Innovations

Application of machine learning models to improve the accuracy of earnings management prediction

Dr P Varadharajan*

*Professor of Finance, MS Ramaiah Institute of Management, Bengaluru, Karnataka, India.

Corresponding Author: **Dr P Varadharajan**

Abstract

Earnings Management carried out continuously can affect the accuracy and reliabil- ity of financial statements. This paper focuses on developing a machine learning model to pre- dict earnings management. Financial data of NIFTY500 companies from the year 2012 to 2021 collected from Prowess Database is used in this study. The study used random forest approach to screen the variables by finding their significance using the mean decrease gini. The C5.0 decision tree theory was established to identify the firms' level of earnings management. The results indicate that the proposed hybrid model developed using random forest and decision tree has an accuracy of 85.1%. The neural network model developed had an accuracy of 78.2%. The decision tree model was found to be more effective than the neural network model. It could be seen that operating profit margin, total assets turnover and corporate size are decisive factors in determining the level of earnings management.

Keywords: 1.*Earnings Management, 2.C5.0 Decision Tree, 3.Artificial Neural Network and Financial Inclusion.*

Introduction

An earnings management technique typically shows a positive view of the financial situation by manipulating financial reports. Earnings Management (EM) has been a major concern for most organizations for several decades now. Generally, earnings are viewed as a measure of an enterprise's past performance. Stakeholders consider corporate earnings as a vital indicator for a firm's operating performance. The management of earnings has therefore become one of the major objectives of the firm (Chen et al.,2015). Companies often manipulate earnings through discretionary accruals. Such conduct and intent may lead to serious consequences for the stakeholders and create information asymmetry. Earnings management carried out contin- uously by a company can lower its credibility. Additionally, manipulating the cost with earn- ings information may undermine the organization's value. Studies show that earnings manage- ment methods have a negative impact on the brand's success (Ariza et al.,2016). Earnings management can affect accuracy of financial statements by causing users to rely on profit fig- ures from technical financial statements (Kalbuana et al.,2021). Earnings manipulation weak- ens investors' trust and has a negative impact on the country's economy (Ahmed and Naima, Earnings management involves manipulating accruals through discretionary choices in ac- crual accounting in order to achieve specific targets (Kliestik et al, 2021). Researchers have identified two major categories of earnings management research: accrual-based earnings man- agement and real earnings management. Accrual-based earnings management is the most com- mon method of managing earnings since it does not violate accounting principles, and man- agement can adjust the results according to its own discretion (Chen and Shen, 2020). One of the most common ways to manage profits is by manipulating accounting accruals, especially discretionary accruals. Accounting accruals are, therefore, often considered in studies of earn- ings management (Mahmoudi et al., 2017). Real earnings management is a strategy for assist- ing a corporation in departing from the traditional financial statement operation guideline. The drawback of real earnings management is that it necessitates earlier execution, which affects an enterprise's actual value and real economic activity (Chen et al., 2015).

A model to predict earnings management would be useful in finding out the extent of earn- ings management process that is carried out in an organization. Linear Regression models have been widely used in predicting earnings management. To help analysts and decision-makers such as investors predict earnings, a variety of techniques have been created over time. Ma- chine Learning approaches can extract useful information from any

sort of data using super- vised and unsupervised algorithms (Huang and Yen,2019). Machine learning methods have the advantage of providing variable selection approaches that allow us to locate the most relevant predictors from vast collections of financial variables while avoiding the frequent overfitting concerns associated with models with many predictors (Zadeh et al.,2020).Using hybrid ma- chine learning methods, this study aims to develop an efficient and accurate hybrid machine learning models for predicting earnings management by using accrual-based earnings manage- ment.

Literature Review

Several machine learning models were developed and tested in an attempt to establish a model of high accuracy to predict earnings management. The Linear Regression (LR) model has been used extensively in traditional approaches to study earnings management. This technique how- ever, has several limitations, such as linearity and assumptions like non-flexibility of regression models, nonexistence of correlation and homoscedasticity. Hence, other machine learning tech- niques could be employed to predict the degree of earnings management carried out in an or- ganization (Namazi and Maharluie, 2015).

In a study conducted in electronic companies listed in the Taiwan Stock Exchange, earnings management prediction model was developed based on elastic net and C5.0 algorithms. It was found that C5.0 has a better accuracy and provides best classification performance (Chen and Shen, 2020). In an attempt to evaluate the predicting ability of neural networks, Iranian com- panies from five different industries listed on the Tehran Stock Exchange were studied. The model developed consisted of multilayer perceptron neural network with two hidden layers. The Multilayer perceptron neural network has the ability to predict earnings management at various levels in different industries (Mahmoudi et al., 2017). Hunt et al., (2019) used stepwise logistic regression, elastic net and random forest algorithms on a sample consisting of 1,16,904 firm year observations, and found that the random forest-based model is most effective.

