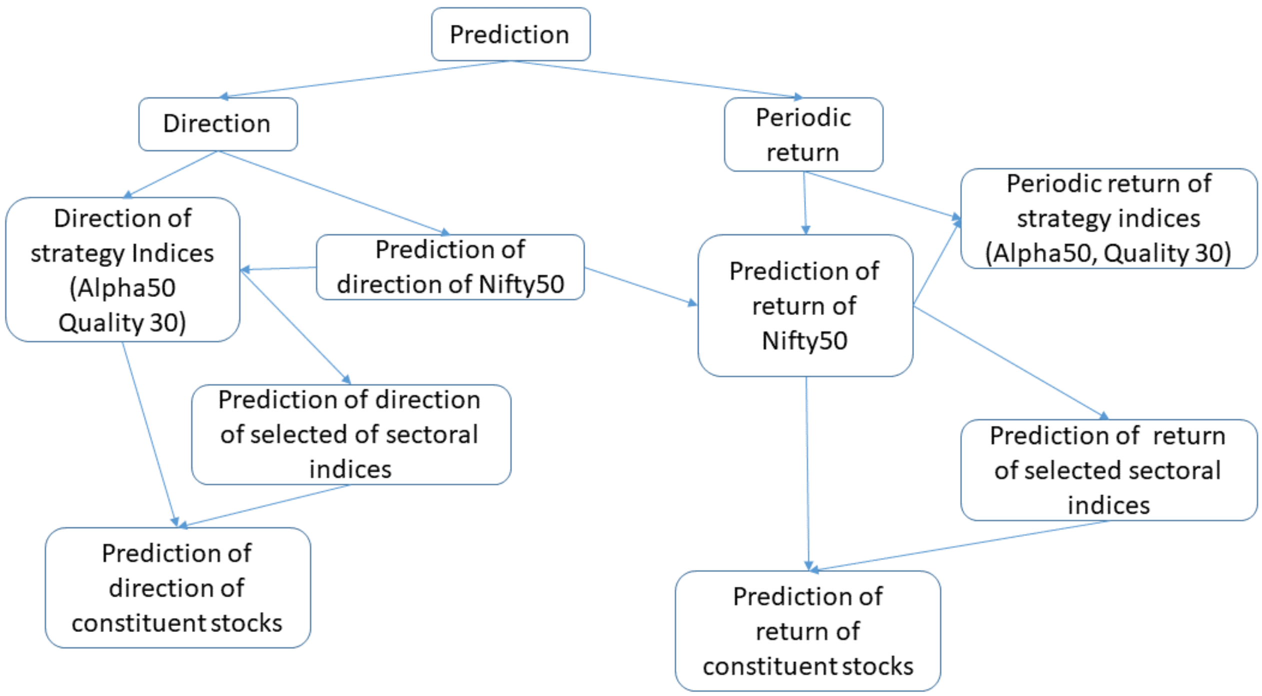

Figure 2: Logical framework

Objectives: Given the ease of computing, technical indicators have been increasingly being used for various predictions (direction of movement, return, value of index or stock(. (Alsubaie et al., 2019), (Nazário et al., 2017), (Beyaz et al., 2018)). Technical indicators are capable of capturing various characteristics of market such direction of movement, momentum, volatility etc. But there is an ever growing list of such indicators and there is no merit in applying all the technical indicators.

As indicated in the previous studies that mobile phones are increasingly being used for trading purposes which leaves us with the options that are computationally not very demanding such as decision trees with ensemble learning, neural network etc. In this context the objectives of the study are as follows -

- to examine the performance aforesaid supervised learning models with respect to prediction of direction and prediction of periodic returns over short term (over 1 , 3 , 5 days)

- to examine how of technical indicators with standard evaluation and customised evaluation period impact the performance of the supervised models for Nifty50 , Nifty Alpha50 , Nifty200 Quality30.
- to explore whether strategy indices can outperform broad market index e.g. Nifty50
- to estimate the impact of change in evaluation period on cumulative return for the aforementioned indices

2 Literature Review

Researchers have adopted various approaches for short term prediction. This ranges from graph based techniques(Tan et al., 2021), technical indicator based statistical and machine learning methods (Tang et al. (2019), Yao et al. (2021)), use of evolutionary computation techniques to generate trading rules (Chen et al., 2021). This also incorporates use of wavelet principles , Gann's time cycle (Zhou et al. (2021), principles of direction change (Kampouridis and Otero (2017). In the section some relevant past works on prediction of direction shall be presented.

Chandar and Punjabi (2021) proposed CSO to train multi-layer perceptron to predict the closing price of the stocks on the same day. The model with 9 technical indicators was tested 10 stocks and the performance of the model is superior compared to GA-LSTM, DE-FLNN, GA-RBF, CSO-ARMA, PSO-ELMAN, PSO-MLP, BBO-MLP. (Chen et al., 2023) proposed a model where turning points are predicted using PLR and trained the model with WSVM , backward propagation neural network , gradient boosted decision trees, random forest and long short term memory. It was observed that PLR with weighted support vector machine performed the best. Lv et al., 2019 compared the performance of six machine learning (ML) algorithms (i.e. support vector machine , logistic regression, random forest, naive bayes, XG boost, classification and regression trees) with six deep neural network models (DNN) (i.e. MLP, DBN, LSTM, SAE, GRU, RNN). It is observed that traditional ML algorithms at par with DNN models with respect to prediction of direction. But trading strategies based on DNN result into better profit after consideration of transaction cost.

