ISSN No. 0976-5697

Volume 8, No. 3, March – April 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Decoding Investment Pattern of FIIs and DIIs in Indian Stock Market using Decision Tree

Dr. Manminder Singh Saluja*
Assistant Professor (Senior Scale)
IIPS, Devi Ahilya University
Indore, India

Dr. Yasmin Shaikh
Assistant Professor (Senior Scale)
IIPS, Devi Ahilya University
Indore, India

Abstract: Investing in stock market has always been a riskier venture. Market participants have always tried to correctly time the market to make more money. It is not only about the timing in the market but also about a correct decision to buy or sell the stock. Big institutional players had always been more successful in making money compared to the smaller ones like retail investors. The main objective of this paper is to decode the investment pattern of these big payers like Foreign Institutional Investors (FIIs) and Domestic Institutional Investors (DIIs) using decision tree method of machine learning techniques. Using J48 technique of C4.5 classification program it was found that FIIs have more predictable and correct investment pattern compared to DIIs. The information gain ratio made *High* attribute as the prime node of the tree for both FIIs and DIIs with accuracy level of 66.17% and 59.35% respectively. FIIs are the net buyers if *High* and *Close* attributes are positive and net sellers if *High* and *Low* are negative. However, DIIs ironically are the net buyers with a negative *High*.

Keywords: Decision Tree, Stock Market, FIIs, DIIs, Investment, Machine Learning.

I. INTRODUCTION

Stock market is a place where transfer of existing financial ownership takes place. This transfer is done at an agreed price between buyer and seller, which is prevailing at the time of transaction in the market. Interested buyers always gains on buying at a lower price and seller gains by selling at a higher price. Before deciding on the price at which the stock should either be bought or sold, it is more important to decide whether at that moment on that particular day should the stocks be bought or sold. Even if an investor buy a stock which though is less than his expected price but if market falls from that point and takes a very long time to recover and again reach to that particular transacted value, it will not only stuck the capital of the investor but he will also loose on the returns. It may also happen, that an impatient investor sells that stock in loss. This is actually, what happens with most of the retail participants in the market. However, there are players in the market who not only gain in the market but also decide the direction of the market. These are Foreign Institutional Investors (FIIs) and Domestic Institutional Investors (DIIs). These players are the professional participants of the market with their prime business of earning from price movements in stock market. These players hold a very big amount of liquidity creating the ability to take big or multiple positions in the market and thus changing the pace of the market. The investment pattern of these players if decoded can help retail and High Net worth Investors (HNIs) to take more prudent and wise decision in the market.

Recently, machine learning techniques have been applied on historical stock market data to extract useful patterns from available data. These patterns are then analyzed to predict the future patterns and trends in stock market.

In this paper, NIFTY 50 index data is analyzed to predict the buying and selling behavior of FIIs and DIIs. The NIFTY 50 index is a major stock market index for Indian equity market. It covers 13 sectors of the Indian economy and offers investment managers exposure to the Indian market in one portfolio. The data is analyzed using decision tree classifier. The objective of using this classification technique is to reveal the behavioral patterns of institutional investors in Indian stock market.

The remainder of paper is organized into four sections. Section II contains the literature review about research on stock market, institutional investors, and use of data mining and machine learning techniques in predicting stock market behavior. Section III includes details of data and methodology used to generate the model for classification. Section IV presents the results, performance evaluation of model being generated, and discussion on results. Finally a brief conclusion is included in Section V.

II. RELATED WORK

Prediction of stock market behavior is a subject of interest for researchers all over the world. Since the stock market behavior is governed by many factors such as interest rates, economic outlook, inflation, deflation, economic and political shocks, changes' in economic policies by government, and investors' sentiments and so on. None of the available literature on stock market predictions claims to predict the stock market behavior accurately. However, the price discovery mainly happens with volume based buying and selling that is undertaken by institutional players like FIIs and DIIs in the market. Globally FIIs have always emerged as big players in shaping different markets. In developing countries including India, there has been increased liberalization of domestic financial and capital market and an opening up of the markets to FIIs [1]. The main emerging feature of Indian equity market since 1991 is its gradual integration with global market. FIIs played a major role in shaping the 'equity price movement' in India since 1991. The role of the institutional investors, both domestic and foreign was found in driving the return on the Indian equity market in the last decade [2]. The interaction of foreign institutional investments with market returns and market volatility in India using both static and dynamic models based on daily data are investigated [3]. The findings of both models show foreign investors as positive feedback traders when investing in the Indian market and as negative feedback traders during their withdrawal. Using the impulse response functions based on vector auto regression, strong evidence is found that foreign institutional investments destabilize the market, particularly with selling activities, as they significantly increase the volatility.

