International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

Customer Behavior Analysis using Weighted Selective Naive Bayesian

R.Siva subramanian, D.Prabha

ABSTRACT: The predominant success of each enterprise relies upon the efficient analysis of customer behavior and a deep understanding of customers' needs. Performing better customer analysis provides an effective analysis of potential customers, better decision-making, improved business processes, the measure of customers churn and enhance customer retention. Efficient customer analysis can be performed using the Naive Bayes (NB), which is a simple classifier used for predicting customer behavior. However, the performance prediction of Naive Bayes is strongly affected in some real-time datasets which involve the presence of correlated attributes and this creates a breach of the assumption made by the Naive Bayes model on the dataset. To enhance the performance prediction of the NB and to eliminate correlated variables, this research suggests a simple WSNB (weighted Selective Naive Bayesian) method that uses the C4.5 DT for selecting the attributes with high information values. Then the selected weighted attributes are further used with the Naive Bayes for improving performance prediction. The Experimental approach is tested with a bank client dataset, which indicates that WSNB makes better predictions than the standard Naive Bayes. Also, WSNB reduces the running time of the classifier by eliminating the correlated attributes, which in effect minimize the size of learning and testing data.

Keywords: Customer Analysis, Feature Selection, Decision Tree (DT), Machine Learning, Naive Bayes (NB), Prediction, WSNB

I. INTRODUCTION:

Customer research is a necessary factor for each business to consumer and business to business enterprises business planning. This research seeks to identify potential customers and identifying the depth of customer satisfaction with the enterprises.

The study of customer data will provide better marketing strategies, strengthened enterprise standards, enhanced services and increases customer acquisition & retention [4]. In customer analysis, the need for customer relationship management (CRM) is considered an important asset. CRM manages customer interaction with the enterprises and uses the stored historical data to analyze the potential customers and, in effect, to improve the business value with the customers.[2]. With a significant quantity of available data, better predictive modeling is required to anticipate future outcomes, a better insight into customer experience and better decision-making strategies. [13].

Revised Manuscript Received on February 28, 2020.
***** Correspondence Author

R.Siva subramanian*, Research Scholar, Anna University, Chennai, India. Sivasubramanian12@yahoo.com

DR.D.PRABHA, Associate Professor, Department of Computer science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India. Prabha@skcet.ac.in

The analysis of customer data can be performed using different machine learning strategies. In machine learning, classification is considered as one of the important tasks in decision aid for many fields such as pattern recognition, medical diagnosis, handwriting & speech recognition, and biometric identification. One of the ML supervised learning method NB, is a simple effective & efficient classifier that performs surprisingly in many classification tasks.**[** 1**]**.Consider L the learning set with $L = \{x_1, x_2, \dots \dots, x_n \mid y\}$ with n instances and where $\{x_1, x_2, \dots \dots, x_n\}$ refer to input attributes and y the class label. With the new observation samplex_i, the goal of the NB classifier to predicts the y class variable by

$$y = argmax_y(P(y|x_i)) \qquad \text{----- (1)}$$

Based upon the assumption made by NB that variables are independent of the class label,

$$y = argmax_y(P(y)\prod_{i=1}^{n} P(x_i \mid y)) \qquad \text{----- (2)}$$

Compare to other classification, NB can be easily constructed, fast, require less learning data, and highly scalable. The main primary assumption of the NB model is conditional independence- attributes in the dataset are independent of the given target label and all attributes in the datasets are equal.

This assumption makes to perform the goodness of the classifier. But in a certain condition, the performance of NB is degraded, this happens due to violation of conditional independence between attributes in some real-time datasets. The existence of redundant, irrelevant & noisy features lets the model perform poorly in prediction. **[6].** This makes researchers pay greater consideration to improve NB's performance prediction. Different researchers have proposed a different methodology to improve unpractical assumptions. In this research we suggest, the selection of features is considered to be a simple and effective technique for optimizing the prediction accuracy of the NB model. Feature selection is considered since the NB model cannot directly manage the correlated attributes during its learning or testing phrase. Feature Selection improves the performance prediction of the NB model by removing the correlated attributes from the datasets and makes the NB assumption to be satisfied with the dataset and helps to perform better. From the analysis of various procedures, we suggest a methodology called weighted selective naive Bayesian (WSNB), simple and effective procedure to improve the Naive Bayes performance prediction. The WSNB method is based upon selecting the features with high information values and from the selected attributes that are extensively applied to construct NB.