The combination of Principal Component Analysis, LightGBM and Hyperopt method pre- dicted earnings better than the traditional Logistic Regression when tested on 3000 companies with highest market capitalization over a 30-year period. It was also found that, using neural networks and decision trees together provides not only better predictability but also important decision rules compared to using only neural network (Xinyue et al., 2020). Huang and Yen, (2019) found that the XGBoost algorithm provides the most accurate prediction among GA fuzzy clustering, XG Boost, HACT and contrastive divergence algorithm. Chen and Howar (2015) attempted to detect earnings management practices by using soft computing methods. The hybrid model constructed using stepwise regression, decision trees and random forest is the best choice as it provided optimal results in terms of accuracy in predicting earnings ma- nipulation of publicly listed electronic companies. Financial statements from big corporations over a period of four years were studied using Bayesian Naive Classifier (BNC) to enhance the decision-making process by identifying earn- ings manipulation (Dbouk and Zaarour, 2017). Qualitative and quantitative models to detect earnings manipulation were developed and tested on two private and state-owned entities in Fijian. The performance of both the discretionary accrual model and the qualitative model were not consistent in the measurement of earnings management (Naidu and Patel, 2013). Data relating to 117 firms from eight different industries were collected and machine learn- ing techniques were used to predict earnings management. Quarterly Earnings Prediction using epsilon support vector regression QEPSVR produced more significant results than the Brown and Rozeff Auto-Regressive Integrated Moving Average (BR ARIMA) model (Fischer et al., 2020). Difference between the neural network models and regression model in detecting earn- ings manipulation of 94 listed companies in the Tehran Stock Exchange were analyzed. The results show that multi-layer perceptron and generalized regression neural networks (GRNN)

are more accurate than linear regression (Namazi and Maharluie, 2015). Out of the two models developed to predict earnings management of bio technology industries, the GRNN-based model had the best performance and the linear regression-based model had the least perfor- mance and the Bayesian Network screening method, together with the C5.0 decision tree, yields the best results (Chen et al., 2015). Based on the review of literature, this study adapts a combination of Random Forest and C5.0 Decision tree. Random Forest is used for variable screening and decision tree algorithm is used to develop a prediction model.

Random Forest

A common goal in prediction modelling is to reduce the number of variables required to obtain a forecast in order to increase efficiency of the model. Random forests are tree-based prediction models in which each tree is based on a random vector sampled independently and whose dis- tribution matches that of all other trees in the forest. As a result of accurate calculation, random forests can determine an important index of independent variables. The index can capture the interaction between the predictors and the response variable. Chen and Howard, (2015) in their study found that, the random forest model was stable in terms of accuracy and chose the most relevant predictive variables in identifying earnings management. In the random forest archi- tecture, variable selection is a key factor for many applications in expert systems. For datasets with known underlying relationships between predictors and outcomes, conditional random forest methods for variable selection may be preferred since conditional random forest is often better at correctly finding significant associations between predictors (Speiser et al, 2019). This study adopted random forest as a vital indicator to evaluate the variables of earning manage- ment due to its high performance in carrying out classification tasks.

C5.0 Decision Tree

Decision trees can be used for both classification and regression problems. This algorithm is used to analyse large datasets based on division rule, resulting in the best predictions. In addi- tion to not being susceptible to any statistical hypothesis of the sample data, decision tree's key characteristics are its capacity to treat partial data and examine the potential relationships among large and sophisticated input and output variables. As a result, C5.0 is one of the most widely utilized decision tree algorithms (Chen and Shen, 2020). When applied to detect earn- ings manipulation in the biotechnology industry, the C5.0 decision tree produced excellent

outcomes with very high accuracy (Chen et al., 2015). In this study, the C5.0 method is used for prediction.

Artificial Neural Network

In comparison with statistical approaches, neural networks have three significant advantages. For starters, neural networks can learn any complex design or nonlinear mapping. Second, neural networks take no defaults into account when distributing data, and third, neural networks are particularly flexible when dealing with partial, missing, or ambiguous data (Mahmoudi et al., 2017). An influential neural network model is the multilayer perceptron (MLP), which con- sists of many layers of nodes. The input, the output and the hidden layer consist of nodes cor- responding to input nodes, output nodes, and hidden nodes (Tsai and Chiou, 2009). A good layering and input selection can approximate a good result. Input, hidden, summation, and out- put layers make up a generalized regression neural network. The parameters and structures of the neural networks are not clearly defined when using neural networks, and the way in which they are defined is often based on trial and error. On comparing the neural network model with other machine learning models, neural network model provides more accuracy in prediction (Namazi and Maharluie, 2015).

