



International Journal of Financial Management and Economics

P-ISSN: 2617-9210
E-ISSN: 2617-9229
IJFME 2025; 8(1): 323-332
www.theeconomicjournal.com
Received: 10-02-2025
Accepted: 15-03-2025

Ranjith K
Research Scholar Kerala
University of Fisheries and
Ocean Studies Panagad,
Kochin, Kerala, India

Dr. Rajesh K
Assistant Professor and Head,
Department of Business
Administration and
Management, Kerala
University of Fisheries and
Ocean Studies (KUFOS),
Kochi, Kerala, India

Dr. Suraj ES
Associate Professor, KCT
Business School, Kumaraguru
College of Technology,
Coimbatore, Tamil Nadu,
India

Dr. Nitha KP
Guest Faculty, Department of
Computer Applications, Sree
Kerala Varma College,
Thrissur, Kerala, India

Corresponding Author:
Ranjith K
Research Scholar Kerala
University of Fisheries and
Ocean Studies Panagad,
Kochin, Kerala, India

Machine learning algorithm in Indian stock market for revising and refining the equity valuation models

Ranjith K, Rajesh K, Suraj ES and Nitha KP

DOI: <https://doi.org/10.33545/26179210.2025.v8.i1.500>

Abstract

Prediction of stock prices with machine learning approach is one of the attractive matter for researchers over the period. Statistical methods which are highly traditional are no longer viable for appropriate analysis of huge amount of data. With the help of machine learning approach, it is able to expose hidden patterns and forecast future trends and behaviour in stock market. In this paper, supervised machine learning algorithms, are proposed for classification of Indian stock market data. This study mainly focused on the machine learning approach for revising and refining the existing equity valuation models to bring more prediction accuracy. These results from the models will help the investor to make decision on selling, purchasing or holding the stocks in Indian stock market

Keywords: Machine learning, classification, supervised machine learning

1. Introduction

Stock market consists of several buyers and sellers of stock. Lot of researches have already been done to predict stock market prices using traditional and modern methods. This research is done to know the advantages of machine learning approach in stock price prediction after revising the equity models. Predicting the stock market is a difficult task which consists of several factors including economic, accounting and market sentiments etc. Machine learning is used in many sectors to know the pulse of the stock market. One of the most popular application of machine learning is stock market prediction itself. Machine learning algorithms are either supervised or unsupervised. In Supervised learning, labelled input data is trained and algorithm is applied. Classification and regression are types of supervised learning. It has a higher controlled environment. Unsupervised learning has unlabelled data but has lower controlled environment. It analyses pattern, correlation or cluster. The studies also cover various unsupervised machine learning algorithms as well as some statistical methods used in the stock price prediction. Study has been covered on generating portfolio of stocks that maximizes the profit with minimum risk.

Multiple regression machine learning algorithm

The multiple regression is used in statistics to understand the association between a dependent variable and more than one independent variables. In machine learning it is used to predict a dependent variable with two or more independent variables. The main application of the multiple regression machine learning algorithm is

- Effectiveness of independent variables on prediction,
- Predicting the impact of changes of the model that will impose on the results.

1.1 Classification

Classification, a field in statistics, which is the process of assigning an observation into the appropriate categories or class. Through machine learning the classifiers are trained with the appropriate features that represents each class. This trained model then predicts an unlabelled observation x , into a predicted label y . The model is then evaluated for its performance with the measures such as accuracy, precision, recall, f1-score. In this study the Random Forest classifier is used for the classification of the stocks into one of the four class labels. Random forest classification algorithm is an ensemble learning method for classification.

The advantage of random-forest classifiers is that it reduces overfitting. This study mainly focused on revising and refining the existing equity valuation models using machine learning approach for Sensex stock price prediction

2. Objectives of the study

1. To know the machine learning algorithm approach for stock price prediction in Indian stock market.
2. To study, revise and refine the selected equity valuation model using multiple regression models.
3. To bring the improvement in stock price prediction for Sensex stocks using revised and refined model using machine learning approach.

3. Methodology

An analytical study with the secondary sources of data including the yearly reports, official websites of BSE (www.bseindia.com), NSE (www.nseindia.com), Confederation of Indian Industries (CII) library, Centre for Monitoring Indian Economy (CMIE) Prowess and annual report library services like Ebsco Services is considered.

Table 1: List of Sensex stocks

Sl. No	Name of the Stock
1	Axis Bank
2	Asian Paints
3	Bajaj Auto
4	Bajaj Finance
5	Bharathi Airtel
6	Coal India
7	HCL Tech
8	HDFC
9	HDFC Bank
10	Hero Motors
11	Hindustan
12	ICICI Bank
13	IndusInd Bank
14	ITC
15	Kotak Bank
16	Larsen & Toubro (L&T)
17	Maruti
18	Mahindra & Mahindra (M&M)
19	National Thermal Power Corporation Ltd (NTPC)
20	Oil and Natural Gas Corporation Limited (ONGC)
21	Power Grid
22	Reliance
23	State Bank of India
24	Sun Pharma
25	Tata DVR
26	Tata Motors
27	Tata Steel
28	Tata Consultancy Services (TCS)
29	Vedanta
30	Yes Bank

3.1. Software's Used

To analyze the data, R Software 4.0.2 is used. To refine the best three equity valuation models and build an integrated model using machine learning or sentimental indicators, R python, Cocalc is used

4. Review of Literature

Patel *et al.* (2018) ^[1] using Machine Learning Mechanism for stock price forecasting. This study tried improve the predictive performance of the stock market and verifies

targeted value of forecast with the comparison of machine learning mechanism. Machine learning has been applied for different sectorial indices to predict to invest in a particular time in that company or not. Das *et al.* (2017) ^[3] utilized Hybrid Clustering-GWO-NARX neural organization strategy in anticipating stock cost and considered that K-implies bunching calculation has been utilized to decide the most encouraging group. This strategy can be utilized for foreseeing the stock cost by applying NARX neural organization calculation. Execution of the forecast through experimentation is contrasted and evaluated with assistance the financial backers in settling on a venture choice. He zeroed in on develop a hybridized model for stock expectation and assessed the presentation of forecast in other financial exchanges.