A descriptive study has been done to investigate about the relationship between FIIs and Stock Market Volatility by finding the impact of FIIs on the market capitalization and turnover of National Stock Exchange of India (NSE) [4]. The result of the study suggests that FIIs have a positive impact on the performance. It also concluded that FIIs are responsible for the volatility of Indian Stock Market.

With the advent of the digital computer, stock market prediction has moved into the technological realm. The most prominent technique used to predict stock market behavior involves the use of classification methods. Classification methods include neural network, decision tree, naive Bayes k-nearest neighbors and many more.

Inductive machine-learning classifiers including artificial neural network, decision tree, and k-nearest neighbor wereused to predict average index of stock market [5]. Through appropriate collaboration of these models, prediction accuracy up to 65 percent was achieved. In [6], SVM, ANN, GA-SVM, and GA-ANN machine learning algorithms were used in a standalone andintegrated form to predict the direction of stock market with fundamental, technical and hybrid methods. In [7],a study is conducted to help the investors in the stock market to decide the better timing for buying or selling stocks based on the knowledge extracted from the historical prices of such stocks. Decision tree classifier was used to take the appropriate decision based on historical data. An approach for predicting future market direction based on chart patterns recognition by using data mining classification was proposed [8]. The model aimed to predict whether the market index will go up, go down or stay in the next 1, 6 and 21 days.

Several data mining techniques have been applied for predicting and detecting the behavior of stock market prices, index, and behavior of investors. In the current study decision tree classification technique is used to detect the behavior of FIIs and DIIs in Indian stock market. Decision tree technique is selected since the accuracy achieved with decision tree is far better than other machine learning classifiers. In a study, the decision tree model outperformed the Support Vector Machine and the Neural Network methods[9]. According to a 10-fold cross validation evaluation, decision tree achieved average accuracy rate of 83.73%. The performances of the quadratic polynomial Support Vector Machine and the Neural Network were 79.29% and 75.44% respectively [10]. The performance of different types of decision trees was considerably better than other binary classifiers.

There is a vast gap in the literature on the relationship of DIIs with market in India. Even there was hardly any research, which can predict or decode the investment pattern of institutional players in the market. The above-mentioned literature clearly highlights a wide gap in detection of investment behavior of institutional players in the market. Therefore, the objective of this research is to identify the attributes for analyzing the investment pattern of FIIs & DIIs and to decode the investment pattern on the basis of historical data.

III. DATA AND METHODOLOGY

A. *Data Selection*

The study is done on secondary data collected from NSE, NSDL and SEBI. The data on six attributes namely *Open*, *High*, Low, *Close*, *Volatility*, and *Action* was taken for three-year period ranging from 1st April 2014 to 24th March 2017 making in all 733 working days of the market. The data of O*pen*, *High*, *Low* and *Close* is taken for NIFTY 50 Index, which represents the average buying and selling price of top 50 listed Indian companies by market capitalization on National Stock Exchange of India (NSE). Market volatility is measured by the data of INDIAN VIX index – volatility index of NSE for Indian markets. The *Action* attribute reflects the buying and selling behavior of FIIs and DIIs measured through their daily net position in the market.

Table I presents the sample data for the above mentioned attributes from 733 observations ranging from 1st April 2014 to 24th March 2017. On 1st April 2014 NIFTY opened at 6729.5 touching a *High* of 6732.2 falling 57 points from its peak to make a *Low* of 6675.4 before recovering and closing at 6721. The volatility index on this day fell by 4.34 % compared to previous days' value. FIIs on the day were net buyer to a tune of 385.66 cr. while DIIs were net sellers of 247.88 cr.