Research Gap

Traditional auditing systems, which are constrained by time, human resources, price, and ef- fects, struggle to detect anomalous activities in vast and complicated financial data. As a result, developing a prediction model for the level of earnings management is quite useful for auditors and investors in determining the degree of financial statement manipulation.

Research Objectives

The purpose of this study is to develop hybrid machine learning models to predict earnings management and to find the accuracy of the models.

Methodology

In the research process, data was identified and pre-processed, a training set was defined, al- gorithms were selected, training parameters were identified, and the test set was evaluated. In order to predict earnings management, potential predictive variables were identified. Consid- ering the fact that the study has many variables, Random Forest method was used for variable

screening and arriving at final list of variables which was used as input for the decision tree model (C5.0) and neural network model for predicting earnings management. Data was col- lected from NIFTY 500 companies from 2012 to 2021 using the Prowess database. Mean im- putation was used to treat missing values during the pre-processing stage.

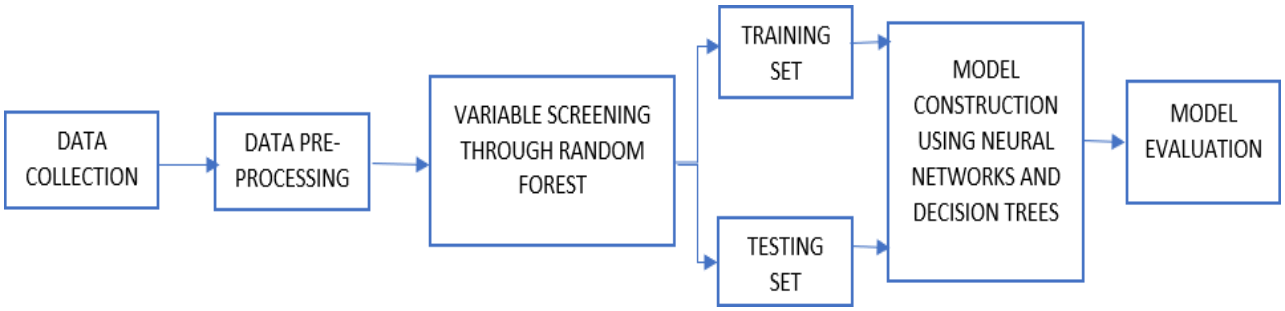

Fig. 1. Research Methodology

After estimating the alpha values, the non-discretionary accruals were computed. Reducing the non-discretionary accruals from the total accruals gives the discretionary accruals. Natural logarithm values for the discretionary accruals were calculated and the degree of earnings man- agement was classified based on those values. The average and standard deviation of the dis- cretionary accruals were calculated and the ceiling and floor values were set by adding and subtracting the standard deviation from the average respectively. The observations had an av- erage of 8.90 and standard deviation of 3.42. If the value is above the ceiling or below the floor, it is classified as Very High Earnings Management or Very Low Earnings Management. If the values fall between the average and floor, it is classified as Low Earnings Management and if it falls between average and ceiling, it is classified as High Earnings Management (Chen et al., 2015).

Table 1. Earnings Management Classification

Classification	**Description**	**No of samples**
Very Low Earnings Management	DA<5.47	313
Low Earnings Management	5.47<DA<8.90	1282
High Earnings Management	12.33>DA>8.90	2635
Very High Earnings Management	DA>12.33	435
Total		4665

Independent Variables. Based on previous studies potential predictive variables were identi- fied. 19 variables were selected which could probably affect earnings management. The varia- ble names and formula are mentioned below.

Table 2. Independent Variables

Variable Code	Variable Name	Formula	Source
X1	Leverage Coefficient	Total Liabilities / Total As- sets	Mahmoudi et al., 2017 , Chen and Howard,2015
X2	Corporate Size	Sales revenue of the firm	Mahmoudi et al., 2017, Chen et al., 2015
X3	Corporate Performance	Cash from Operations	Chen and How- ard,2015
X4	Return on Equity	Net Income / Shareholder's Equity	Chen and How- ard,2015
X5	Return on Assets	Net Income / Total Assets	Chen and How- ard,2015
X6	Operating Cash Flow	Cash flow from operating activities	Chen and How- ard,2015, Chen et al., 2015
X7	Total Assets Turnover	Net Sales / Average Total Assets	Chen and Shen,2020
X8	Current Ratio	Current Assets / Current Li- abilities	Chen and Shen,2020
X9	Net Profit Margin	Revenue – Cost / Revenue	Chen et al., 2015
X10	Operating cash flow ratio	Operating cash flow / Cur- rent liabilities	Chen and Shen,2020
X11	Long-term funds ap- propriate rate	(Total stockholders' eq- uity+ Long term liabilities) ÷Total fixed assets	Chen and Shen,2020
X12	Sales-to-equity ratio	Sales revenue / Total equity	Chen and Shen,2020