Henrique *et al.* (2018) ^[5] contemplated the stock value expectation utilizing support vector relapse on an everyday and up to the moment costs. This is the AI method called Support Vector Regression (SVR) to anticipate stock costs for enormous and little capitalization organizations in three unique business sectors. Expectation blunders of the model are estimated and the model is contrasted and the arbitrary walk model proposed by the proficient market theory. The outcomes demonstrated that the SVR has prescient force, particularly when utilizing a methodology of refreshing the model intermittently. There characteristic consequences of expanded forecast exactness can be seen uniquely during lower unpredictability periods. Wang *et al.* (2019) ^[7] utilized distinctive novel ways to deal with Sentiment Analysis for Stock Prediction. It's anything but a managed learning model can be utilized to foresee stock development bearing which can join specialized data and assessment through the news. This investigation in India on prediction of the bearing of offer costs, utilizing various procedures including SVMs and neural organizations by making an AI approach. Augby *et al.* (2018) ^[6] used Sentiment Analysis dependent on the informal communities for stock value expectation in the securities exchange. This examination zeroed in via online media, for example, "Twitter and Facebook" for financial exchange expectation purposes. This investigation tracked down that the web-based media impact can be utilized for transient value forecast and wistful examination can be considered as a help factor in the expectation of stock costs moreover. Adam Atkins (2018) ^[8] utilized monetary news to foresee financial exchange unpredictability. It was discovered that unpredictability developments are more unsurprising than resource value developments when utilizing monetary news as AI input. It was additionally discovered that expectation results depend just on data got from the news source.

Naik *et al.* (2012) ^[9] contemplated the beneficial exchanging rules for financial exchange ventures. He utilized the man-made brainpower methods to the inclined bearing of the cost for BSE file (India concretes stock value file (ICSPI)) prospects with a few specialized markers for corporate share warning reason. Then, at that point he mining the exchanging rules to decide the contention among the yields of the main stage utilizing develop learning strategy. Nyman *et al.* (2018) ^[10] contemplated the monetary determining that is to appraise the qualities for yield, swelling, work, spending, and so on, which are valuable to direct activity. He considered the audits on large information which gives numerous possibly energizing chances to finding information through whimsical ways.

Lima *et al.* (2013) ^[11] analysed the relationship between news stories and stock varieties which can be improved with different markers. This examination gathered news stories from three unique sites and the stock history from the New York Stock Exchange to continue to information mining order calculations. Articles in papers and magazines were named by their feeling, the immediate connection to a particular organization and geographic market impact. This investigation can recognize the examples between the pointers through this proposed model to anticipate stock value varieties with exactness's of 100%. Ding *et al.* (2013) ^[7] suggested profound realizing which will be valuable for occasion driven stock value development forecast by proposing a novel neural tensor organization. This examination additionally utilizing a profound convolutional neural organization to display the consolidated impact of long-haul occasions and momentary occasions on stock value developments. It was tracked down that a profound convolutional neural organization can catch the more extended term impact of information occasions than a standard feed-forward neural organization.

Wanzala, *et al.* (2017) ^[16] studied on the exact investigation of traditional microstructure-based liquidity measures to inside and out examination of market strength proxied by LR3, LR4 and VR. All the more explicitly, the examination was coordinated towards deciding if the market strength was an indicator of the monetary development contingent interaction model (CPM). This model can be utilized for examination since the model gives space to both the immediate and aberrant impacts of market versatility on monetary development. Kaushik *et al.* (2018) discovered that stock market prediction is an important issue for investors in financial market. Comparative performance of efficiency techniques will be useful for efficiently predicting the stock price in the market. This research will also useful in implementing the economic growth model in future for stock price prediction. Kanade *et al.* (2012) ^[17] used historical data examination expectation model worked with online media, enormous information logical capacities investigation and AI to occasionally foresee the pattern about financial exchanges. Future execution and unpredictability of the business sectors of the framework is influenced by the political and monetary news and impact of the web-based media. Calculation utilized for supposition examination utilizes summative evaluation of the assumptions in a specific tweet or news story which Will improves for the exactness of the forecast.

Liu *et al.* (2017) ^[13] utilized enormous information investigation to build a calculation stage for stock determination and exchanging methodologies. It receives world class molecule swarm improvement (EPSO) to explain ideal exchanging openings and consolidates becoming progressive self-arranging map (GHSOM) and EPSO in its stock determination technique. This investigation tracked down that stock examples, for example, star development stocks, and loads of ineffectively working organizations, stocks supported by steady or rising working productivity, and stocks supported by falling working benefit can be grouped for choosing protected, quality stocks by financial backers in securities exchange. Bach *et al.* (2019) ^[18] examined the large information innovations, data put away in different wellsprings of semi-organized and unstructured information. This examination zeroed in on s the scholarly centre of the field; strategies

utilized in monetary area for printed mining, in the time of the Internet, large information, and online media; wellsprings of information are regularly utilized for text mining in the monetary area, and for which purposes Finally, it was discovered that profundity investigation for text mining in money ought to be directed, zeroing in additional on the interior records as the space of the investigation.

S. Kavitha (2017) ^[19] focused on the huge information examination particularly prescient investigation which will impact in the monetary market with its arising advances stock trade. This examination investigated fluctuated procedures utilized for dissecting the unstructured data like stock trade audits and chronicled measurement of monetary data and shows that SVM calculation is essentially utilized for the wistful grouping due to its solid and exact outcomes contrasted with AI calculations. Yablonsky (2017) ^[21] provided an inside and out examination of the advanced stages utilizing IBM Watson Analytics as guide to foster prescient models at stock costs determining. This examination made a correlation with traditional scientific econometric stages and models with present day insightful stage models. It was tracked down that prescient models for stocks depend on the ideas of IBM Watson Analytics that have shown results, which are better than any remaining models.