(Zhou et al., 2021) predicted the direction of daily return using 60 indicators for deep learning models. They has used three variants - first all the raw indicators (with out transformation), secondly all transformed indicators and subsequently 31 principal components. The structure of model varied from 10 to 1000 layers. Dataset with 31 components has marginally done better compared dataset with 60 indicators (raw). The highest accuracy achieved was 59.6% . Kilimci and Duvar (2020) did a comparative studies where 3 deep learning algorithms (convolutional neural network , recurrent neural network , long short term memory) along with 4 types (Word-to-vec , Glove, BERT, FasText) of word embedding for 9 most traded bank stocks to predict the direction BIST100 index. It observed that a combination with word embedding and a deep learning algorithms outperforms individual performance of word embedding and deep learning techniques. Performance vary with source of texts and the accuracies varied from 77.82% to 86.39%. Tang et al. (2019) used weighted support vector machine (WSVM) coupled with piece-wise linear representation (PLR) to predict change in direction. Along with WSVM and PLR, the proposed model also added another RSI (relative strength index) based trading rule along with over or under sampling correction for turning point prediction. The proposed method outperformed models where only PLR & WSVM, PLR , WSVM & RSI, PLR & ANN are used. Directional change is another approach to examine movement of data points. The underlying concepts rests on the use of θ , which is a threshold value used to confirm the change in direction. (Palsma and Adegboye, 2019) in their study used particle swarm optimization (PSO) , continuous shuffled frog leaping algorithm (CSFLA), genetic algorithm (GA) and genetic programming (GP) to optimise the value of θ for 4 currency pairs. It was observed GP performed the best followed by CSFLA. Joginpally (2023) proposed Spiking Quantum Neural Network (SQNN) for prediction of S&P100 , ASX200 and SSE50 indices. It is observed thatS QNN outperformed support vector regression (SVR), long-short term memory (LSTM), particle swarm optimization (PSO) with LSTM amd Improved PSO with LSTM. The accuracy of SQNN for US , Australia and China markets are 95.2% , 94% and 93.56 respectively. It also shows that as look back period increases errors metric for different metrics (e.g. MAPE and RMSE) behave differently. As look-up period increased errors increased for LSTM and PSO-LSTM and the same decreased for SVR & IPSO- LSTM. RMSE decreased and MAPE increased for SQNN as look back period increased. Ohaegbulem and Kalu (2023) used logistic regression to predict the direction NSX ALL Share index using other sectoral index such as banking, insurance, oil & gas, consumer goods. In the process they could achieve an accuracy of 70.17%. Ren et al. (2018) predicted next day market direction using support vector machine and also factored in sentiment analysis and day of the week effect. Apart 8 sentiment index they also used open, high, low, close, volume, financial volume, and changes in volume. The study shows that without sentiment component accuracy for both SVM and logistic regression was around 71% . After factoring in sentiment indices, the accuracy went up to 89.93% and 86.59% for SVM and logistic regression respectively.

3 Materials & Methods

Strategy indices but helps investors to allocate the fund which is in sync with the chosen trading strategy whether it is chasing risk-adjusted return (e.g maximising Alpha), minimising volatility or investing in financially sound companies. NSE has developed 35 such strategy indices [8]. But all the strategy indices are not traded at NSE. As one date, there are 13 strategy indices that are traded in the exchange [9]. Nifty Alpha 50 was the older strategy indices which was launched on 19th November 2012. For the curent study, two strategy indices Nifty Alpha50 and Nifty200 Quality30 have been selected and Nifty50 has been selected as index of the broader market.

3.1 Methods & Target variable

There are numerous studies with respect to application of machine learning algorithms for prediction of direction and values (Lv et al. (2019); Faghihi Nezhad and Minaei Bidgoli

[8] $https://www.niftyindices.com/indices/equity/strategy-indices$

[9] $https://www.nseindia.com/reports-indices-historical-index-data$

(2021), Gandhmal and Kumar (2019)). Also there are studies which deploys multiple architectures or methods such as (Ouyang et al., 2020). But, in the context of retail traders where interpretability (i.e. importance of various TIs) and simplicity of the model are the key concern, logistic regression has been used for prediction of direction (i.e up or down) . And given the non-linearity of stock market data regularised regressions and adaptive multiple regression (splines) have been used for predicting returns. Splines not only help in addressing non-linearity issues , but also looks into interaction among technical indicators. Since the study is on short term trading , the forecasting horizon has been limited to 1 day, 3 days, 5 days.

The split criteria for cross-validation of the aforementioned models stands at 70:30 and the number of folds of cross validation has been set at 10. Further , for the adaptive regression (Spline) model , a hyper grid has deployed which can take care of linear to cubic relations among the various technical indicators and its interaction variables. Number of such terms has been limited to 100 and only top most important 10 terms will be displayed.

The target or the dependent variable has been defined as the logarithm of periodic return for 1 , 3 and 5 days i.e. $Y = log(Y_{t+i}/Y_t)$. To capture the direction of return, the values of Y has been discretized such as values greater than 0 has been considered as 1 else 0.

3.1.1 Logistic Regression:

When the dependent variable in a regression equation, is not continuous i.e. binary or multi-leveled (or multinary) then logistic regression is used. It yields the probabilities of various observation belonging to one of the classes or levels. And the following equation is used -

$$p(X) = \frac{\exp^{(\beta_0+\beta_i*\sum x_i)}}{1+\exp^{(\beta_0+\beta_i*\sum x_i)}}$$

x_is are the predictor of the model and p(x) is the probability of belonging to a class. The logit transformation of the above equation lead to the following

$$g(x) = \frac{p(x)}{1-p(x)} = \beta_0 + \beta_i * \sum x_i$$

3.1.2 Regularised Regression:

Regularised regression is another simple supervised model which is primarily used in cases where assumptions of linear models can not be maintained. As the number of features in a model grows assumptions like estimates of coefficients may not be unbiased or have lowest variance. Regularisation helps to minimise the variance and sampling error.

In regularised regression the following term is minimised -

$$(\sum_{i=1}^{n}(y_i - \hat{y})^2 + Penalty$$

There can be 3 types of penalty which can be applied. For ridge regression the objective is to minimise $[\sum_{i=1}^{n}(y_i - \hat{y})^2 + \lambda \sum_{i=1}^{p} \beta_i^2]$. In case of least absolute shrinkage and selection operator (Lasso) the objective is to minimise $[\sum_{i=1}^{n}(y_i - \hat{y})^2 + \lambda \sum_{i=1}^{p} |\beta_i|]$. For elastic nets the term $[\sum_{i=1}^{n}(y_i - \hat{y})^2 + \lambda_1 \sum_{i=1}^{p} \beta_i^2 + \lambda_2 \sum_{j=1}^{p} \|\beta_j|]$ is minimised.

3.1.3 Multivariate Adaptive Regression Splines:

Spline helps to addresses the non-linearity issues present in stock market data. Further, this also helps to factor in interaction between various predictors. It creates various linear segments (thereby creates a hinge function) using various knots. Such linear These knots are the intersection point between two linear segments.

$$y_i = \beta_0 + \beta_1 C_1(x_i) + \beta_2 C_2(x_i) + \beta_3 C_3(x_i) + ... + \beta_d C_d(x_i) + \epsilon_i$$

In the above equation , $C_1(x)$ represents x values in between c_1 and c_2. Similarly, $C_d(x)$ represents x values between $c_(d-1)$ and c_d. In the process a search is initiated among x-values to two different linear relationship between y and x results in smallest error(Bradley and Brandon, 2020).