In the WSNB approach, for selecting the attributes $C4.5\ DT$ is applied. Using the learning set, the C4.5 model is used to construct a decision tree and weight is assigned to each feature based upon the depth of the features is present in the DT. This is carried out to measure the dependency of features with others. Optionally, features with lower values mean higher dependency with other features and from the build tree, features that appear in topmost depth with high values are considered as the relevant one. Then finally, with selected weighted features are used to construct the Naive Bayes. The suggested method proves prediction performance is improved compared to the standard Naive Bayes model.

II. RELATED WORKS

[8**]** apply SBC method to enhance the NB model, intending to improve prediction by eliminating redundant attributes in the datasets.

The methodology uses a greedy search with forwarding direction to elect the best attributes from the whole datasets.

The features selected are used for constructing NB and tested using UCI datasets. The experimental conducted are compared with $C4.5DT$, NB & SBC. Results reveal improvement in performance prediction of Naive Bayes is achieved using the SBC method. **[**16**]** Uses a method to eliminate the redundant attributes by applying preprocessing procedure (using decision tree) with NB. The methodology is to set weights to attributes by computing the degree of dependency with the other attributes. To compute the dependencies between attributes, DT is constructed and the weight of the attributes is computed based upon depth where attributes are tested. The lower weight is assigned to attributes that have more dependency and the attribute with high value is selected for evaluation of NB. The experimental method is tested using UCI datasets. **[9]** Uses SBC methodology to improve the NB performance prediction by combing C4.5 with NB. In the SBC method, the first construct $C4.5\ DT$ using the learning set and from the C4.5 tree, the first three-level features are selected as relevant attributes for NB evaluation. The idea is the attributes with high information values will appear in topmost of the tree. The experiment is tested using UCI datasets and results are compared with SBC, NB, C4.5, and ABC. [17**]**A new variable selection IIG proposed to use with NB to enhance the performance prediction. The IIG method is proposed to eliminate the drawbacks of IG (non- negative & Threshold value). The experiment is conducted using UCI datasets and findings are compared to IG and IIG. [18**]**New model called NBtree which is a hybrid method of DT and NB. The methodology is used to optimize the prediction in large datasets. The experimental procedure is conducted on datasets obtained from UCI and the results are compared with NBtree, NB & C4.5 .**[**11**]** A new improved version of Naive Bayes is being designed to boost the model without taking into account the conditional independence between the features. The correlation between the features is added through the edges. The proposed methods use conditional probabilities to identify dependency between the attributes. The empirical findings show that the performance prediction of Naive Bayes improves and also retains robustness.

[14**]** Explore a new method SNB(Semi-Naive Bayes) to join the features which have dependencies using the Chebyshev theorem.

[15**]** finds the dependencies between the features and joins the features using FSSJ and BSEJ. From the two procedures, BSEJ performs better.

III. NAIVE BAYES

Naive Bayes is simply straightforward and applied in a variety of classification problems. The primary advantage of NB is the independence assumption between the attributes and each attribute is considered as equal. The naive model performance may degrade if the presumption made by the classifier is violated. In some real-time domain datasets, the NB presumption is breached due to the presence of redundant attributes and also the presence of irrelevant, noisy and missing data makes the model perform poorly [9**]**. By eliminating these attributes in the datasets, the NB classification can be improved. To resolve the above problem with the redundant attributes, we use the C4.5 DT model with the NB model. The C4.5 model, using the learning the set construct the DT and weights are assigned to features based upon the level it depends on other attributes values. Then, the selected weighted features obtained from DT are further applied for NB evaluation to enhance the performance prediction. Using attribute weight in equation 3

$$y = argmax_y (P(y) \prod_{i=1}^{n} P(x_i \mid y)^{w_i} \qquad \text{----- (3)}$$

w_i Weight of feature x_i

IV. C4.5 DECISION TREE

The decision tree is non -parametric, top-down approach, induction and statistical classifier used for classification. C4.5 is an enhanced version of the ID3 classifier (by Ross Quinlan) and works well with missing, discrete & continuous data. Various reasons to use of decision tree include good interpretability, data processing, easily visualize and interpret. Decision tree can be labeled into ID3 (Iterative Dichotomisor), CART (Classification & Regression Tree), CHAID, MARS, C4.5, Conditional Inference Trees [10**].** In this methodology, the C4.5 DT is applied.