K.C *et al.* (2018) ^[23] contemplated the presentation of Support Vector Regression Kernels (direct, polynomial and RBF) to know the prescient force. Distinctive execution pointers, for example, Moving Average technique, Correlation of loan costs, and Stock explicit news slant examination are broken down and assessed. Approval is performed by taking diverse time span conduct of the stock and looking at the real and anticipated stock costs. The venture choice is based on speculation methodology created to permit least danger venture for financial backer.

M Umer Ghani (2019) ^[24] utilized Machine Learning Algorithm zeroed in on Linear Regression (LR), Multi month Moving Average(3MMA), Exponential Smoothing (ES) and Time Series Forecasting utilizing MS Excel as best measurable device for graphical portrayal of expectation results. Subsequent to getting the information from Yahoo Finance for Amazon (AMZN) stock, AAPL stock and GOOGLE stock, Linear relapse effectively anticipated financial exchange pattern for the following month and furthermore estimated exactness as per estimations. K Hiba Sadia (2019) ^[22] evaluated and present and audit a more plausible technique to foresee the stock development with higher exactness. This paper inspects the utilization of the expectation framework in genuine world and issues related with the precision of the general qualities. The paper introducing an AI model to anticipate the life span of stock in a cutthroat market.

Aryendra Singh (2020) ^[26] studied on the recorded information on stock costs of recorded firms dependent on the reason to apply AI calculations, for example, Linear Regression progressed profound learning procedures like Long-Term Short Memory and applying specialized apparatuses to apply the Modern Portfolio Theory on the lookout. This methodology would be an extraordinary assistance for representatives and financial backers to put cash in the securities exchange since they depend on a tremendous scope of chronicled information.

Suchitha Bhorkar (2021) ^[25] assessed the papers, and saw

that AI calculation can be utilized making financial exchange forecast and programming language for making expectation is python. Studies utilized the dataset for expectation is taken from either Yahoo Finance or Google Finance. Kind of information contains data about the stock, for example, High, Low, Open, Close, Volume and so on Backing Vector Machine (SVM), Artificial Neural Network (ANN) are generally utilized for the forecast of stock. Calculated Regression and Random Forecast is utilized and gives great outcomes in papers.

5. Machine learning for equity valuation models

In this study, three equity valuation models are selected for the study purpose. P/E Model, P/B Model and CAPM has been used in this study. These three models are revised and refined using machine learning approach.

Linear regression is a simple algorithm first developed in the field of statistics. It is an approach to model the relationship between a response variable and one or more predictor variables. A linear relationship exists between input (independent) variables and a single (continuous, dependent) output variable. Therefore, the output variable, was calculated as a linear combination of the input variables.

Two main types of linear regression exist:

i) Simple linear regression, when there is only one independent variable

It was studied as a model for understanding relationships between a single input and output variable. For example, If the repo rate explains the market price, then the regression is expressed as,

$$\text{market price} = \beta_0 + \beta_1 * \text{repo rate} + \epsilon.$$

ii) Multiple linear regression, when there are multiple independent variables.

However, multiple linear regression goes one step ahead further and adds more than one independent variable. For example, if additional variables like DPR, EPS influence the market price, then the equation becomes,

$$\text{market price} = \beta_0 + \beta_1 * \text{repo rate} + \beta_2 * \text{DPR} + \beta_3 * \text{EPS} + \epsilon.$$

There are mainly three domains where the Regression-based machine learning algorithms Pai *et al.* (2021) can be applied.

- Quantifying the relations between the variables: Correlation identifies the strength and direction of the association between two variables. These relationships are quantified using multiple regression.
- Prediction: Regressions' ability to quantify the relationships/association can be turned into solving prediction problems.
- Forecasting: Estimation of some variables of interest at some specified future period is another domain where multiple regressions display outstanding performance. Prediction is similar to forecasting, but the estimations occur for a prolonged period when compared to the ones done for a specific date in forecasting

Multiple linear regression works by changing the parameter

values to reduce the cost, which is the degree of error between the predicted value and the values of the training set. The general equation of multiple linear regression is expressed as,

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n$$

- Y represents the dependent variable
- b_0 represents the axis intercept
- n signifies the number of variables in the dataset
- x_1 through x_n are independent variables in the dataset.
- b_1, b_2, \dots, b_n are coefficient parameters

The equation represents a hyperplane. A line of best fit for data in 3 or more dimensions is called a hyperplane. The attempt of the regressor will always be to create an equation of a hyperplane that best represents the training data. The regressor tries out several different equations to see which equations best fits the data.

A calculating metric known as cost function determines how well a hyperplane represents a training set. The cost function represents the degree of error between the hyperplane's values and those of the training set. Different formulas can calculate the cost. The one that the regression model uses is called the mean squared error (MSE). This cost function determines how well a hyperplane represents the data. The regression model can fine-tune the parameters' values to find the best fit hyperplane that defines the data the gradient descent function is used.

The gradient descent function helps the machine learning models to refine the parameters so that the optimal values for the coefficient of the hyperplanes can be identified. The algorithms use gradient descent in order to converge upon a parameter value that produces the lowest error. The models can later use these values for future predictions with a new set of data; in this study, multivariate gradient descent function to be used as we are dealing with a multilinear regression model.

After applying the gradient descent, the model will converge on the coefficients $b_0, b_1, b_2, \dots, b_n$ and finalize the parameter values and create the equation of the optimal hyperplane. Finally, plugging in the actual values to this optimized hyperplane equation will give the estimated value of the model.