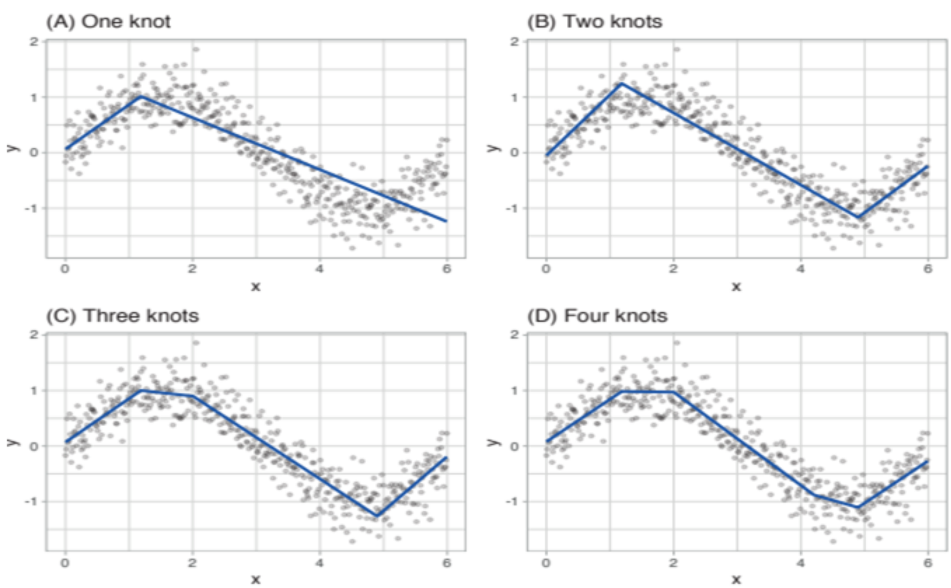

Figure 3: Multivariate Adaptive Regression Spline

3.2 Data

This study involved one broad market index CNX Nifty50 and two other strategy indices, Nifty200 Quality30 and Nifty Alpha 50. To understand the pattern of movement of constituents of the strategy indices , major constituent sectoral indices such as Nifty IT (information technology), Nifty FMCG (Fast Moving Consumer Goods), Nifty FS (Financial Services) have bee considered. And to understand the volatility of the market Indix VIX has been used.

The time frame of the study is 01/03/2019 to 28/03.2024. All the information has been taken from www.niftyindices.com , www.nseindia.com and www.finance.yahoo.com. In line on the study of Zhou et al. (2019) technical indicators like AD, ADX, ATR, NATR, ADX , ADXR, Stochastics, William's R, RSI, MFI, OBV, T3 has been considered. The standard formulae of technical indicators imply that the evaluation period generally ranges in between 14 to 26 days Colby (2003). For the purpose of the study, the standard evaluation period stands at 14 days. Further, those indicators have been estimated with 3 day, 5 day and 7 day evaluation period.

4 Results & Discussion

4.1 Preliminary Observations:

From table no 14 to table no 19 (in Appendix) show, the extent of co-movements of broader market index (i.e. Nifty50) and strategy indices over different period of prediction horizon. From the tables it is apparent that the probability of co-movement is higher for Quality30 compared to Alpha50. The probability of co-movement in the same direction of Alpha50 hovers around 75% and that for Quality30 centered around 80%. Highest probability was observed for 5 day ahead movement between Nifty50 and Quality30. The results corroborates the figure 1 which underlines the consistency of the movement in the same direction of the strategy indices in the context of movement of broad market index. As indicated Nifty Capital Goods , Nifty Financial Services are the sectoral constituents of Alpha50 index and the same for Quality30 are Nifty FMCG & Nifty IT. Table no 20 to table no. 28 summarizes the co-movement of sectoral indices with Nifty 50. The table show that the probability of co-movement is highest for Financial Services which stands at 85% . The same for IT and FMCG hovers around 71% and 72% respectively.

4.2 Prediction of direction

4.2.1 Direction prediction with standard formulae

In the context of prediction of direction of movement, following table summarizes the performance of logistic regression on Alpha50 and Quality30 with standard formulae of the technical indicators. Its also evident that as prediction horizon increased accuracy of the model decreased except for Nifty50 which stayed almost constant for all forecasting period.

	Prediction Horizon		
	1 day ahead	3 day ahead	5 day ahead
AIC - Alpha 50	1187.9	1494.9	1510.1
AIC - Quality30	807.61	1427.5	1505.6
AIC - Nifty50	1675.1	1639.7	1675.1

Table 1: Performance of logistic regression - Strategy Indices

Table no 29 to table no 31, summarizes the outputs of application of logistic regression for prediction of direction of Nifty50 , Alpha50 and Quality 30 indices. The initial observations are as follows -

- AIC values of the prediction for nifty are higher than that of Alpha50 and Quality30 indices. The AIC value for Nifty (**??** remains constant for all the entire forecasting horizon. For strategy indices , AIC value is least for 1 day ahead forecasting and its gradually increase with the length of the forecasting horizon.
- across all prediction horizons, AIC values are least for Quality30, followed by Alpha50 and Nifty50.
- the number of significant indicators of Nifty50 are lesser than the number of significant variables for Alpha50 and Quality30 indices.

4.2.2 Direction prediction - customised evaluation periods

To explore the possibilities to enhance the performance of logistic regression, the evaluation periods are tweaked to 3 day , 5 day and 7 days. The implication of changes in the evaluation period is summarised in the following tables -

		1 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	732.08	1028.3	1167
		3 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	1508	1555.2	1552.7
		5 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	1536.5	1537.3	1508.7

Table 2: Logistic regression - Performance - Nifty 50

The table clearly shows that change in the evaluation period of the technical indicators improved the performance of the models as average AIC values of all prediction horizon decreased and more number of technical indicators are found to be significant. It shows, 1 day and 3 day ahead movement of Nifty50 is best predicted with 3 day evaluation period. For 5 day ahead movement prediction, evaluation period of 7 days works the best.

Following tables summarizes the performance logistic regression for Quality30 and Alpha50.