DT visualizes like flowchart tree structure and tree consists of leaves, branches, and root nodes.

The node appears in topmost is referred to as the top or root node, Internal root represents test on attributes, branch represents the test outcome and leaves represents the target label.

The construction DT is based upon the selection of attributes from the datasets. The random picking of attributes for modeling DT results in poor performance of prediction. Based on the splitting criteria, the attribute is selected for building the DT. The splitting criteria compute the information values for each attribute in the datasets and depend upon the values nodes is selected. In this research gain ratio splitting criteria are adopted and it is extended version of Information Gain and reduces its bias GR is computed by

$$\text{Gain Ratio} = \frac{\text{Gain(attribute)}}{\text{split info(attribute)}} \quad \text{-----} \quad (4)$$

Split information of features is measured using

$$split\ info(D) = -\sum_{j=1}^{v}\left(\frac{|D_j|}{|D|}\right)\log_2\left(\frac{|D_j|}{|D|}\right) \quad \text{-----} (5)$$

Gain for an features is computed using

$$Gain(A) = I(D) - E(A) \quad \text{-----} (6)$$

$$E(A) = \sum_{i=1}^{n} \mathrm{I(D)} \frac{\mathrm{d_{1i}+d_{2i},+\cdots+d_{mi}}}{\mathrm{d}} \quad \text{-----} (7)$$

$$I(D) = -\sum_{i=1}^{n} p_i log_2 p_i \quad \text{-----} (8)$$

In C4.5 classifier, features with more information values are taken as node and by applying the recursive method to form initial DT [**7**].

A. Pruning

Using the learning set, C4.5 DT is completely grown and learning sets are possibly classified correctly. With the constructed decision tree there may be possible of over fitting risk with the large data and also badly generalizing to new instances [**5**]. Over fitting occurs due to reduced error in learning set and increased error in test instances. The process of obtaining optimal size DT is a very important one, to reduce over fitting issues and enhance the prediction accuracy of DT. To eliminate the over fitting issues, C4.5 applies the pruning procedures. Pruning can be performed either earlier (pre pruning) or after constructing the DT(post-pruning). In our study post pruning is applied in C4.5 DT.

First, the DT is constructed completely and pruning (post) is performed from the bottom to up approach. Pruning aims to minimize the size of the DT by eliminating some branches of the DT which do not contribute more information on the classification of instances.

Pruning performs in such a way to decrease the prediction error rate, the complexity of classifier, reduces over-fitting of the tree and removes the features which provide a small amount of information to classify data. Pruning is performed from bottom to top [3].

V. WEIGHTED SELECTIVE NAIVE BAYESIAN (WSNB):

The basic design of Naive Bayes is conditional independence between the attributes and the model is oversensitivity to correlated and irrelevant attributes. The presence of such attributes in the datasets makes the NB classifier to perform poorly. While in the same correlated attributes datasets, decision tree performs wisely. This is possible due to, the redundant features are not possible to split into the learning data because the two correlated attributes will get more weight in the decision to the class which they belong and DT will consider one set redundant feature for splitting in the learning set. From the constructed tree using a C4.5 DT, some branches may hold noisy and over fitting in the learning set. To remove such branches, C4.5 employs pruning methods to remove such branches from the tree. Then from the pruned $C4.5\ DT$ topmost features are considered because the features in the topmost layers hold higher information values. These selected weighted attributes would help to enhance the Naive Bayes model performance prediction and with reduced running time.