The empirical analysis of the three equity valuation models used in the study had errors. So, in order to revise and refine the model, the multiple linear regression machine learning algorithm is used. The various steps used involved in creating the machine learning model are:

1. Data Preprocessing
2. Creating Train and Test Datasets
3. Creating and training the linear regressor
4. Testing the Linear regressor

Data Preprocessing: All the variables used in the data set were numerical. Identifying, removing inaccurate records from the dataset, recognizing unfinished, unreliable, inaccurate, non-relevant parts of the data was performed to clean the data. The inconsistencies identified or removed may have been caused by user entry mistakes, errors in the storage.

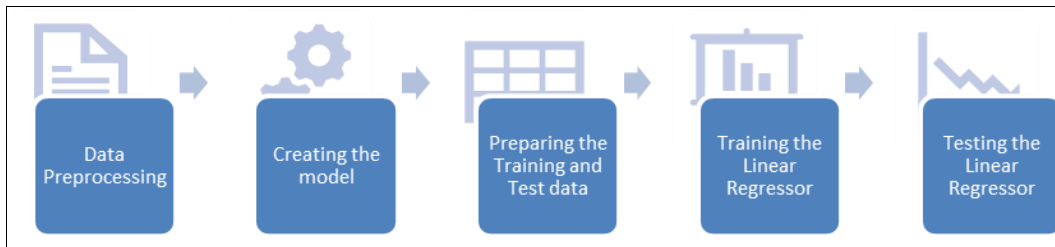


Fig 1: Steps used in the multiple linear regression machine learning model

Fig 1 shows the different steps involved in the multiple linear regression machine learning model. The data is then split into two arrays, one for independent variables and one for the dependent variable and given to the next step of the model.

Creating the Model: The multiple linear regressors is created for revised P/E, P/B and CAPM. The array of the independent variables used to create the model for the predictor or the dependent variables. For example, for the dependent variable, market price, the independent variables to the regressor are:

For P/E, the independent variables are EPS, EPS Growth, DPR, repo, ROE

For P/B, the independent variables are BVPS, repo, EPS, EPS Growth, DPR, ROE

For CAPM, the independent variables are R_f , R_m , beta.

Preparing the Training and the Test data: The whole dataset is divided into two training sets and the test set. Then, the ratio of 80:20 is selected, which divides 80% of

the data into the train data and the rest 20% to the test set.

Training the Linear Regressor: A linear regression model essentially adds a coefficient to each input variable, determining its importance. The value of input variables is then multiplied with the corresponding coefficient, and the bias (intercept) term is added to the sum. That is essentially the predicted value.

Testing the Linear regressor: To test the regressor, use the model to predict the test data. RMSE, MSE, R-squared metric are the commonly used metric used to evaluate the performance of a model

Architecture of refined equity models

This section covers the architecture of the revised equity valuation models for predicting the prices of Sensex stocks in Indian stock market.

Architecture of refined price to earnings model (P/E) using multiple regression machine learning approach

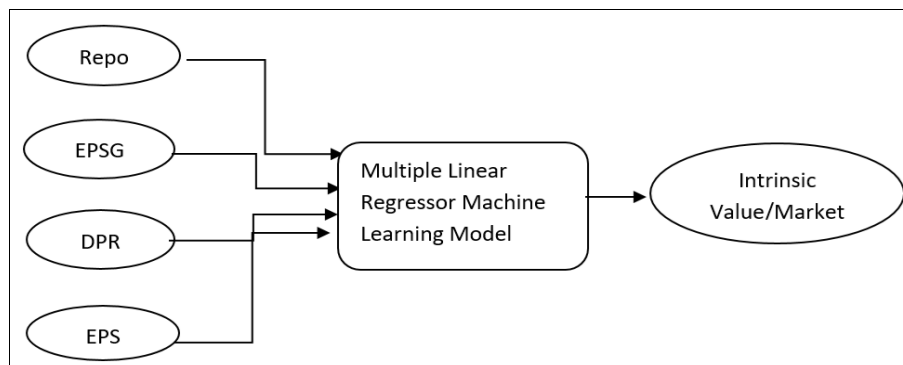


Fig 2: Architecture of Refined P/E Model using multiple regression machine learning

Fig 2 shows the architecture of the Refined P/E model that consists of four input variables-Repo rate, EPS Growth rate, DPR and EPS; Market Price as the output variable in the model.

6.2 Architecture of refined price to book value model using multiple regression machine learning approach

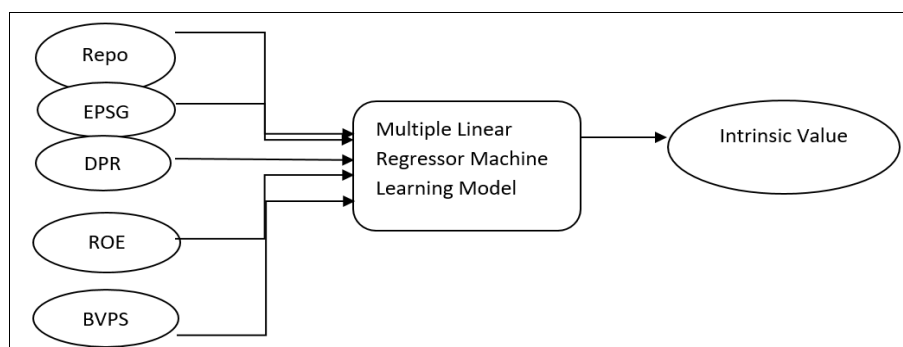


Fig 3: Architecture of refined P/B model using multiple regression machine learning approach

Fig 3 shows the architecture of the refined P/B model consists of five input features, repo, EPSg, DPR, ROE, and BVPS as input variables in the model, and IV as the output variable.