		1 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	762.07	1522	1170
		3 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	1463.1	1522	1497.6
		5 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	1534.8	1522	1524

Table 3: Logistic regression - Performance - Nifty200 Quality 30

		1 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	703.51	1010.2	1197
		3 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	1452.9	1506.8	1518.9
		5 day ahead prediction	
	3 day data	5 day data	7 day data
AIC	1565.7	1557.3	1561.7

Table 4: Logistic regression - Performance - Nifty Alpha50

For 1 day and 3 day ahead prediction, an evaluation period of 3 days works best and for 5 day ahead prediction , 5 day based evaluation outperforms other. Such trend is applicable for both the strategy indices.

4.2.3 Performance of cross validated models

		1 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.89, 0.87	0.81, 0.76	0.78, 076
Sensitivity	0.85, 0.82	0.76, 0.66	0.69, 0.67
Specificity	0.91, 0.91	0.85, 0.85	0.91, 0.85
		3 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.71 , 0.67	0.67 , 0.66	0.66 , 0.63
Sensitivity	0.56, 0.56	0.49 , 0.48	0.49 , 0.49
Specificity	0.82 , 0.76	0.80 , 0.81	0.79 , 0.74
		5 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.69 , 0.66	0.69, 0.67	0.66, 0.65
Sensitivity	0.52 , 0.48	0.53 , 0.52	0.50 , 0.52
Specificity	0.82, 0.80	0.81 , 0.77	0.79 , 0.73

Table 5: Prediction of direction with cross validated logistic regression model - Nifty50

		1 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.79, 0.76	0.75, 0.74	0.73, 072
Sensitivity	0.68, 0.64	0.62 , 0.64	0.58, 0.58
Specificity	0.86 , 0.84	0.84, 0.81	0.84, 0.81
		3 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.70, 0.69	0.71 , 0.64	0.69 , 0.68
Sensitivity	0.44 , 0.42	0.40 , 0.35	
0.38, 0.38 Specificity	0.85 , 0.88	0.88 , 0.83	0.87 , 0.86
		5 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.68 , 0.62	0.66, 0.65	0.65, 0.65
Sensitivity	0.26, 0.27	0.22 , 0.23	0.28 , 0.24
Specificity	0.92, 0.86	0.92 , 0.92	0.88 , 0.87

Table 6: Prediction of direction with cross validated logistic regression model - Nifty Alpha50

	3 day data	5 day data	7 day data
		1 day ahead prediction	
Accuracy	0.88 , 0.86	0.83 , 0.77	0.76 , 0.78
Sensitivity	0.86 , 0.82	0.78 , 0.75	0.72 , 0.74
Specificity	0.90 , 0.89	0.87 , 0.79	0.80 , 0.82
		3 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.70 , 0.69	0.70 , 0.64	0.68 , 0.62
Sensitivity	0.60 , 0.58	0.58 , 0.50	0.53 , 0.43
Specificity	0.77 , 0.77	0.78 , 0.74	0.80 , 0.78
		5 day ahead prediction	
	3 day data	5 day data	7 day data
Accuracy	0.68 , 0.65	0.64 , 0.65	0.65 , 0.64
Sensitivity	0.51 , 0.45	0.44 , 0.46	0.40 , 0.33
Specificity	0.80 , 0.74	0.79 , 0.76	0.82 , 0.85

Table 7: Prediction of direction with cross validated logistic regression model - Nifty200 Quality30

Analysis of Nifty50: Based on accuracy of test data, it is observed that for 1 day and 3 day ahead prediction of direction, 3 day evaluation period based model provides best result. The same for 5 day ahead prediction is 5 day based evaluation period .

Analysis of strategy indices: In the line with the observations of Nifty50, for strategy indices also models based on 3 day evaluation period of technical indicators leads to highest accuracy for test data for all the prediction horizon.

Further, its evident for all the indices that sensitivity decreased across all prediction horizon. But specificity of the models for all indices has almost remained constant or marginal decreased.

4.3 Prediction of periodic return

4.3.1 Use of Regularised Regression

Following tables shows the performance of ridge regression and LASSO for all the indices for all prediction horizon. The methods have been cross-validated in the similar manner as for logistic regression. Under regularised regression Ridge regression and Least Absolute Shrinkage and Selection Operator (LASSO) have been applied to predict periodic return of the index. Root Mean Square Error (RMSE) is parameter to asses the performance of the models.

	Ridge Regression								
	1 day ahead prediction			3 day ahead prediction			5 day ahead prediction		
Train , Test	3 day data	5day data	7 day data	3 day data	5day data	7 day data	3 day data	5day data	7 day data
RMSE	0.008 , 0.013	0.001 , 0.011	0.012, 0.009	0.012, 0.019	0.020, 0.020	0.021, 0.017	0.009, 0.013	0.026 , 0.025	0.026 , 0.023
Variable Importance (First 3)	WR RSI FASTK	WR RSI FASTK	WR RSI FASTK	RSI FASTK ATR	WR RSI	WR RSI FASTK	WR RSI FASTK	VIX close RSI WR	VIX close RSI FASTK
	LASSO								
RMSE	0.008 , 0.013	0.010 , 0.011	0.012 , 0.009	0.020, 0.019	0.020, 0.02	0.021, 0.017	0.011, 0.011	0.026 , 0.025	0.025 , 0.023
Variable Importance (First 3)	RSI FASTD FASTK	WR RSI FASTK	FASTD RSI WR	RSI FASTD NATR	WR RSI	WR VIX close RSI	RSI WR FASTD	VIX close WR RSI	RSI VIX close NATR

Figure 4: Performance of Ridge and LASSO - Nifty50

Analysis of Nifty50: The error rate (RMSE) is marginally different between ridge regression and LASSO for all indices.Error rates are found to increases with length of prediction period. Both under ridge and LASSO , it is observed that for 1 day , 3 day and 5 day ahead prediction, error rate is least when evaluation period of the technical parameters is 3 day. Across all prediction horizon , Willam's R, RSI & FastK are found to be the common significant indicators governing periodic return.