A. Algorithm (WSNB):

1.WEIGHTED SELECTIVE NAIVE BAYESIAN - procedure ($D =$ Dataset, $t_n =$ number of trees, $d = depth\ of\ the\ tree$)
2. Dataset D is spitted into learning set T & testing set
3. Learning set T is split to $\{t_1, t_2, t_3, \dots, t_n\}$ learning subsets
3.1 Learning set $T = \{x_1, \dots, x_n \mid y_n\}$
3.2 Sample randomly (replacement) from T with n instances
3.3 add n to t_1 (repeat upto t_n $times$)
4. for i ranges from $t_1$to t_n learning subsets do
5.Learn a C4.5 (DT) from the learning subsets i
6. Record the weight of the attributes as1/ $\sqrt{d}$ (d $is\ the\ depth\ of\ the$ tree) and if the feature is not tested then record weight as zero.
7. for x ranges from 1 to d level in DT do
8. Place features on layers x into the selected attribute set.
9. then calculate final attributes weight(for the selected depth) as the average weight of $t_1$to t_n DTs
10. Remove the features appear beyond the selected depth.
11. Lastly, run an NB classifier using the weighted selected attributes

The dataset is spitted into Learning T and testing set. Both data are mutually exclusive (Testing set are not used in the learning and learning set are not used for testing). Then form the learning set T using bagging (ensemble) procedure the learning set is split into $\{t_1, t_2, t_3, \dots, t_n\}$ learning subsets. Randomly sampling from the original learning set $T = \{x_1, \dots, x_n \mid y_n\}$ with replacement, n percentage instances are added to t_1. Then the procedure is repeated up to generating $\{t_1, t_2, t_3, \dots, t_n\}$ learning subsets. The use of bagging procedure helps to reduce the variance in the decision tree and possibly avoid over fitting. The ensemble method aims to build a strong decision tree by combining the weak learns into a single DT and possibly improving the accuracy of the classifier. A small proportion of n instances is chosen to build a DT. If the percentage of data split is too larger, then time complexity increases and moreover decision tree becomes more complex. Then, C4.5 DT is build for each learning subsets $\{t_1, t_2, t_3, \dots, t_n\}$ and from the constructed C4.5 DT, using 1 \ $\sqrt{\mathrm{d}}$ ($d\ depth\ of\ the\ tre$) the weight is assigned to the attributes at the level tested.

The weight is assigned to the attributes to check how much it depends upon other attributes values. Higher weights are assigned to attributes which are highly correlated with the class label and lower weight is assigned to the features which have many dependencies with the other predictors. The attributes will receive zero weight if do not exist in the C4.5 DT.

x denotes the depth of features that are selected into the attribute set. For example, if the x denotes depth one means the features in that layer are considered and if x denotes depth two means the attributes in the depths 1& 2 are considered. Optionally any level can be chosen, but getting deeper trends to get more volume of features selected with less weight.

So for that reason features appear in the topmost layers are considered which has high information values. Then, the final feature weight is computed using the aggregation of the average weight of $\{t_1, t_2, t_3, \ldots, t_n\}$ the DT for the selected depth. Remove the attributes that do not appear in the selected d depth and attributes with the zero weight. Then with the Weighted Selected attributes are considered for the assessment of Naive Bayes. This procedure possibly makes to improve the performance prediction of Naive Bayes, since the attributes that are lower information values and correlated, irrelevant, noisy/missing attributes are eliminated.

VI. EXPERIMENTAL DESIGN

A. Data Sets:

The evaluation of the WSNB procedure is experiment with the bank customer datasets obtained from the UCI respiratory. The data sets are linked to a direct marketing drive and consist of 45211 instances with 16 input attributes and one output variable with two classes[19]. The experimental results are compared with WSNB and standard Naive Bayes.

B. Experimental Procedure (WSNB)

1. The dataset D is spitted into a learning set T (70%) and testing set (30%) (Step 2)
2. From learning set T, $t = 10$ learning subsets are generated and n =10 percentage instances in each learning subsets(with replacement) (step 3)
3. Then, $C4.5$ DT is constructed for the $i = 10$ learning subsets (step 4 & 5)
4. Weight is assigned to the attributes accordingly to level the attributes are tested using $1/\sqrt{d}$ (step 6)
5. Top three layers features($x = 1$ to 3) are considered as selected attribute set (step 7)
6. From the weighted selected attributes set for each depth from ($x = 1$ to 3) , the features set are used to construct the Naive Bayes Model.
7. Performance metrics from each depth is noted and results are compared with standard Naive Bayes