6.3 Architecture of refined capital asset pricing model using multiple regression machine learning approach

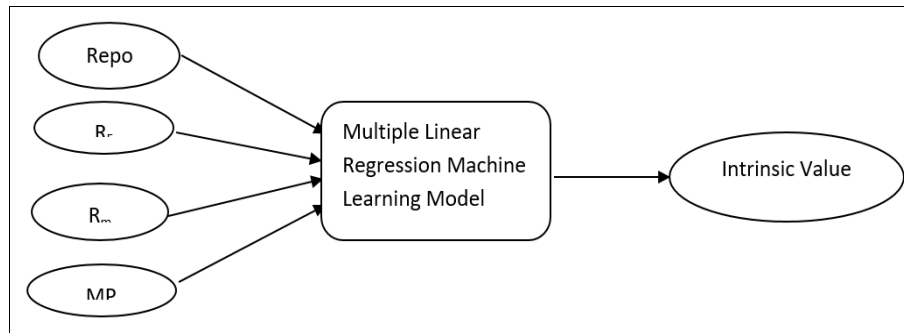


Fig 4: Architecture of refined capital asset pricing model using multiple regression ML approach

Fig 4 shows the architecture of refined CAPM which consists of three input features such as repo, Rf, Rm in the input layer, and IV as the output variable.

7. Analysis-checking the valuation accuracy of revised & refined models using machine learning Approach

This section covers checking the valuation accuracy of the revised and refined three superior valuation models. First, three equity valuation models such as P/E, P/B, CAPM models are selected and revised, substituting the model with the most appropriate feature that improves the model's accuracy. Then, the revised model is refined using a multiple regression machine learning model. Finally, the revised and refined model is tested for its valuation

accuracy.

Table 5 shows the RMSE of revised and refined models for Sensex stocks. The mean RMSE of the refined price to earnings model (P/E) is 1.71. It was found that the refined price to earnings model that resulted in the least RMSE with machine learning algorithm was adjudged as one of the best models that can give an accurate prediction. RMSE of Sensex stocks decreased significantly after using the multiple regression ML algorithm. This proved that regression is the most accurate than other methods. Power Grid and ITC have the lower RMSE value under refined price to earnings model bringing more accuracy to the intrinsic value of those stocks.

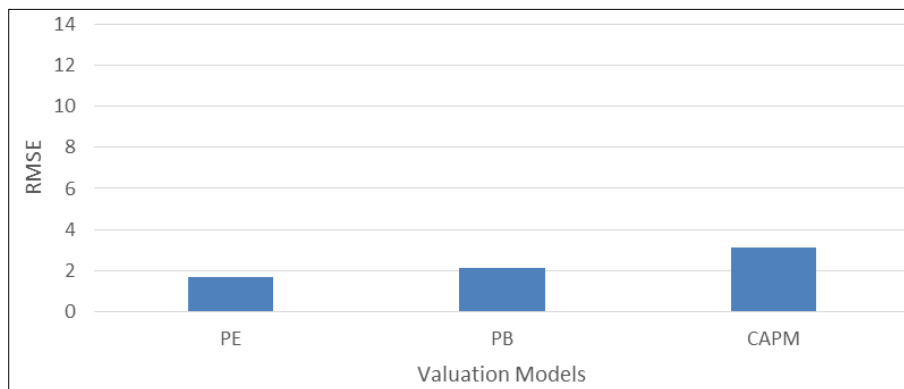


Fig 5: Comparison of select three Revised and Refined Equity Valuation Models for Sensex Stocks in India based on RMSE

Fig 5 shows the comparison of the RMSE value of the refined P/E, P/B and CAPM. There was a extreme decrease in the RMSE value of the base models to that of the refined models. This demonstrates the effect of the variable repo on

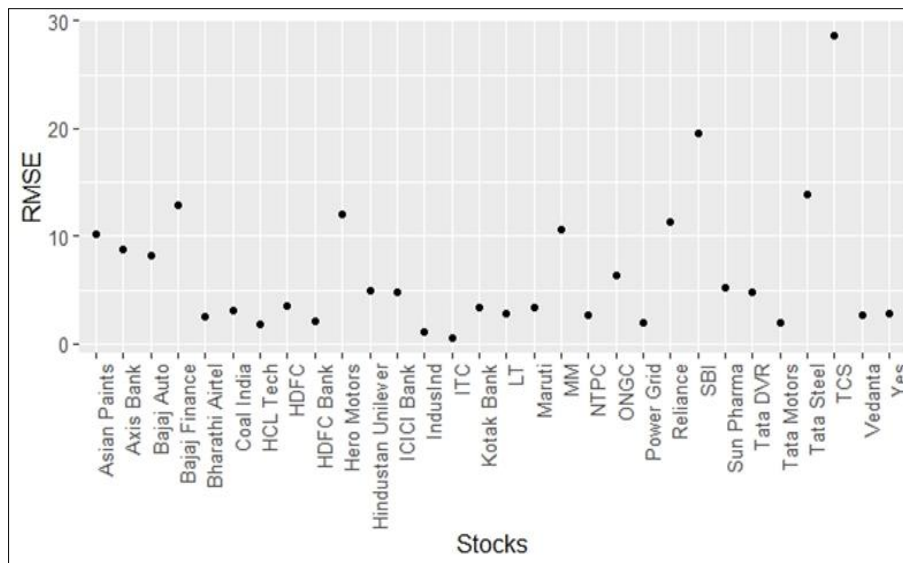
the valuation models and their accuracy. The mean RMSE of the refined price to book value model (P/B) model is 2.12, and that of CAPM is 3.15. When compared to the RMSE of the base models, it is very low.