	Ridge Regression								
	1 day ahead prediction			3 day ahead prediction			5 day ahead prediction		
Train , Test	3 day data	5day data	7 day data	3 day data	5day data	7 day data	3 day data	5day data	7 day data
RMSE	0.008 , 0.008	0.008 , 0.010	0.10 , 0.10	0.017 , 0.17	0.017 , 0.17	0.016 , 0.017	0.023 , 0.022	0.024 , 0.019	0.022 , 0.024
Variable Importance (First 3)	wr , rsi fastK	wr , rsi fastK	wr fastK rsi	wr rsi fastK	wr rsi fastK	wr rsi fastD	wr rsi fastK	wr rsi fastK	wr rsi fastK
	LASSO								
RMSE	0.008 , 0.008	0.008 , 0.009	0.010 , 0.009	0.017 , 0.017	0.017 , 0.017	0.017 , 0.017	0.023 , 0.021	0.023 , 0.020	0.021 , 0.023
Variable Importance (First 3)	wr	wr fastD	wr	wr	wr	wr	wr	wr rsi natr	rsi, fastD wr

Figure 5: Performance of Ridge and LASSO - Quality30

	Ridge Regression								
	1 day ahead prediction			3 day ahead prediction			5 day ahead prediction		
Train , Test	3 day data	5day data	7 day data	3 day data	5day data	7 day data	3 day data	5day data	7 day data
RMSE	0.014 , 0.013	0.013 , 0.012	0.013 , 0.012	0.010 , 0.015	0.013 , 0.012	0.011 , 0.015	0.031,, 0.034	0.032 , 0.029	0.034, 0.030
Variable Importance (First 3)	NATR MFI VIX close	WR RSI FastK	WR RSI FastK	RSI FastK SlowD	WR RSI FastK	WR RSI FastD	VIX close RSI FAstK	VIX close RSI WR	VIX close WR RSI
	LASSO								
RMSE	0.0.14 , 0.013	0.013, 0.013	0.013 , 0.12	0.009 , 0.014	0.012, 0.012	0.011 , 0.015	0.031, 0.034	0.032 , 0.029	0.031 , 0.030
Variable Importance (First 3)	SlowD RSI NATR	WR	WR RSI	FastD fastK RSI	WR RSI SlowD	WR RSI FastD	RSI VIX close	VIX close NATR RSI	VIX close NATR WR

Figure 6: Performance of Ridge and LASSO - Alpha50

Analysis of strategy indices: The RMSE, for Ridge and LASSO for strategy indices are marginally different. Considering the error rate of test data, it is observed that for 1 day and 3 day ahead prediction , models with 3 day evaluation period performs better. For 5 day ahead prediction , 5 day based evaluation period shows superior performance. For Quality 30 index, William's R, RSI and Fast K are top three predictors of periodic return using ridge regression. With respect to LASSO, in most cases the number of influential predictors is limited only one i.e. William's R .

The periodic return of Alpha50 is majorly governed by William's R, RSI , Fast K in all cases. Considering RMSE of test data, it can be seen that 1 day ahead return is not sensitive to evaluation period. 3 day ahead return can be better estimated when indicators are estimated on the basis of 5 day. An evaluation period of 7 days better predict 5 day ahead return.

4.3.2 Use of Multiple Adaptive Regression Spline

Spline not only captures the non-linearity aspect of the data , but it also captures the interaction among the independent variables. For the current study a hypergrid has been designed to search for best suited models. The hypergrid allows to address from linear to cubic relations, from 2 to 100 terms.

	Spline								
	1 day ahead prediction			3 day ahead prediction			5 day ahead prediction		
Train , Test	3 day data	5day data	7 day data	3 day data	5day data	7 day data	3 day data	5day data	7 day data
RMSE	0.012 , 0.012	0.012 , 0.012	0.011 , 0.015	0.017 , 0.022	0.017 , 0.023	0.017 , 0.025	0.018 , 0.035	0.019 , 0.034	0.019 , 0.034
Best Prune Model	nprune - 12 degree - 1	nprune - 2 degree - 1	nprune - 12 degree - 1	nprune - 23 degree - 2	nprune - 23 degree - 3	nprune - 23 degree - 1	nprune - 23, degree 2	nrprune - 34 degree - 2	nprune - 23 degree - 3
Variable Importance (First 3)	NATR T3 VIX close	NATR T3 VIX close	VIX close T3 NATR	VIX close BB_UP AD	VIX close BB_UP AD	VIX close T3 NATR	VIX close T3 OBV	VIX close ADXR AD	VIX close BB_UP NATR

Figure 7: Performance of Spline - Nifty50

Analysis of Nifty50: Based on the error rate of the test dataset, 1 day and 3 day ahead returns can be better predicted when LASSO is used with the parameters evaluated over 3 days. 5 day ahead return can be better predicted using 5 day evaluation period. The important factors governing periodic return are very different that of ridge regression. It can be observed that VIX close, T3 and NATR are the important predictors across all prediction horizon. Moreover, linear models with 2 to 12 terms can be used to predict 1 day ahead return. The best pruned models to predict 3 day and 5 day ahead return are mostly quadratic or cubic in nature with 23 to 32 terms.

	Spline								
	1 day ahead prediction			3 day ahead prediction			5 day ahead prediction		
Train , Test	3 day data	5day data	7 day data	3 day data	5day data	7 day data	3 day data	5day data	7 day data
RMSE	0.012 , 0.013	0.008 , 0.011	0.009 , 0.011	0.02 , 0.022	0.020 , 0.022	0.021 , 0.021	0.026 , 0.028	0.029 , 0.026	0.024 , 0.031
Best Prune Model	nprune - 2 degree - 1	nprune - 23 degree - 3	nprune - 23 degree - 2	nprune - 23 degree - 1	nprune - 23 degree - 1	nprune - 23 degree - 2	nprune - 23 degree - 1	nprune - 23 degree - 2	nprune - 23 degree - 3
Variable Importance (First 3)		RSI ATR FASTD	WR RSI NATR	RSI VIX close T3	RSI WR BB_UP	RSI NATR BB_UP	VIX close BB_UP RSI	VIX close T3 RSI	VIX close ADX RSI

Figure 8: Performance of Spline - Alpha50

Analysis of strategy indices: For Alpha50, being a momentum based strategy, 1 day, 3 day and 5 day ahead return can best predicted using 5 days evaluation period. RSI , VIX close , ATR or NATR governs the periodic return for Alpha50. But, no significant predictor of 1 day ahead return can be identified. Most best pruned models have 23 terms and approximately 44% models are linear.