Table I. Sample of Historical Net Investment of FIIs and DIIs in Cash Segment of Indian Indices.

Date	Open	High	Low	Close	% Change in India VIX Index	FII Net Buying/ Selling (in₹ cr.)	DII Net Buying/ Selling (in ₹cr.)
01-04-14	6729.5	6732.2	6675.4	6721.0	-4.34	385.66	-247.88
02-04-14	6757.6	6763.5	6723.6	6752.5	2.55	594.67	-471.78
03-04-14	6772.0	6776.7	6696.9	6736.1	4.22	717.39	-716.57
27-03-15	8396	8413.2	8269.1	8341.4	-4.37	-320.52	674.76
30-03-15	8390.9	8504.5	8380.7	8492.3	-1.64	-240.34	651.67
31-03-15	8527.6	8550.4	8454.1	8491	1.61	356.07	283.71
29-02-16	7050.4	7094.6	6825.8	6987.0	-7.81	-2018.02	1445.2
01-03-16	7038.2	7235.5	7035.1	7222.3	-7.37	1760.98	317.02
02-03-16	7321.7	7380.3	7308.1	7368.8	1.7	1437.5	-593.67
22-03-17	9047.2	9072.9	9019.3	9030.4	2.74	356.64	-779.91
23-03-17	9048.7	9099.0	9048.6	9086.3	-3.55	1094.44	-590.78
24-03-17	9104	9133.5	9089.4	9108	1.2	543.35	116.5

B. *Data Preprocessing*

Since the data presented in table I is numeric but the decision tree using J48 requires a categorical data, table II therefore presents thesample of the categorical data that was converted from numeric data using following rules:

1) Firstly, the returns for the four attributes namely *Open*, *High*, *Low* and *Close* were calculated from the given data.

2) If the return on opening price is greater than zero, it is earning positive returns; thereby 'POSITIVE' value was assigned to *Open* attribute and if return on opening price is less than zero 'NEGATIVE' value was assigned to this attribute.

3) Similarly, POSITIVE and NEGATIVE values were assigned to other three attributes namely *High*, *Low* and C*lose*.

4) In case of *volatility* attribute, if the percentage change in INDIA VIX index was more than zero, 'MORE VOLATILITY' value was assigned and if this percentage change was less than zero, 'LESS VOLATILITY' value was assigned to the *Volatility* attribute.

5) Action attribute was created by assigning 'BUY' or 'SELL' value. If FIIs or DIIs on the *t* or current day were net buyers they were given 'BUY' value and if they were net sellers they were assigned 'SELL' value.

Table II. Sample of Categorical Data after Preprocessing the Research Data

Date	Open	High	Low	Close	Volatility	FII Action	DII Action
01-04-14	POSITIVE	POSITIVE	POSITIVE	POSITIVE	Less Volatility	Buy	Sell
02-04-14	POSITIVE	POSITIVE	POSITIVE	POSITIVE	More Volatility	Buy	Sell
03-04-14	POSITIVE	POSITIVE	NEGATIVE	NEGATIVE	More Volatility	Buy	Sell
27-03-15	NEGATIVE	NEGATIVE	NEGATIVE	NEGATIVE	Less Volatility	Sell	Buy
30-03-15	NEGATIVE	POSITIVE	POSITIVE	POSITIVE	Less Volatility	Sell	Buy
31-03-15	POSITIVE	POSITIVE	POSITIVE	NEGATIVE	More Volatility	Buy	Buy
29-02-16	POSITIVE	POSITIVE	NEGATIVE	NEGATIVE	Less Volatility	Sell	Buy
01-03-16	NEGATIVE	POSITIVE	POSITIVE	POSITIVE	Less Volatility	Buy	Buy
02-03-16	POSITIVE	POSITIVE	POSITIVE	POSITIVE	More Volatility	Buy	Sell
22-03-17	NEGATIVE	NEGATIVE	NEGATIVE	NEGATIVE	More Volatility	Buy	Sell
23-03-17	POSITIVE	POSITIVE	POSITIVE	POSITIVE	Less Volatility	Buy	Sell
24-03-17	POSITIVE	POSITIVE	POSITIVE	POSITIVE	More Volatility	Buy	Buy

C. *Decision Tree Induction*

A decision tree is a classification scheme that generates a tree and a set of rules, representing the model of different classes from a given data set [11]. The set of records available for developing classification method is generally divided into two disjoint subsets – a *training set* and a *test set*. Training set is used to build the classifier/model while test set is used to measure the accuracy of classifier. The accuracy of classifier is determined by the percentage of test examples that are correctly classified.