X13	Employee profitability	Net profit before tax / total number of employees	Chen et al., 2015
X14	Quick Ratio	(Current Assets – Inventory) / Current Liabilities	Chen and Shen,2020
X15	Managerial Ownership	Shares held by promoters	Chen et al., 2015
X16	Loss	1 if loss, else 0	Chen et al., 2015
X17	Financing Activities	Change in outstanding shares more than 10% - 1 else 0	Mahmoudi et al., 2017, Chen and Howard,2015
X18	Operating profit margin	Operating Income / Sales revenue	Chen and Shen,2020

Analysis and Interpretation

Variable Screening Using Random Forest

Given the fact that the study has many variables for the prediction of earnings management, random forest method was used to find out the significance of the variables before establishing decision tree model. The random forest packages were installed, and the data was split into training and testing sets in an 8:2 ratio. Thereafter, the random forest algorithm was used to determine the significance of the variables employed.

The random forest approach considers mean decrease gini as an important indicator in de- termining the significance of the variable in earnings management prediction. Higher the value of mean decrease gini, greater the influence of the variable on earnings management. Table 3 shows the mean decrease gini values of the variables used in this study. The order of importance of the variables are Operating Cash Flow, Corporate Size, Operating Profit Margin, Corporate Performance, Long Term Funds Appropriate Rate, Managerial Ownership, Total Assets Turn- over, Return on Assets, Net Profit Margin, Leverage Coefficient, Quick Ratio, Return on Eq- uity, Current Ratio, Operating Cash Flow Ratio, Sales to Equity Ratio, Employee Profitability. The variables Loss and Financing Activities do not have any significance and hence will be removed in the further study. The accuracy of the random forest model was found using the confusion matrix. The random forest model developed in the study to find out the significance of the variables had an accuracy of 75%. Mean decrease accuracy specifies the level of decrease in accuracy of the model when the specific variable is dropped. With three variables tried at

each split and the number of trees being 500, the random forest model had an out of bag esti- mate of error rate as 24.59% resulting in an accuracy of 75%.

Table 3. Variable importance output from Random Forest

Variable	Mean Decrease Accuracy	Mean Decrease Gini
Operating Cash Flow	0.053	207.9973
Corporate Size	0.068	175.3198
Operating Profit Margin	0.068	168.2167
Corporate Performance	0.051	153.5326
Long Term Funds Appropriate Rate	0.054	147.8837
Managerial Ownership	0.037	136.1943
Total Assets Turnover	0.031	112.1690
Return on Assets	0.026	107.8518
Net Profit Margin	0.024	102.3296
Leverage Coefficient	0.023	102.0554
Quick Ratio	0.023	98.3986
Return on Equity	0.012	97.1454
Current Ratio	0.020	97.1209
Operating Cash Flow Ratio	0.020	96.5214
Sales to Equity Ratio	0.022	94.1772
Employee Profitability	0.018	93.9009
Loss	0.000	4.4778
Financing Activities	0.000	0.1753

Table 4. Random Forest Confusion Matrix

	High EM	Low EM	Very High EM	Very Low EM
High EM	1655	150	21	9
Low EM	374	493	11	16
Very High EM	102	12	186	0
Very Low EM	39	66	3	129
OOB estimate of error rate: 24.59%				Accuracy: 0.75

C5.0 Decision Tree Model

The decision tree model was built using discretionary accruals as the dependent variable, which is a proxy for earnings management and independent variables chosen using random forest. The observations were divided into training and testing set at a ratio 8 to 2. (80% to the training dataset and 20% to the testing dataset). Decision Trees can be used to analyze the link between variables and select the most important factors for predicting a specific outcome. The variables Loss and Financing Activities were removed after variable screening through Random Forest and the study was carried out with 17 variables. Decision tree algorithm was employed and the variables used in the construction of the tree are Operating Profit Margin, Corporate Size, Long Term Funds Appropriate Rate, Corporate Performance, Managerial Ownership, Operating Cash Flow. Initially, the model's training set accuracy was 80 percent, while the testing set accuracy was 71.6 percent. The accuracy was improved by tuning the hyperparameters. Hy- perparameter tuning was carried out using the grid search algorithm. The optimal values for the parameters were found by using grid search. RF + C5.0 The training set had 88 percent accuracy, whereas the testing set had 85.1 percent accuracy. The model had type I error rate of 14.89%.