C. Validity Scores

The performance metrics are compared using various scores that include true positive and negative, false positive and negative **[12]**

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{----- (9)}$$

$$\text{Sensitivity or Recall} = \frac{TP}{TP+FN} \quad \text{----- (10)}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad \text{----- (11)}$$

$$\text{Positive Predictive} = \frac{TP}{TP+FP} \quad \text{----- (12)}$$

$$\text{Negative Predictive} = \frac{TN}{Tn+Fn} \quad \text{----- (13)}$$

$$\text{False Positive Rate} = \frac{FP}{FP+TN} \quad \text{----- (14)}$$

$$\text{False Negative Rate} = \frac{FN}{FN+TP} \quad \text{----- (15)}$$

where FN measure amount of false negative, FP measure amount of false positive, TP measure amount of true positive and TN measure amount of true negative.

D. Experimental Results

The experiment is conducted for two methodologies, one is WSNB method (section 6.2) and other one is Standard Naive Bayes. The findings are shown below

Table 1: Results of Accuracy, Recall & Specificity for Naive Bayes and WSNB (features selected at depth 1,2 &3).

Classifier	Accuracy	Sensitivity	Specificity
Naive Bayes	8.796	0.52	0.92
WSNB(depth 1)	8.904	0.33	0.96
WSNB(depth 2)	8.902	0.45	0.949
WSNB (depth 3)	8.871	0.47	0.942

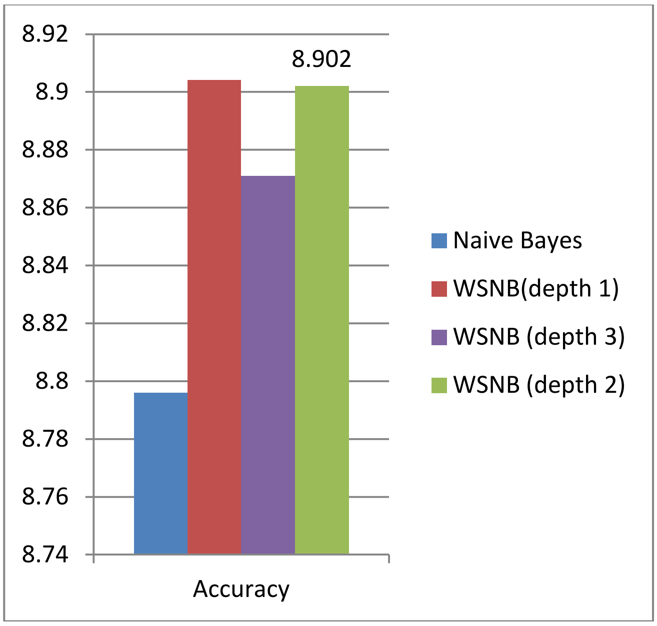

Fig 1: Accuracy comparison of Naive Bayes and WSNB (with depth 1,2,& 3selected features)

Table 2: Results of PPV, NPV, FNR & FPR for Naive Bayes and WSNB (features selected at depth 1,2 &3).

Classifier	PPV	NPV	FNR	FPR
Naive Bayes	0.49	0.934	0.47	0.07
WSNB(depth 1)	0.57	0.91	0.66	0.03
WSNB (depth 2)	0.54	0.92	0.54	0.050
WSNB (depth 3)	0.53	0.930	0.52	0.057

Retrieval Number: D1636029420 /2020©BEIESP
DOI: 10.35940/ijitee.D1636.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