In this study, decision tree for data under analysis is generated using J48, which is the implementation of C4.5 classification program [12]. This classification is done using Weka software [13]. To generate a decision tree, the attribute with highest gain ratio is selected for splitting. Pruning is applied in the induction of decision tree. Pruning decision trees is a fundamental step in optimizing the computational efficiency as well as classification accuracy of such a model [14]. Applying pruning methods to a tree usually results in reducing the size of the tree (or the number of nodes) to avoid unnecessary complexity, and to avoid over-fitting of the data set when classifying new data. According to [14], the parameter to test the effectiveness of classifier is labeled as *confidence factor*. The lower values of confidence factor not only reduces the tree size, but also helps in filtering out statistically irrelevant nodes that would otherwise lead to classification errors.

To achieve higher accuracy of classifier, the J48 classifier is tested on data with confidence factor ranging from 0.1 to 1.0. The number of minimum instances per node (MinNumObject) was set to 2, and cross validation folds for test set (CrossValidationFolds) was set to 10.Few results on FIIs are reported in Table III and results on DIIs data are reported in Table IV.

The measures used to evaluate the performance of classifier are namely – correctly classified instances, Kappa statistics, precision, recall, F-measure, and ROC area [15]. The accuracy of results is evaluated on above mentioned performance measures. The study considered the weighted average of precision, recall, F-measure, and ROC area for evaluation of results.

Table III. Values of Performance Measures for FII data on different values of Confidence Factors

Confidence Factor	*Correctly Classified Instances %*	*Kappa Statistics*	*Precision (wt. Avg.)*	*Recall (wt. Avg.)*	*F-Measure*	*ROC area*
0.15	66.43	0.3254	0.602	0.606	0.582	0.589
0.25	65.76	0.3107	0.659	0.658	0.655	0.643
0.35	66.17	0.3198	0.662	0.662	0.660	0.647
0.45	66.57	0.3282	0.666	0.666	0.664	0.646
0.55	66.03	0.3174	0.660	0.660	0.659	0.648

Table IV. Values of Performance Measures for DII data on different values of Confidence Factors

Confidence Factor	*Correctly Classified Instances %*	*Kappa Statistics*	*Precision (wt. Avg.)*	*Recall (wt. Avg.)*	*F-Measure*	*ROC area*
0.15	59.35	0.1379	0.587	0.593	0.569	0.555
0.25	58.78	0.1257	0.580	0.588	0.563	0.558
0.35	57.98	0.1096	0.570	0.580	0.556	0.565
0.3	58.25	0.1135	0.576	0.583	0.557	0.568

For FIIs, maximum accuracy of classifier is achieved at confidence factor *0.35*. So, the decision tree for FII is induced at confidence factor *0.35*. For DIIs, maximum accuracy of classifier is achieved at confidence factor *0.15*. So, the decision tree for DIIs is induced at confidence factor *0.15*.

IV. RESULTS & DISCUSSION

The basic objective of the research is to decode the investment pattern of institutional investors. Stock market investment is very complex and sensitive investment exercise that requires a strong analytical backup. The frequent buying and selling decision has to beprudent in nature rather than a game of gambling. Institutional Investors has a backup of specialized research team and the vast experienceof handling the market. This advantage results in a positive outcome in the form of higher returns. The paper therefore tries to decodethis investment theory of institutions by using classification techniques of machine learning models. J48 decision tree was used as a classifier for splitting the decision nodes of these institutions.

A. *Performance Evaluation:*

The performance of the decision tree is measured by correctly classified instances, Kappa statistics, precision, recall, F-measure, and ROC area. The J48 algorithm that was run with value of *confidence factor*, C=0.35 & M=2 for FIIs action and *confidence factor*, C=0.15 & M=2 for DIIs action produced the following results:

1) Correctly Classified Instances: The percentage of correctly classified instances often called accuracy or sample accuracy is 66.1664% for FIIs and 59.3452% for DIIs.