Table 2: RMSE Comparison of Revised & Refined Equity Models

Stocks	Refined P/E Model	Refined P/B Model	Refined CAPM
Asian Paints	3.01	3.19	10.20
Axis Bank	1.07	1.75	3.28
Bajaj Auto	1.66	1.94	2.54
Bajaj Finance	4.22	22.41	2.54
Bharathi Airtel	0.55	0.63	0.44
Coal India	0.61	0.42	0.60
HCL Tech	0.73	0.95	0.90
HDFC	2.19	1.07	2.10
HDFC Bank	3.86	2.45	4.69
Hero Motors	2.09	2.72	4.86
HUL	0.59	0.69	3.06
ICICI Bank	1.64	0.92	3.26
IndusInd	0.78	0.69	4.89
ITC	0.22	0.20	0.40
Kotak Bank	0.86	1.10	4.89
LT	2.29	2.56	1.47
Maruti	3.20	2.33	16.60
MM	1.91	0.57	2.15
NTPC	0.14	0.14	0.20
ONGC	0.64	0.33	1.92
Power Grid	0.10	0.15	0.38
Reliance	1.09	2.13	1.43
SBI	5.44	3.24	6.42
Sun Pharma	2.16	2.00	1.73
Tata DVR	0.46	0.26	0.62
Tata Motors	2.83	2.78	1.04
Tata Steel	1.11	0.11	2.68
TCS	1.87	2.89	4.95
Vedanta	0.74	1.45	0.48
Yes	3.19	1.67	3.87
RMSE	1.71	2.12	3.15

* Results computed using R-4.0.2

It was found that RMSE of Refined P/E Model works very well for Sensex stocks due to the predictive power of the variables in the model.



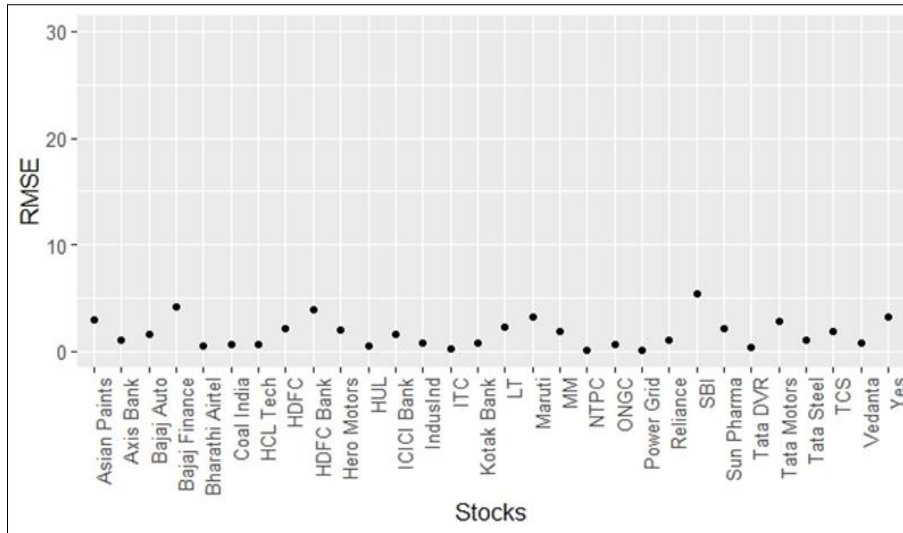


Fig 6: RMSE of the P/E Model and Refined PE Model

Fig 6 compares the RMSE of the P/E model to that of the revised and refined P/E model. The plot shows that the RMSE of the stocks that had a span from 0-30 for the 30 Sensex stocks reduced to 0-10 with the help of the multiple regression machine learning models. The P/E model gave an RMSE of 6.65 (Reference Table 3.15), whereas that of the

refined PE model is 1.71. RMSE of 93% stocks were reduced except for the HDFC Bank and Tata Motors. The Sensex returns and the stock returns of these are highly correlated; hence market risk represented by beta is highly significant for these stocks.

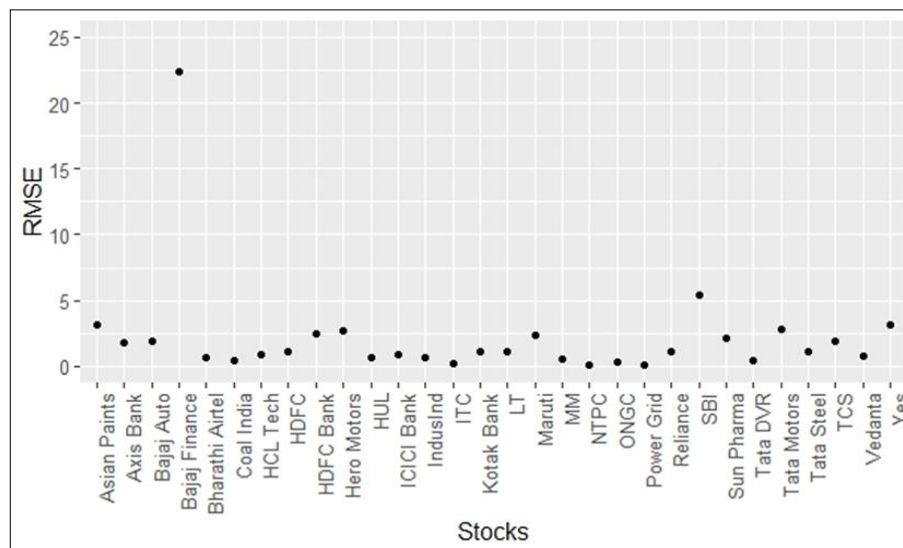
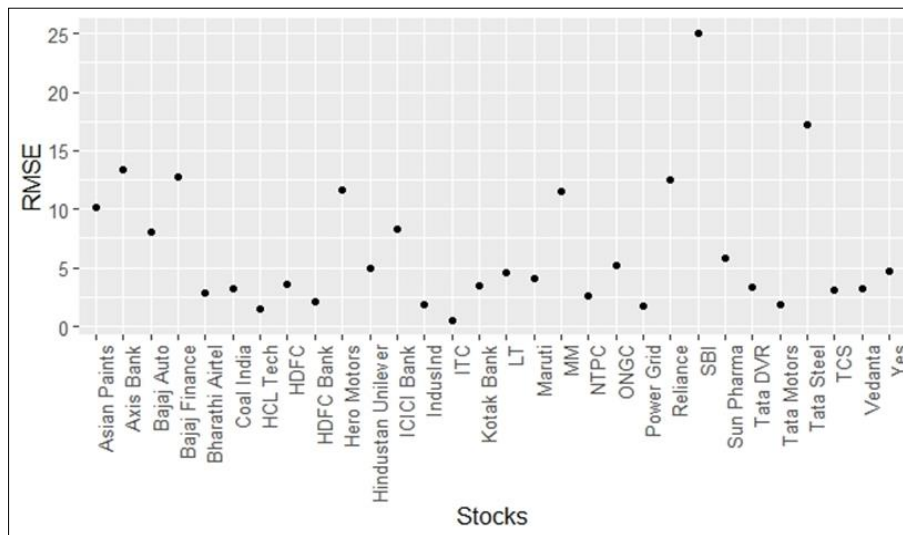