The error rate is least for Quality 30 compared other two indices. 1 day and 3 day ahead returns can be better predicted using 3 day evaluation period. And 5 day ahead return can

	Spline								
	1 day ahead prediction			3 day ahead prediction			5 day ahead prediction		
Train , Test	3 day data	5day data	7 day data	3 day data	5day data	7 day data	3 day data	5day data	7 day data
RMSE	0.005 , 0.005	0.004 , 0.006	0.009 , 0.008	0.014 , 0.014	0.014 , 0.016	0.014, 0.016	0.018 , 0.020	0.018 , 0.020	0.017 , 0.019
Best Prune Model	nprune - 12 degree - 3	nprune - 34 degree = 3	nprune - 2 degree - 2	nprune - 12 degree - 1	nprune - 23 degree - 1	nprune - 23 degree - 1	nprune - 23 degree - 1	nprune - 23 degree - 1	nprune - 23 degree - 1
Variable Importance (First 3)	wr natr rsi	wr atr mfi	NA	t3 natr fastK	fastD t3 natr	wr rsi bb_up	t3 natr vix_close	rsi natr mfi	bb_up ad wr

Figure 9: Performance of Spline - Quality30

be better predicted using 7 day evaluation period. William's R, NATR / ATR, T3 are the most common predictors across all prediction horizon. Most of the best performing models have 23 terms and 67% of the models are linear.

4.4 Comparison among Ridge , LASSO & Spline

There is no significant difference in the performance of ridge and LASSO with respect to prediction of periodic return. RMSE of test data of Nifty 50 is least for ridge regression , followed by LASSO and Spline across all prediction period. Momentum indicators have been found to be influential predictors of periodic return in the case of regularised regression (i.e. ridge and LASSO). But for the adaptive regression , But in case Spline, its volatility indicators like average true range (ATR) / normalised true range (NATR), VIX close and Bollinger band (upper) and T3 (a smoothening variable) found to be influential.

For Alpha50, Spline outperforms Ridge and LASSO to estimate return over 1 day and 5 days. But, Ridge and LASSO outperforms to predict return of 3 days. For Quality 30, spline outperforms ridge and LASSO for all prediction horizon.

Further, momentum based indicators such as William's R, RSI, Fast K, Fast D dominate the performance of ridge and LASSO for all indices. In case of Spline, periodic return of Nifty 50 is governed by volatility based indicators as such VIX close, NATR , BB-UP and a smoothening variable T3.

But for the strategy indices its a mix of momentum and volatility based indicators which governs the periodic return of the indices.

4.5 Estimation of return

To estimate the periodic return, a buy side trading rule has been developed based on the fitted value of logistic regression of respective models. If the fitted value of an observation is greater than or equal to 0.50 , a buy signal is triggered for the respective prediction period (i.e. 1 day, 3day or 5 day).

All the variants of the model (i.e. with standard evaluation period and also under 3 day , 5 day and 7 day evaluation periods) have been considered for the comparison. Following tables summarises the performance of the indices -

	Prediction		
	1 day ahead	3 day ahead	5 day ahead
Cum.Ret	0.644	1.667	-94

Table 8: Cumulative Return Nifty50: Under standard evaluation period

	1 day ahead prediction		
	3 day data	5 day data	7 day data
Cum.Ret	0.405	0.711	-0.754
	3 day ahead prediction		
	3 day data	5 day data	7 day data
Cum.Ret	1.721	1.478	1.348
	5 day ahead prediction		
	3 day data	5 day data	7 day data
Cum.Ret	2.444	2.970	2.893

Table 9: Cumulative return - Nifty50: Under different evaluation periods

Its evident from above tables the buy signal triggered under standard evaluation period seriously under perform as the cumulative return of Buy & Hold (henceforth B&H) strategy is 200%. When Nifty50 is traded based on a trading signal of 5 day evaluation period, it generated a cumulative return of 297%.

	Prediction		
	1 day ahead	3 day ahead	5 day ahead
Cum.Ret	0.817	2.781	5.603

Table 10: Cumulative Return Alpha50: Under standard evaluation period

	1 day ahead prediction		
	3 day data	5 day data	7 day data
Cum.Ret	0.657	1.385	1.09
	3 day ahead prediction		
	3 day data	5 day data	7 day data
Cum.Ret	3.048	4.05	3.133
	5 day ahead prediction		
	3 day data	5 day data	7 day data
Cum.Ret	6.077	6.473	5.859

Table 11: Cumulative return - Alpha50: Under different evaluation periods

The cumulative return under B& H for Alpha50 is 300%. It generates a maximum cumulative return of 647% when traded for 5 day period based 5 day evaluation period.

	Prediction		
	1 day ahead	3 day ahead	5 day ahead
Cum.Ret	0.543	1.342	2.703

Table 12: Cumulative Return Quality30: Under standard evaluation period

		1 day ahead prediction	
	3 day data	5 day data	7 day data
Cum.Ret	0.777	0.687	0.346
		3 day ahead prediction	
	3 day data	5 day data	7 day data
Cum.Ret	1.832	2.095	0.808
		5 day ahead prediction	
	3 day data	5 day data	7 day data
Cum.Ret	2.411	3.491	2.423

Table 13: Cumulative return - Quality30: Under different evaluation periods

The cumulative return on B& H strategy over the period for Quality30 is 95% . And it generates a maximum cumulative return of 349% when traded for a 5 day period based on 5 day evaluation period. Therefore, 5 day emerges as the most profitable trading period . But the choice of evaluation period differs, for Nifty50 its 7 days , for Alpha50 and Quality30 it is 5 days.

Incorporating trend of Nifty50: As indicated in figure2 direction of the strategy indicator can be predicted based technical indicators alone. And in the second approach, along with technical indicators return of Nifty50 can also be Incorporated. It is already underlined that both the strategy indices have generated highest yield for 5 day trading period (i.e. prediction period) based on 5 day evaluation period of the technical indicators.

When buy signals are generated for the best performing models after incorporating nifty return , it is is found that the cumulative return stands at 349% for Quality30 and the same for Alpha50 stands at 525% (approx). Both the are less than cumulative returns generated by the stand alone models.

5 Conclusion

Table no. 1 to 4 shows that 1 day and 3 day ahead forecasting of direction of all indices can be best performed with 3 day evaluation period. For 5 day ahead prediction, 5 day based evaluation suits the best for Quality 30 and Alpha50. Evaluation period of 7 days suits Nifty 50.

In line with forecasting of direction, it is seen that for 1 day and 3 day ahead returns are best predicted when technical parameters are evaluated over 3 days. This is applicable for all indices. For Nifty 50 and Alpha 50, it is 7 days of evaluation period which outperforms others for 5 day ahead prediction. And for Quality 30 , 5 days based evaluation suits the best. With respect to performance of methods Spline outperformed the rst for Nifty 50 and Alpha 50. Moreover, Vix close, T3 and important predictor of periodic return of NIfty 50.