2) Kappa Statistics: The Kappa statistic (or value) is a metric that compares an observed accuracy with an expectedaccuracy (random chance). There is no standardizedinterpretation for Kappa statistic. A value greater than 0 means that the classifier is doing better than chance. The value of Kappa statistic for FII is 0.3198 and for DII, it is 0.1379. According to [16], the value of Kappa statistic is fair if it lies between 0.21 and 0.40. Thus, the resulting values show that the classifier fairly classifies the data.

3) Precision: Precision is proportion of instances that are truly of a class divided by the total instances classified as that class. In the present context, we have *Action* attribute as the class label. There are two possible values of *Action* – buy and sell. Formally, *precision (n, cb)* for any node *n* and total number of instances classified as buy, *cb*, at node *n* is the ratio of number of instances correctly classified as buy to the total number of instances classified as buy. *precision (n, cs)* for any node *n* and total number of instances classified as sell, *cs*, at node *n* is the ratio of number of instances correctly classified as sell to the total number of instances classified as sell. The precision for buy is computed by (1) and precision for sell is computed by (2).

$$precision\ (n, cb) = \frac{correct\ \ (n,cb)}{cb} \qquad (1)$$

$$precsion(n, cs) = \frac{correct\ (n,cs)}{cs} \qquad (2)$$

In case of FIIs the precision value for buy is 65.8% and for sell is 66.7% making weighted average of 66.2%. For DIIs the mentioned rates are 60.3% and 56.5% with weighted average of 58.7%.

4) Recall: Recall is proportion of instances classified as a given class divided by the actualtotal in that class. Formally, *recall (n, cb)* for any node *n* and total number of instances classified as buy, *cb*, at node *n* is the ratio of number of instances correctly classified as buy to the total number of instances actually having *action* as buy, *btotal*. *recall (n, cs)* for any node *n* and total number of instances classified as sell, *cs*, at node *n* is the ratio of number of instances correctly classified as sell to the total number of total number of instances actually having *action* as sell, *stotal*.The recall for buy is computed by (3) and for sell is computed by (4).

$$recall\ (n, cb) = \frac{correct\ \ (n,cb)}{btotal} \qquad (3)$$

$$recall(n, cs) = \frac{correct\ (n,cs)}{stotal} \qquad (4)$$

Similar to precision rate, Recall in case of FIIs for buy is 72.4% and for sell is 59.5% making weighted average of 66.2%. For DIIs Recall is 79.8% and 33.4% for buy and sell respectively with weighted average of 59.3%.

5) F-measure: In statistical analysis of binary classification, F-measure is a measure of test's accuracy. F-measure can be interpreted as a weighted average of the precision and recall. F-measure is considered best when its value is 1 and is worst when its value is 0. The value of F-measure for FIIs' is 0.660 and for DIIs' its value is 0.569, which implies that the test accuracy of FIIs' classifier is much better than the test accuracy of DIIs' classifier.

6) ROC Curve: The ROC curve is a plot of true positive rate (tpr) against false positive rate (fpr). The lower left point (0, 0) represents the strategy of never issuing a positive classification; such a classifier commits no false positive errors but also gains no true positives [17]. The opposite strategy of unconditionally issuing positive classifications is represented by the upper right point (1, 1).The point (0, 1) represents perfect classification. This is one of the most important value output by Weka. An "optimal" classifier will have ROC area values approaching 1, with 0.5 being comparable to "random guessing" (similar to a Kappa statistic of 0). The weighted average of ROC area for FIIs is 0.647, which means that the model for FII is a good predictor of FII's behavior. The weighted average of ROC area for DIIs is 0.555, which means that the model for DIIs' is just better than random guessing model.

B. *Decision Rules*

On applying J48 algorithm on FII data under analysis with confidence factor 0.35, using 10-folds cross validation and allowing pruning on decision tree, the study derived decision tree as depicted in Fig. 1.