Fig 7: RMSE of the P/B Model and refined PB Models

Fig 7 compares the RMSE of the P/B Model to that of the refined P/B Model. The plot shows that the RMSE of the refined P/B model was significantly reduced after refining

the model. The RMSE at a maximum level of 25.04 was brought to the range of 0-5 except for Bajaj finance.

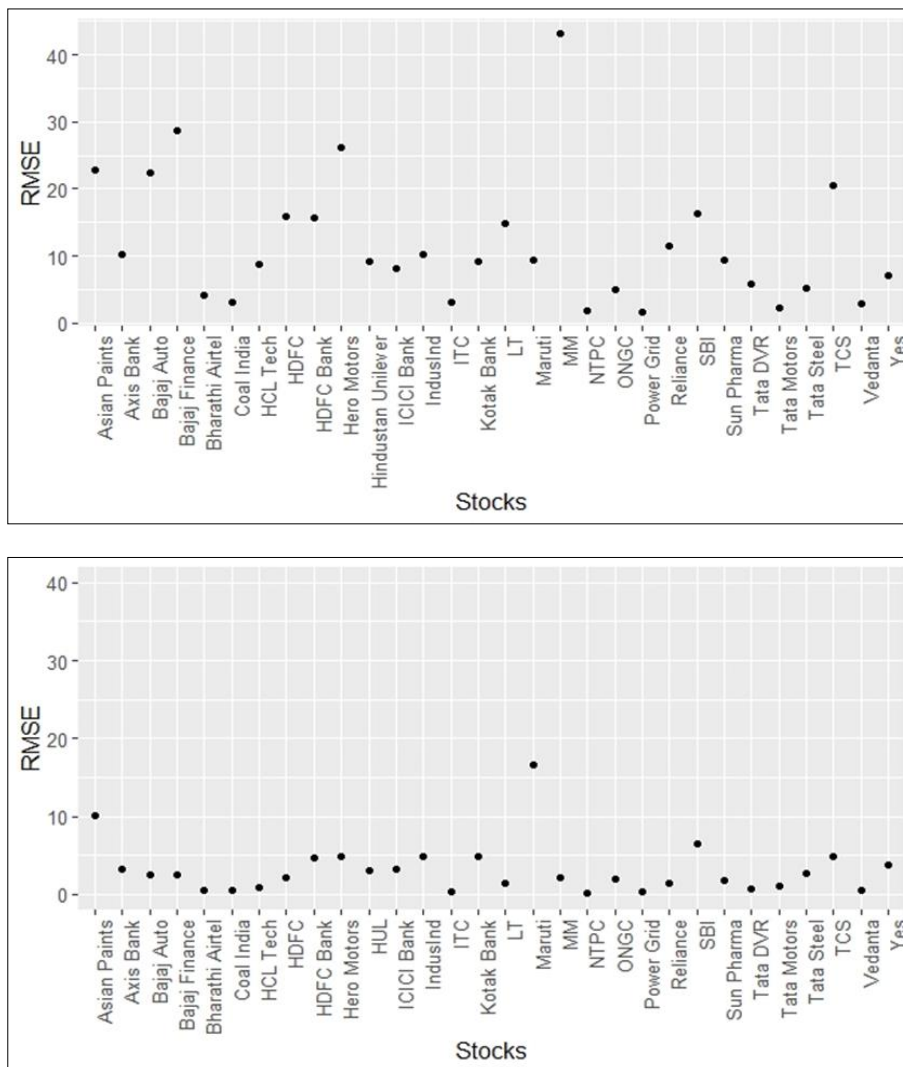


Fig 8: RMSE of the CAPM and Refined CAPM Model

Fig 8 gives the comparison of CAPM to the refined CAPM model. The RMSE values in a range of 0-40 were reduced to 0-10, with Maruti Suzuki as an outlier at the value of 16.6. This shows that market risk is the most relevant factor

that determines the price movement for this stock. Hence for the prediction of these stocks Maruti Suzuki, Tata Motors, HDFC and Bajaj Finance, the traditional models to be followed