Table 5. Decision Tree Confusion Matrix

	High EM	**Low EM**	**Very High EM**	**Very Low EM**
High EM	505	24	1	6
Low EM	46	195	0	6
Very High EM	24	3	46	3
Very Low EM	8	17	1	48

Table 6. Accuracy of RF + C5.0

	Overall Accuracy Rate	**Type I error**
Training Set	88%	12%
Testing Set	85.1%	14.89%

Fig.3 helps in determining the rules to predict the degree of earnings management carried out in a firm. Numerous rules can be formed using the model and few of the rules are mentioned below.

- When the Corporate Size is greater than Rs.290000 million (or) the managerial ownership is greater than 485000000 and corporate size is greater than Rs.520000 million and total assets turnover ratio is less than 0.27, the firm is said to have very high degree of earnings man- agement practices.
- Firms with corporate size less than Rs.12000 million and leverage coefficient less than 0.054 (or) firms with corporate size greater than 52000, total assets turnover greater than 0.27 and long-term funds appropriate rate greater than 1.7 are said to have high degree of earnings management practices
- Firms with corporate size less than Rs.12000 million, leverage coefficient greater than 0.054 and net profit margin greater than 48% are said to have low degree of earnings management practices.
- If the corporate size is less than Rs.12000 million, leverage coefficient is greater than 0.054, net profit margin is less than 48%, employee profitability is less than 10, long term funds appropriate rate is less than 5.4, operating cash flow is less than Rs.746 millionand total assets turnover ratio is greater than 2.1, the firm is said to have very low levels of earnings management practices.

Artificial Neural Network

The dataset was randomly split into training set (80%) and testing set (20%).Fig.4 represents the neural network created in the study. The neural network constructed has an input layer, two hidden layers and an output later. The input layer receives the variables whereas the output layer generates the prediction results. The hidden layers constitute neurons and serves to in- crease the complication of the network. Variable Importance for the neural networks was found using Olden's method. Fig.5 shows the variable importance plot for the neural network. The accuracy of the model was found to be 78.2%.