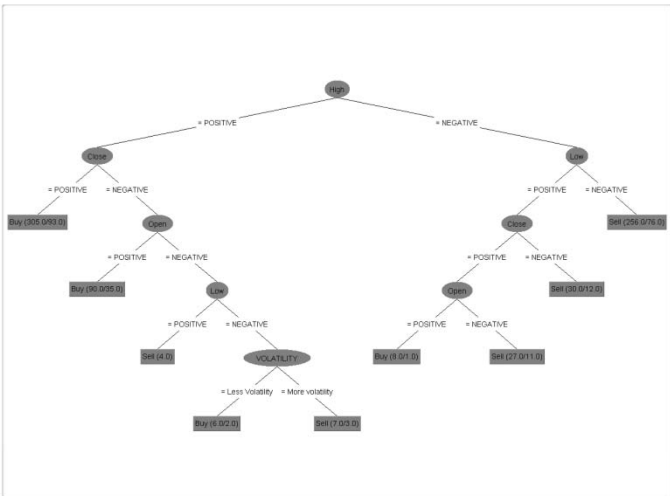

Figure 1. Predicted Investment Pattern of Domestic Institutional Investors

Decision tree can be transformed to a set of rules bymapping from the root node to the leaf nodes one by one. The set of rules resulting from FII's decision tree is as follows:

R1: IF (*High* = POSITIVE) AND (*Close* = POSITIVE) THEN *Action* = Buy.

R2: IF (*High* = POSITIVE) AND (*Close* = NEGATIVE) AND (*Open* = POSITIVE) THEN *Action* = Buy.

R3: IF (*High* = POSITIVE) AND (*Close* = NEGATIVE) AND (*Open* = NEGATIVE) AND (*Low* = POSITIVE) THEN *Action* = Sell.
R4: IF (*High* = POSITIVE) AND (*Close* = NEGATIVE) AND (*Open* = NEGATIVE) AND (*Low* = NEGATIVE) AND (*VOLATILITY*= Less Volatility) THEN *Action* = Buy.
R5: (*High* = POSITIVE) AND (*Close* = NEGATIVE) AND (*Open* = NEGATIVE) AND (*Low* = NEGATIVE) AND (*VOLATILITY* =More Volatility) THEN *Action* = Sell.
R6: IF (*High* = NEGATIVE) AND (*Low* = POSITIVE) AND (Close = POSITIVE) AND (*Open* = POSITIVE) THEN *Action* = Buy.
R7: IF (*High* = NEGATIVE) AND (*Low* = POSITIVE) AND (Close = POSITIVE) AND (Open = NEGATIVE) THEN *Action* = Sell.
R8: IF (*High* = NEGATIVE) AND (*Low* = POSITIVE) AND (Close = NEGATIVE) THEN *Action* = Sell.
R9: IF (*High* = NEGATIVE) AND (*Low* = NEGATIVE) THEN *Action* = Sell.

It is visible from the induced decision tree and decision rules that the classification for investment pattern of FIIs started by splitting the attribute, '*High*'. J48 uses highest gain ratio for splitting an attribute. This explains that the buying and selling of FIIs in cash market of NSE gets main cue from the *High* point of the day of index. If the current days' *High* is greater than the previous days' *High* i.e. if *High* is positive, FIIs takes second cue from the closing values. If there is a significant upward movement in the market that can lead to a positive *Close* as compared to previous days' *Close*, than FIIs have a strong tendency to be a net buyer in the market. The behavioral pattern is shown by decision tree in Fig. 1where 305 observations are classified until the first leftmost node is reached. This node is correctly classifying the described behavior for 212 observations out of 305 observations.

However if the closing price on t day seems to be getting lower, FIIs have looked at the opening of the market. If the opening on t day is higher to what it was on t-1 day, then again they have been the buyers on the day. The tree here has again correctly classified 55 out of the 90 such instances. In case, where *High* is positive close is negative and the difference of opening price is also negative then FIIs analyze the low point of the day before taking any action. If the *low* point of t day seems to be more than the *Low* value on t-1 day, then all 4instances are correctly classified as Sell. If this *low* is negative, they looked into the market's volatility. Further in continuance to previous behavior in times of low volatility they bought into the market and in case of more *Volatility* they sold in the market.