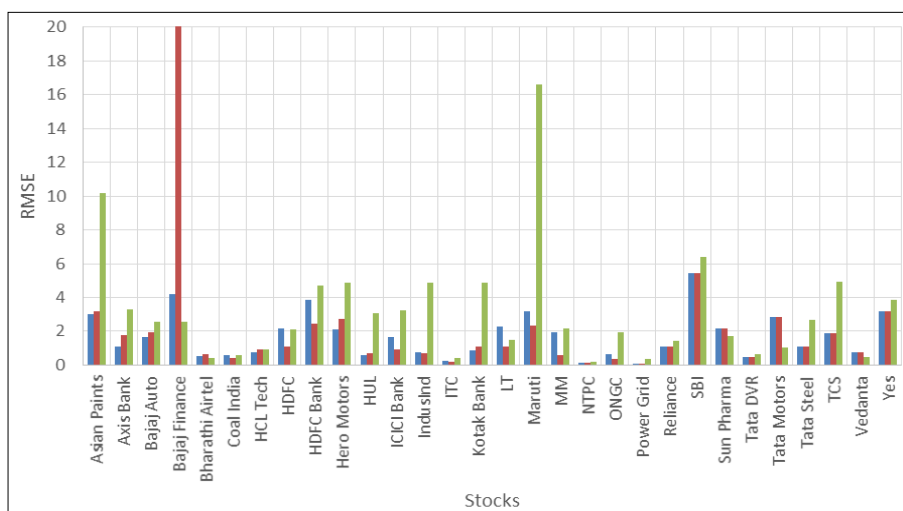


Fig 9: Plot of the refined P/E, P/B, CAPM models for Sensex stocks in India based on RMSE

Fig 9 shows the overall RMSE values of the three refined models. The RMSE value got reduced extremely which shows the significance of the substituted variables. From some sectors like automobile industry, financial firms the market risk is a highly significant variable. On consideration of such variables may give an adverse result than the expected one, the outliers shown in the plot is an example

8. Conclusion

The results of the revised and refined model shows that the valuation errors when compared to the base model are reduced. The P/E model with least RMSE 1.71 predicts the Intrinsic value with highest accuracy. To support the investors to choose the appropriate Sensex stocks and to make profitable investments the market sentiments also play a vital role. This study improved the predictive power of the models through multiple regression approach for Sensex stocks in Indian stock market.

9. References

- Patel K. Analysis of stock market forecasting using machine learning mechanism. *JASC: Journal of Applied Science and Computations*. 2018;5(6):112-119.
- Patel R. Stock market prediction using sentiment analysis: Testing the method's accuracy and efficiency. *International Journal of Latest Technology in Engineering, Management & Applied Science*. 2016;5(3):143-145.
- Das D. Hybrid clustering-GWO-NARX neural network technique in predicting stock price. *IOP Conference Series: Journal of Physics: Conference Series*. 2017;755:1-14.
- Das S. Real-time sentiment analysis of Twitter streaming data for stock prediction. *Procedia Computer Science*. 2018;132:956-964.
- Henrique BM, Sobreiro VA, Kimura H. Stock price prediction using support vector regression on daily and intraday prices. *The Journal of Finance and Data Science*. 2018;4(3):183-201.
- Al Augby S. Stock market prediction using sentiment analysis based on social network: Analytical study. *Journal of Engineering and Applied Science*. 2018;13(6):2388-2402.
- Wang C. Novel approaches to sentiment analysis for stock prediction. *GitHub Repository*. 2019;1-12.
- Atkins A. Financial news predicts stock market volatility better than close price. *The Journal of Finance and Data Science*. 2018;4(2):120-137.
- Naik RL. Prediction of stock market index using genetic algorithm. *Computer Engineering and Intelligent Systems*. 2012;3(6):162-172.
- Nyman R. News and narratives in financial systems: Exploiting big data for systemic risk. *Staff Working Paper-Bank of England*. 2018;1-18.
- Lima L. Big data for stock market by means of mining techniques. *Algoritmi Research Centre, University of Minho, Portugal*. 2013;1-14.
- Ding X, Zhang Y, Liu T, Duan J. Deep learning for event-driven stock prediction. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*. 2015;2327-2333.
- Liu W. EPSO-GHSOM stock selecting and trading strategy on big data. *International Journal of Intelligent Information Systems*. 2017;6(6):7-20.
- Sharma S, Kaushik B. Quantitative analysis of stock market prediction for accurate investment decisions in future. *Journal of Artificial Intelligence*. 2018;1(1):48-54.
- Wanzala RW. Estimation of market immediacy by coefficient of elasticity of trading: Three approaches. *The Journal of Finance and Data Science*. 2018;4(2):139-156.
- Wanzala RW. Market resiliency conundrum: Is it a predictor of economic growth? *The Journal of Finance and Data Science*. 2017;3(1):1-15.
- Kanade V. Stock market prediction using historical data analysis. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2012;2(12):267-270.
- Bach MP, Zoroja J, Vukšić VB. Text mining for big data analysis in financial sector. *Sustainability*. 2019;11(17):1-27.
- Kavitha S. A review on big data analytics via social media. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*. 2017;2(6):35-45.
- Kavitha S. Big data analysis in financial market. *International Journal of Research in Engineering and Technology*. 2015;4(2):422-427.
- Yablonsky S. Modern advanced analytics platforms and predictive models for stock price forecasting: IBM Watson Analytics case. *Proceedings of the 50th Hawaii International Conference on System Sciences*. 2017;1083-1092.
- Sadia K. Stock market prediction using machine learning algorithms. *International Journal of Engineering and Advanced Technology*. 2019;9(1):25-31.
- K.C EH. Performance analysis and prediction of Nepal stock market (NEPSE) for investment decision using machine learning. *International Journal of Computer Science Engineering*. 2018;7(3):15-27.
- Ghani MU. Stock market prediction using machine learning algorithms. *Advances in Distributed Computing and Artificial Intelligence Journal*. 2019;8(4):97-116.
- Bhorkar S. Study of market predictions using machine learning. *International Journal of All Research Education and Scientific Methods*. 2021;9(4):971-977.
- Singh A. An empirical research and comprehensive analysis of stock market prediction using machine learning and deep learning techniques. *IOP Conference Series: Materials Science and Engineering*. 2020;1022(1):1-12.
- Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell stem cell*. 2013 Apr 4;12(4):393-394.