Table no. 9 to 11 shows that cumulative return estimated based on 14 day period generates considerably less return compared to return generated by B& H strategy. It also clearly evident as the evaluation period is shortened the probability of outperforming both return from B&H strategy as well as return of Nifty50 increases.

In case of Alpha50, the probability of outperforming both return of Nifty 50 and the return from B & H strategy of Alpha50 is 67%. But for Quality30, the probability of outperforming Nifty50 is 44% (approx) and the probability to outperform the return from B& H strategy of Quality 30 is 55% (approx).

Further, the pattern of specificity and sensitivity for all indices shows that logistic regression is more efficient in capturing the downward movement compared to upward movement. This corroborated by the fact as sensitivity consistently declined and specificity remained almost constant for all models.

Thus, if retail traders can factor in the above issues while placing their orders, even with simplistic models like logistic regression and few technical indicators can save them from the existing precarious position and unnecessary losses as reflected in SEBI reports.

6 Appendices

6.1 Tables

	Nifty50	
Alpha50	up	down
up	375	184
down	119	582

Table 14: Nifty50 & Alpha50: 1 day ahead movement: Accuracy 75.95%

	Nifty50	
Alpha50	up	down
up	360	185
down	116	597

Table 15: Nifty50 & Alpha50: 3 day ahead movement: Accuracy 76.07%

	Nifty50	
Alpha50	up	down
up	348	182
down	120	606

Table 16: Nifty50 & Alpha50: 5 day ahead movement: Accuracy 75.96%

	Nifty50	
Quality30	up	down
up	431	128
down	135	563

Table 17: Nifty50 & Quality30: 1 day ahead movement: Accuracy 79.08%

	Nifty50	
Quality30	up	down
up	424	121
down	118	592

Table 18: Nifty50 & Quality30: 3 day ahead movement: Accuracy 80.96%

	Nifty50	
Quality30	up	down
up	412	118
down	100	623

Table 19: Nifty50 & Quality30: 5 day ahead movement: Accuracy 82.6%

	Nifty50	
IT	up	down
up	398	161
down	197	501

Table 20: Nifty50 & Nifty IT: 1 day ahead movement: Accuracy 71.52%

	Nifty50	
IT	up	down
up	376	169
down	173	537

Table 21: Nifty50 & Nifty IT: 3 day ahead movement: Accuracy 72.75%

	Nifty50	
IT	up	down
up	355	175
down	179	544

Table 22: Nifty50 & Nifty IT: 5 day ahead movement: Accuracy 71.75%

	Nifty50	
FS	up	down
up	482	77
down	106	592

Table 23: Nifty50 & Nifty FS (Financial Services) : 1 day ahead movement: Accuracy 85.44%

	Nifty50	
FS	up	down
up	447	98
down	91	619

Table 24: Nifty50 & Nifty FS (Financial Services) : 3 day ahead movement: Accuracy 84.94%

	Nifty50	
FS	up	down
up	455	75
down	185	618

Table 25: Nifty50 & Nifty FS (Financial Services) : 5 day ahead movement: Accuracy 85.63%

	Nifty50	
FMCG	up	down
up	400	159
down	195	503

Table 26: Nifty50 & Nifty FMCG: 1 day ahead movement: Accuracy 71.84%

	Nifty50	
FMCG	up	down
up	394	151
down	200	510

Table 27: Nifty50 & Nifty FMCG: 3 day ahead movement: Accuracy 72.03%

	Nifty50	
FMCG	up	down
up	392	138
down	190	533

Table 28: Nifty50 & Nifty FMCG: 5 day ahead movement: Accuracy 73.82%

AIC	1675.1	1639.7	1675.1
		Significance levels	
ad	0.10	0.05	
adr			0.10
natr	0.01	0.10	
bb-up	0.10		
mfi		0.05	
obv		0.001	
wr		0.05	
vix-close	0.05	0.001	0.05

Table 29: Significant technical indicators for Nifty50: 1 , 3 and 5 days ahead prediction with standard formulae

AIC	1187.9	1494.9	1510.1
		Significance levels	
ad		0.05	0.001
adx	0.05		
adxr		0.05	
natr	0.1	0.01	0.05
bb-up	0.001		0.01
t3	0.001		0.05
wr	0.001	0.01	0.1
fastD	0.001	0.001	0.05
slowD	0.01		
rsi	0.001	0.001	0.001
vix-close		0.05	0.1

Table 30: Significant technical indicators for Alpha50: 1 , 3 and 5 days ahead prediction with standard formulae

AIC	808.28	1429.6	1507.9
		Significance levels	
ad			0.01
adx		0.05	0.001
adxr		0.05	
natr			0.01
t3	0.01		0.01
wr	0.001	0.001	0.001
fastD	0.001	0.001	0.05
slowD	0.001	0.01	
rsi		0.001	0.001
vix-close		0.001	0.001

Table 31: Significant technical indicators for Quality30: 1 , 3 and 5 days ahead prediction with standard formulae

Definitions of technical indicators: In the following section standard formulae of various technical indicators have been provided. For several indicators there is a choice offered by creators of these indicators for their estimation i.e. evaluation period. In such cases , in this study the number days has been kept at 14 days as the study is on short term prediction. Colby (2003).

- Average True Range (ATR) - EMA of 27 days of Max (highest price of current perid - lowest price), (highest price - closing price of previous period), (closing price of previous period - lowest price of current period)

- Normalised Average True Range (NATR)

  NATR = Average True Range / closing price * 100

- Average Directional Index (ADX) - Positive Direction Movement (PDM) = (highest price of the current period - highest price of the previous period)

  Negative Direction Movement (NDM) = Lowest price of previous period - lowest of current period)

  Positive Direction Index (PDI) = EMA of 14 days of (PDM/True Range) Negative Direction Index (NDI) EMA of 14 days of (MDM/True Range) DX = 100 * (PDI - MDI) / (PDI + MDI)

  ADX = EMA of 14 days of AD

- Average Directional Index Rating (ADXR) ADXR = (ADX + ADX of 14 days ago)/2

- Stochastics (FastD, Slow D, FastK)
  K = 100 * $\frac{(latestclosingprice - lowestpricein14days}{Highestpricein14days - lowestpricein14days}$
  Fast k = current value of K
  Fast D = 3 period moving average of K
  Slow D = 3 period moving average of Fast D

- William's %R (WR)

$$WR = \frac{(highestin14days - closingprice)}{(Highestin14days - Lowestin14days)}$$