The second half of the tree reflects their action in case where the *High* value of index on t day is less than the *High* value on t-1 day i.e. if *High* is negative. In case if *High* is found to be negative FIIs looks for *Low* value, which if is less than previous days' *Low* results in a direct selling action on the part of FIIs. This is evident from the tree which is showing a correct classification of 180 observations where they have been the sellers in the market. The action seems not so simple when after a negative *High* and when *Low* is found to be positive. For a positive *Low*, if closing seems to be on a negative side then they were again observed to be the net sellers else they take cue from the opening value. The positive opening value results in buying otherwise selling action from FIIs. FIIs have been very balanced in their approach with 380 buying counts and 353 selling counts out of the total 733 observations. Despite their balanced approach as being net buyer or seller, they have been found to have net long positions on the day when markets moved up and were short on falling days of the market.

Contrary to FIIs, DIIs who had in last few years came up as a strong counter force to FIIs had exactly ironical investment pattern to FIIs. This is in a time where they have more buying calls (i.e. 410) in the sample period from April 2014 to March 2017.The J48 algorithm on DII data under analysis with confidence factor of 0.15, using 10-folds cross validation and allowing pruning on decision tree resulted in a decision tree as depicted in Fig. 2.

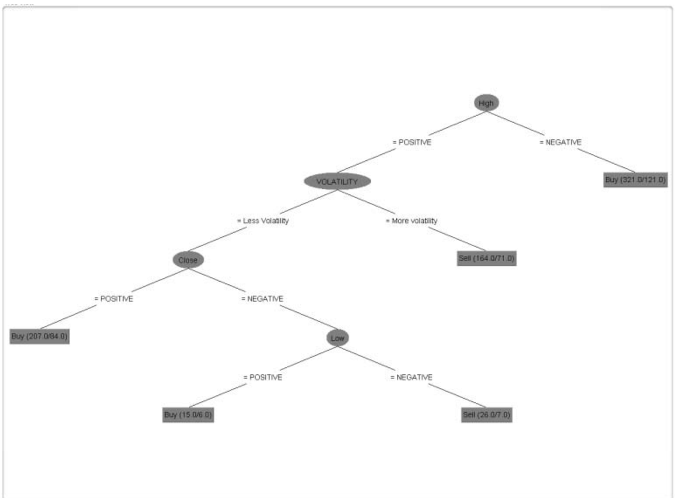

Figure 2. Predicted Investment Pattern of Domestic Institutional Investors

The set of rules resulting from DII's decision tree is as follows:

R1: IF (*High* = POSITIVE) AND (*Volatility* = Less Volatility) AND (*Close* = POSITIVE) THEN *Action* = Buy.
R2: IF (*High* = POSITIVE) AND (*Volatility* = Less Volatility) AND (*Close* = NEGATIVE) AND (*Low* = POSITIVE) THEN *Action* = Buy.
R3: IF (*High* = POSITIVE) AND (*Volatility* = Less Volatility) AND (*Close* = NEGATIVE) AND (*Low* = NEGATIVE) THEN *Action* = SELL.
R4: IF (*High* = POSITIVE) AND (*Volatility* = More Volatility) THEN *Action* = Sell.
R5: IF (*High* = NEGATIVE) THEN *Action* = Buy.

The contrary investment pattern of DIIs is also found in real data. The behavior gets validated from the decision tree and decision rules. The decision tree is again built by splitting the attribute with highest information gain. The classifier for DIIs action splits the *High* attribute first. Ironical to FIIs action where they used to be buying on positive *High*, DIIs are taking a buy call on a negative *High*. It appears that they did not take a direct action when market is positive but look at the *Volatility* in the market. In a more volatile movements occurring in NIFTY – where the change in India VIX is increasing, they were sellers on 91 occasions as correctly classified by tree. On 71 instances when they were buyers tree classified them as sellers. However, in a less volatile environment they look for the closing price, which if sets on positive side results in a buying by these institutional investors. On a closing, that heads for a lower than the previous days' *close* DIIs seems to take a cue from *low* point of the index. If the t days' *low* is positive, DIIs then take a long position in the market otherwise on a negative *Low;* sell their stocks in the market.