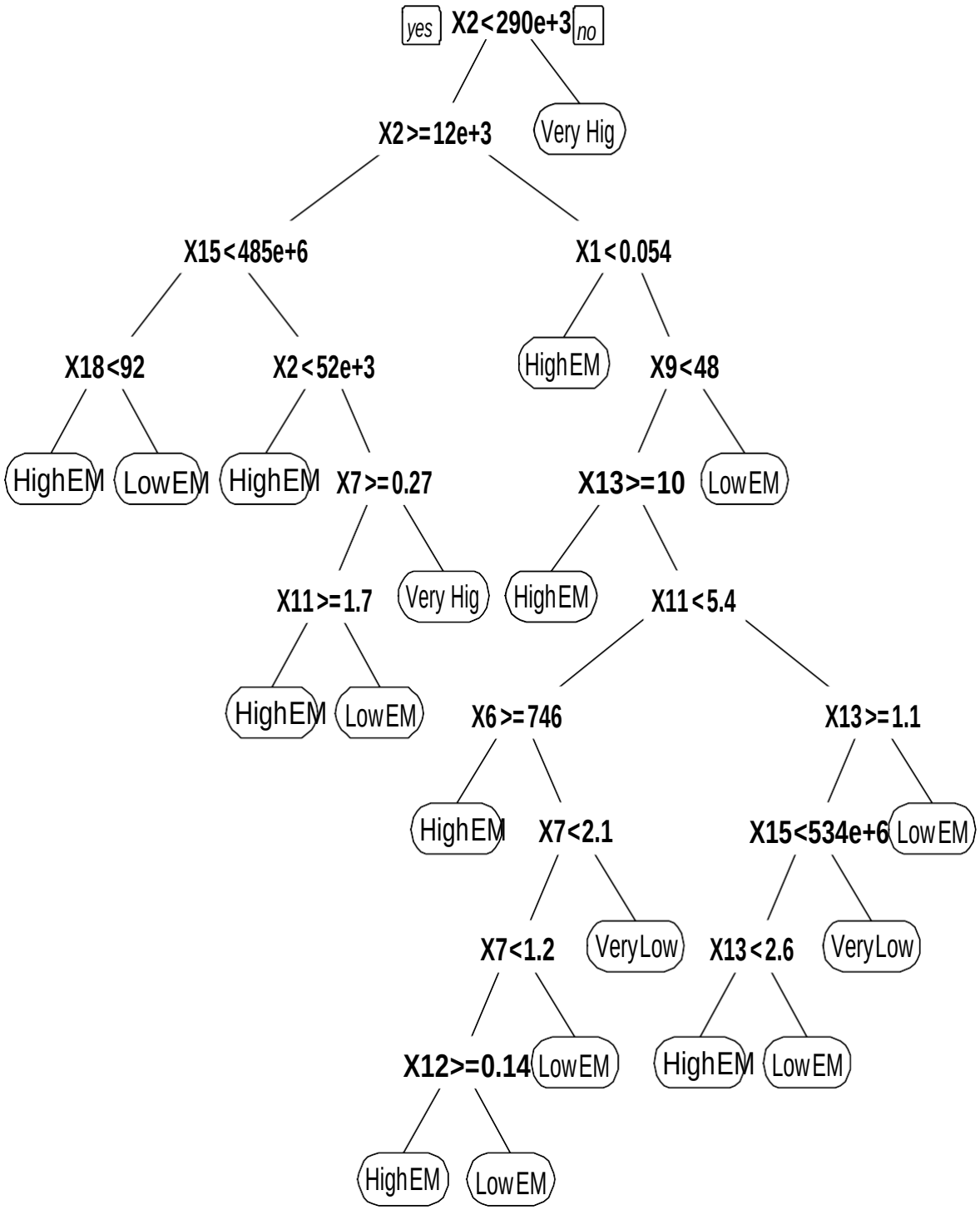

Fig. 3. C5.0 Model

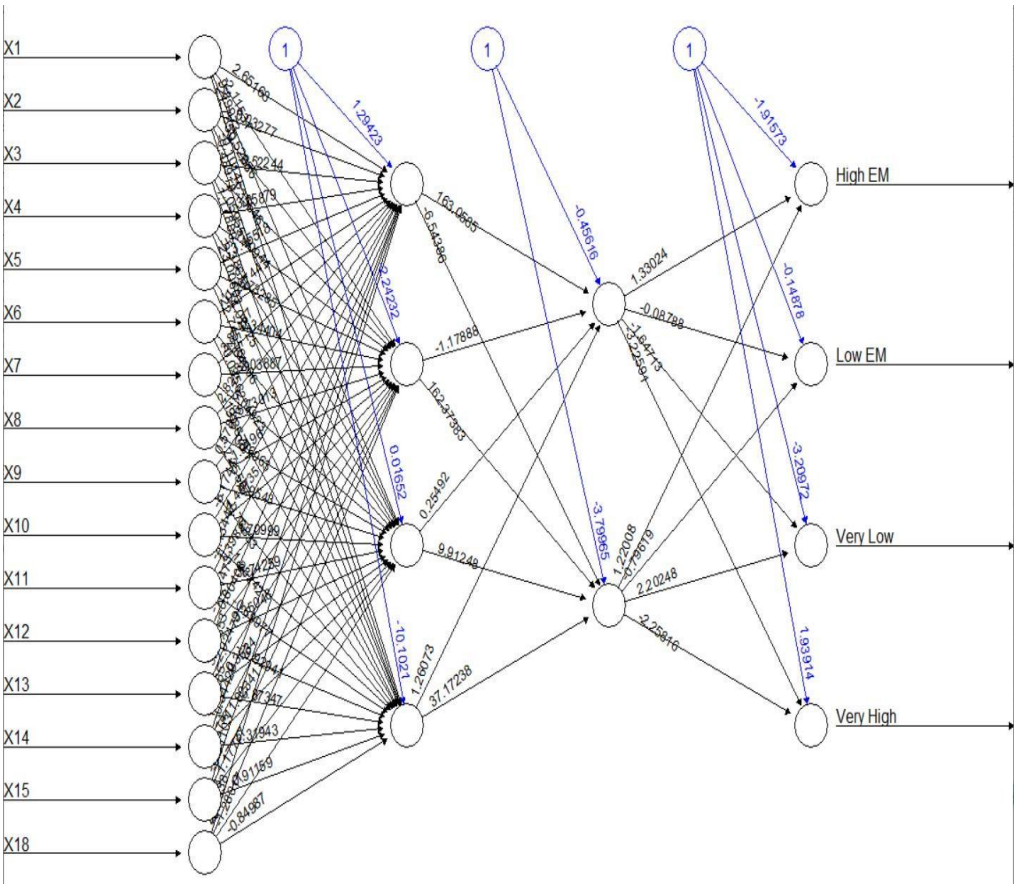

Fig. 4. Neural Network

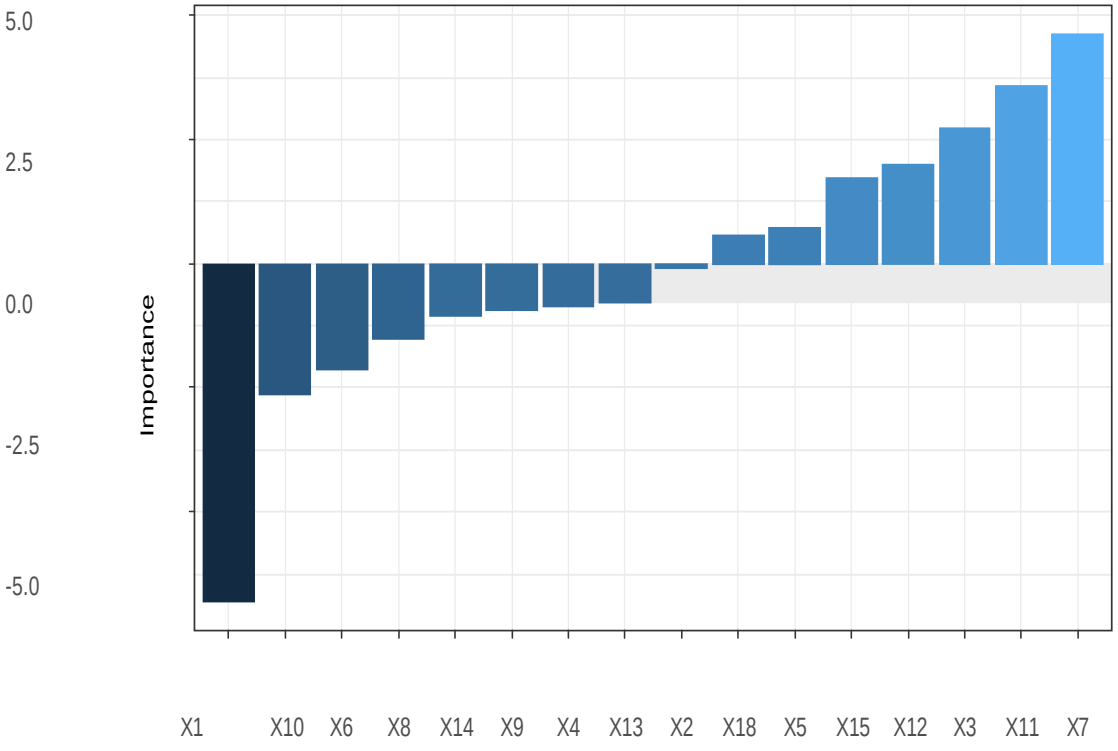

Fig. 5. Variable Importance in Neural Network

Discussion and Implications

Finding important predictive variables would be critical, since it would affect the model's ac- curacy and classification. It could be seen that the usage of random forest for variable screening improves the predictive accuracy of the earnings management prediction models. Variables such as Corporate Size, Leverage Coefficient, Operating Profit Margin and Total Assets Turn- over play a major role in determining the level of earnings management practices carried out by a firm as could be seen in the Fig 3. The results of the experiments shed light on why the recommended model is the optimal in this case. Variable Importance for Neural Networks found through Olden's method suggest that Corporate Size, Managerial Ownership, Return on Assets, Employee Profitability, Corporate Performance and Total Assets Turnover have a sig- nificant influence on the level of earnings management practices of a firm. The variable cor- porate size being an important parameter in earnings management prediction is obvious be- cause it represents the sales revenue of the firm which can be easily manipulated according to the convenience of the firm. Operating profit shows a company's ability to manage its indirect costs. Firms manipulate the operating profit by either accumulating the profit of all quarters in the final quarter or split the profit of the first quarter to all the other quarters to look more profitable (Rahimipour, 2017). Hence operating profit margin is one of the vital factors in in- fluencing earnings management. One issue with evaluating accrual-based earnings manage- ment models is the challenge of gauging the models' success because the exact amount