- Relative Strength Index (RSI)
  RS = (EMA of 14 days of gains) / (EMA of 14 days of losses) RSI = 100 - (100/(1 + RS))

- Accumulation Distribution Indicator (AD)
  $AD = \sum(\frac{(dailyclosingprice - dailylowestprice) - (dailyhighestprice - dailyclosingprice))}{dailyhighestprice - dailylowestprice} * Volume$

- Bollinger Bands (UP) = SMA of 20 days + 2 * standard deviation

- Money Flow Index (MFI)

  Typical price = (daily highest price + daily closing price + daily lowest price)/3

  Raw money flow = Typical price * volume

  Money ratio = Raw money flow for up period / Raw money flow for down period , period of 10 days

  MFI = 100 - 100/(1 + Money Ratio)

- On Balance Volume (OBV) OBV = Previous OBV + current volume when price is increasing OBV = Previous OBV - current volume when price is decreasing

- T3 - Triple Exponential Moving Average

  EMA 1 = EMA of 10 days of closing price EMA2 = EMA of 10 days of EMA1 EMA3 = EMA of 10 days of EMA2

  T3 = 3 * EMA1 - 2 * EMA2 + EMA3

6.2 Abbreviations

- BBO - Biography based optimization
- CSFLA - Continuous shuffled frog leaping algorithm
- CSO - Cat swarm optimization
- DBN - Deep belief network
- FLNN - Functional link neural network
- GRU - Gated recurrent unit
- GA - Genetic algorithm
- RBF - Radial basis function, ,

- MLP - Multi layer perceptron
- PSO - Particle swarm optimization,
- PLR - Piece-wise linear regression
- RNN - Recurrent neural network
- SVM - Support vector machine
- SQNN - Spiking quantun neural network
- SEBI - Security Exchange Board of India
- WSVM - Weighted support vector machine

Computation: For computation of of all technical indicators R-software (version 4.2.1) has been used. Some of the other packages involved in computing and implementing the models are 'TTR (version 0.24.4)', 'earth (version - 5.3.3)', 'caret (version - 6.0 -94)', 'glmnet (version - 4.1 -8)' and 'vip (version - 0.4.1)'.

References

Alsubaie, Y., El Hindi, K., and Alsalman, H. (2019). Cost-sensitive prediction of stock price direction: Selection of technical indicators. *IEEE Access*, 7:146876–146892.

Beyaz, E., Tekiner, F., Zeng, X.-j., and Keane, J. (2018). Comparing technical and fundamental indicators in stock price forecasting. In *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pages 1607–1613. IEEE.

Bradley, B. and Brandon, G. (2020). *Hands-On Machine Learning with R.* CRC Press, Boca Raton, FL, USA.

Chandar, K. S. and Punjabi, H. (2021). Cat swarm optimization algorithm tuned multilayer perceptron for stock price prediction. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, 17(7):1–15.

Chen, C.-H., Shih, P., Srivastava, G., Hung, S.-T., and Lin, J. C.-W. (2021). Evolutionary trading signal prediction model optimization based on chinese news and technical indicators in the internet of things. *IEEE Internet of Things Journal.*

Chen, X., Hirota, K., Dai, Y., and Jia, Z. (2023). A model fusion method based on multi-source heterogeneous data for stock trading signal prediction. *Soft Computing*, 27(10):6587–6611.

Colby, R. W. (2003). *The encyclopedia of technical market indicators.* McGraw-Hill.

Faghihi Nezhad, M. and Minaei Bidgoli, B. (2021). Development of an ensemble learning-based intelligent model for stock market forecasting. *Scientia Iranica*, 28(1):395–411.

Gandhmal, D. P. and Kumar, K. (2019). Systematic analysis and review of stock market prediction techniques. *Computer Science Review*, 34:100190.

Joginpally, B. (2023). An intelligent soft computing framework for forecasting stock movement direction in international stock markets. *Journal of Theoretical and Applied Information Technology*, 101(9).

Kakushadze, Z., Serur, J. A., et al. (2018). *151 Trading Strategies*. Springer.

Kampouridis, M. and Otero, F. E. (2017). Evolving trading strategies using directional changes. *Expert Systems with Applications*, 73:145–160.

Kilimci, Z. H. and Duvar, R. (2020). 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 (bist 100). *IEEE Access*, 8:188186–188198.

Lv, D., Yuan, S., Li, M., Xiang, Y., et al. (2019). An empirical study of machine learning algorithms for stock daily trading strategy. *Mathematical problems in engineering*, 2019.

Nazário, R. T. F., e Silva, J. L., Sobreiro, V. A., and Kimura, H. (2017). A literature review of technical analysis on stock markets. *The Quarterly Review of Economics and Finance*, 66:115–126.

Ohaegbulem, E. U. and Kalu, O. C. (2023). Using logistic regression to predict the direction of nigeria stock market returns (based on some selected sector indexes, 2015-2022). *European Journal of Theoretical and Applied Sciences*, 1(4):1180–1200.

Ouyang, H., Wei, X., and Wu, Q. (2020). Discovery and prediction of stock index pattern via three-stage architecture of ticc, tpa-lstm and multivariate lstm-fcns. *IEEE Access*, 8:123683–123700.

Palsma, J. and Adegboye, A. (2019). Optimising directional changes trading strategies with different algorithms. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 3333–3340. IEEE.

Ren, R., Wu, D. D., and Liu, T. (2018). Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Systems Journal*, 13(1):760–770.

Tan, Z., Liu, J., and Chen, J. (2021). Detecting stock market turning points using wavelet leaders method. *Physica A: Statistical Mechanics and its Applications*, 565:125560.

Tang, H., Dong, P., and Shi, Y. (2019). A new approach of integrating piecewise linear representation and weighted support vector machine for forecasting stock turning points. *Applied Soft Computing*, 78:685–696.

Yao, Y., Cai, S., and Wang, H. (2021). Are technical indicators helpful to investors in china's stock market? a study based on some distribution forecast models and their combinations. *Economic Research-Ekonomska Istraživanja*, pages 1–25.

Zhou, T., Li, X., and Wang, P. (2021). Statistics and practice on the trend's reversal and turning points of chinese stock indices based on gann's time theory and solar terms effect. *Mathematics*, 9(15):1713.

Zhou, Z.-H., Yu, Y., and Qian, C. (2019). *Evolutionary learning: Advances in theories and algorithms*. Springer.