V. CONCLUSION

Stock market is considered as the backbone of any economy as it provides the desired level of investment to the corporate sector. The secondary market depth helps companies raise required money for their expansion. This depth in the market is created by participation of more and more players in the market especially the big guns. These big guns mainly are the institutional investors both domestic and foreigners. Ever since the opening of Indian capital market for external investors, these players had played a dominant role in the market by moving the market in their desired direction through huge quantum of assets under management. The new century had seen the slow but strong emergence of DIIs because of wide participation of mutual fund industry, insurance companies and other big industrial houses. The present study aimed to understand the dynamics of the investment pattern of these big players. Using the decision tree approach the research found that the behavior of FIIs is more predictable and according to the textbook investment strategy. They were found to buy in a market when it is on a positive note and sell when signals are weak. The research was able to correctly classify their pattern to an extent of 66.166%.

Contrary to this, the domestic players mainly were more unpredictable with accuracy rate of 59.34% Even the ROC confirms that their behavior is more of guessing nature. It is therefore concluded that investment pattern of FIIs are found be more matured and acceptable which is also validated from the returns they earn from Indian markets.

VI. REFERENCES

[1] K. C. Samal, "Emerging equity market in India: role of foreign institutional investors," Economic and Political Weekly, pp. 2729-2732, 1997.

[2] P. Mukherjee, and M. Roy, "What Drives the Stock Market Return in India? An Exploration with Dynamic Factor Model," Journal of Emerging Market Finance, vol. 15.1, pp.119-145, 2016.

[3] V. S. Dhingra, S. Gandhi, and H. P. Bulsara, "Foreign institutional investments in India: An empirical analysis of dynamic interactions with stock market return and volatility." IIMB Management Review, vol. 28.4, pp. 212-224, 2016.

[4] M. S. Khan, , and A. Khan, "Impact of FIIs on Indian Stock Market: A Case Study of National Stock Exchange of India," Al-Barkaat Journal of Finance & Management, vol. 8.2, pp. 75-85, 2016.

[5] B. Qian, and K. Rasheed, "Stock market prediction with multiple classifiers," Applied Intelligence, vol. 26.1, pp. 25-33, 2007.

[6] D. Chandwani, and M. S. Saluja, "Stock Direction Forecasting Techniques: An Empirical Study Combining Machine Learning System with Market Indicators in the Indian Context," International Journal of Computer Applications, vol,92.11, pp. 8-17, 2014.

[7] Q. A. Al-Radaideh, A. Assaf, and E. Alnagi, "Predicting stock prices using data mining techniques," The International Arab Conference on Information Technology (ACIT'2013), 2013.

[8] P. Paliyawan, "Stock Market Direction Prediction Using Data Mining Classification,"ARPN Journal of Engineering and Applied Sciences, vol 10.3, pp. 1302-1310, 2015.

[9] E. Kirkos, C. Spathis, and Y. Manolopoulos, "Support vector machines, Decision Trees and Neural Networks for auditor selection," Journal of Computational Methods in Sciences and Engineering, vol. 8.3, pp. 213-224, 2008.

[10] R. Caruana, and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," Proceedings of the 23rd international conference on Machine learning, ACM, 2006.

[11] A. K. Pujari, "Data mining techniques", Universities press, 2001.

[12] J. R. Quinlan, "C4. 5: programs for machine learning," Elsevier, 2014.

[13] E. Frank, M. A. Hall, and I. H. Witten, "The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann, Fourth Edition, 2016.

[14] S. Drazin, and M. Montag, "Decision tree analysis using weka,"Machine Learning-Project II, University of Miami, pp. 1-3, 2012.

[15] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation," Australasian Joint Conference on Artificial Intelligence, Springer Berlin Heidelberg, 2006.

[16] J. R. Landis, G. G. Koch, "The measurement of observer agreement for categorical data," Biometrics, vol. 33.1, 1977, pp. 159–174. doi:10.2307/2529310.

[17] T. Fawcett, "An introduction to ROC analysis," Pattern recognition letters, vol. 27.8, pp. 861-874, 2006.