of earn- ings management is unknown. To some extent, the use of discretionary accruals has assisted in overcoming this problem. The decision tree-based model is quite effective than the neural net- work model in terms of accuracy. Earnings management prediction models assist investors in making educated decisions by revealing the extent of earnings manipulation.

Table 7. Accuracy of C5.0 and ANN

	Overall Accuracy Rate	**Type I error**
C5.0	85.1%	14.89%
ANN	78.2%	21.5%

Implications for the Study

This study is preliminary research focusing on the development of a machine learning model based on C5.0 decision tree to predict accrual-based earnings management. This study pro- poses a method for detecting earnings manipulation that can help stakeholders in corpora- tions make more accurate judgments about their financial statements The sample for this study was the financial data of NIFTY 500 companies from 2012 to 2020. Random Forest method was used to screen the variables and c5.0 decision tree model was developed for the prediction of earnings management. The findings shows that the model constructed using a blend of ran- dom forest and c5.0 decision tree has a test group accuracy of 85.1 and a low rate of type I error. The neural network model had an accuracy of 78.2%. Though the performance of the hybrid model is good, the accuracy of the prediction model could be improved still. Further studies in this area could use more efficient machine learning methods to screen variables and could also consider using more financial variables other than the ones used in this paper. Future studies could consider the scenarios that the neural network model successfully predicts and use it to build a decision tree model that generates effective decision rules and compare the results with that of decision tree and neural network used separately to arrive at a more efficient prediction model.

References

1. *Chen, F. H., Chi, D. J., & Wang, Y. C. (2015). Detecting biotechnology industry's earnings management using Bayesian network, principal component analysis, back propagation neu- ral network, and decision tree. Economic Modelling, 46, 1-10.*
2. *Rodriguez-Ariza, L., Martínez-Ferrero, J., & Bermejo-Sánchez, M. (2016). Consequences of earnings management for corporate reputation: Evidence from family firms. Accounting Research Journal.*
3. *Kalbuana, N., Prasetyo, B., Asih, P., Arnas, Y., Simbolon, S. L., Abdusshomad, A., ... & Mahdi, F. M. (2021). Earnings management is affected by firm size, leverage and roa: Ev- idence from Indonesia. Academy of Strategic Management Journal, 20, 1-12.*
4. *Tahmina, A., & Naima, J. (2016). Detection and analysis of probable earnings manipulation by firms in a developing country. Asian Journal of Business and Accounting, 9(1), 59-82.*
5. *Kliestik, T., Belas, J., Valaskova, K., Nica, E., & Durana, P. (2021). Earnings management in V4 countries: The evidence of earnings smoothing and inflating. Economic Research- Ekonomska Istraživanja, 34(1), 1452-1470.*
6. *Chen, S., & Shen, Z. D. (2020). An effective enterprise earnings management detection model for capital market development. J. Econ. Manag. Trade, 77-91.*
7. *Mahmoudi, S., Mahmoudi, S., & Mahmoudi, A. (2017). Prediction of earnings manage- ment by use of multilayer perceptron neural networks with two hidden layers in various industries. Journal of Entrepreneurship, Business and Economics, 5(1), 216-236.*
8. *Chen, F. H., & Howard, H. (2016). An alternative model for the analysis of detecting elec- tronic industries earnings management using stepwise regression, random forest, and deci- sion tree. Soft Computing, 20(5), 1945-1960.*
9. *Huang, Y. P., & Yen, M. F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. Applied Soft Computing, 83, 105663.*
10. *Amel-Zadeh, A., Calliess, J. P., Kaiser, D., & Roberts, S. (2020). Machine learning-based financial statement analysis. Available at SSRN 3520684.*
11. *Namazi, M., & Maharluie, M. S. (2015). Detecting earnings management via statistical and neural network techniques. International Journal of Economics and Management Engi- neering, 9(7), 2020-2528.*
12. *Xinyue, C., Zhaoyu, X., & Yue, Z. (2020). Using Machine Learning to Forecast Future Earnings. Atlantic Economic Journal, 48(4), 543-545.*
13. *Huang, Y. P., & Yen, M. F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. Applied Soft Computing, 83, 105663.*
14. *Zaarour, B. D. I. (2017). Financial statements earnings manipulation detection using a layer of machine learning. International Journal of Innovation, Management and Technol- ogy, 8(3).*
15. *Naidu, D., & Patel, A. (2013). A comparison of qualitative and quantitative methods of detecting earnings management: Evidence from two Fijian private and two Fijian state- owned entities. Australasian*

Accounting, Business and Finance Journal, 7(1), 79-98.

16. *Fischer, J. A., Pohl, P., & Ratz, D. (2020). A machine learning approach to univariate time series forecasting of quarterly earnings. Review of Quantitative Finance and Account- ing, 55(4), 1163-1179.*
17. *Namazi, M., & Maharluie, M. S. (2015). Detecting earnings management via statistical and neural network techniques. International Journal of Economics and Management Engi- neering, 9(7), 2020-2528.*
18. *Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. Expert systems with ap- plications, 134, 93-101.*
19. *Tsai, C. F., & Chiou, Y. J. (2009). Earnings management prediction: A pilot study of com- bining neural networks and decision trees. Expert systems with applications, 36(3), 7183- 7191.*
20. *Höglund, H. (2012). Detecting earnings management with neural networks. Expert systems with applications, 39(10), 9564-9570.*
21. *Rahimipour, A. (2017). Investigating the Validity of Signaling Theory in the Tehran Stock Exchange: Using Real Earnings Management, Accrual-Based Earnings Management and Firm Growth. International Journal of Economic Perspectives, 11(3).*

*Corresponding Email: **sktrajan007@gmail.com**, **varadharajan@msrim.org***