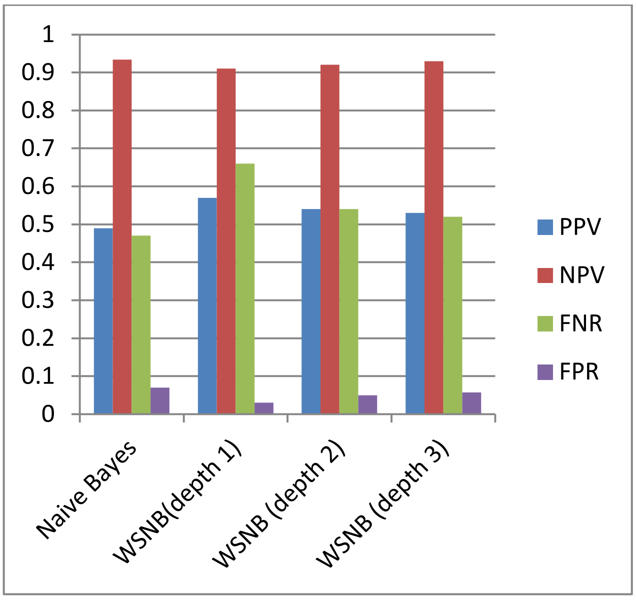

Fig 2: PPV, NPV, FNR & FPR comparison of Naive Bayes and WSNB (with depth 1, 2, & 3selected features)

Table 3: Improvement of accuracy WSNB VS Naive Bayes

WSNB(depth 1) vs. Naive Bayes	WSNB(depth 2) vs. Naive Bayes	WSNB(depth 3) vs. Naive Bayes
+1.08	+1.06	+0.75

E. Result Discussion

The experimental results obtained for WSNB method and Standard Naive Bayes are tabulated in table 1 and table 2, using 70% of instances for training and remaining 30% instances for testing. The table 1 shows WSNB at all depth obtains improved prediction compare to standard Naive Bayes. The WSNB at depth 1 obtained highest accuracy of 8.904 and followed by at depth 2 and 3. But, the standard NB obtains maximum of 8.796 accuracy only and WSNB at depth 1 obtains +1.08(table 3) more accuracy compared to standard Naive Bayes.

The specificity gets improved prediction in WSNB at depth one by 0.96 where standard NB gets only 0.92. But the sensitivity value gets higher in standard NB compare to WSNB. The table 2 indicate WSNB at all depth gets improved results compare to standard NB and WSNB at depth one obtains highest prediction compare to rest of other two depths. The experiment results reveal, C4.5 classifier chooses the best attributes (at topmost level) and the use of that features possibly makes NB to improve performance prediction. The execution time of WSNB is lesser due to the usage of relevant attributes for building Naive Bayes and eliminating the redundant, missing/noisy attributes. WSNB shows better running time compared with the standard Naive Bayes. The WSNB running time for WSNB at depth one is 0.03 seconds. WSNB at depth two is 0.06 seconds and WSNB at depth three is 0.15 seconds. For standard Naive Bayes running time is 0.16 seconds. This clearly indicates the WSNB reduces the running time of the model compared with standard NB. C4.5 DT uses only 10 percent of the data from the original learning dataset T for constructing decision tree. This makes the goodness of the C4.5 and with less learning data C4.5 builds simple DT and avoids complex one.

VII. CONCLUSION

Performing effective customer analysis can be carried using a Naive Bayes. But, the existence of correlated, noisy, irrelevant attributes in the datasets makes the NB model perform poorly. To enhance the performance prediction of the NB classifier, WSNB methodology is suggested. The approach uses C4.5 DT for electing attributes with high information value for constructing a Naive Bayes model to improve the prediction.

The WSNB approach performs effectively and uses the two different classifications methods. The Use of bagging methods helps to lessen the variance in the decision tree and form a strong decision tree by combing multiple single DT models into strong DT. With the few learning data DT find the best relevant attributes and the method perform fast. Assigning a weight to attributes helps to finds out the higher attributes in the tree and possibly makes to select attributes in the depths. WSNB approach performs less execution time compared to standard NB. This is due to the removal of redundant attributes from the datasets that make the learning and testing data to execute in less running time. The methodology of using C4.5 with WSNB for electing attributes to use for constructing Naive Bayes is illustrated. The result analysis reveals WSNB a simple method trend to enhance the performance prediction of the NB by using attributes selected through the C4.5 model. The WSNB achieves higher accuracy at depth one 8.904 where the standard Naive Bayes achieves accuracy of 8.796. This shows the WSNB model performs better prediction compare to standard Naive Bayes.

REFERENCES

1. Abellan, J. & Castellano, Francisco, "Improving the naive bayes Classifier via a Quick Variable Selection Method Using Maximum of Entropy," Entropy . 19.10.3390/e19060247, 2017.
2. Adrian Payne & Pennie Flow, " A Strategic Framework for Customer Relationship Management ," Journal of Marketing. 69(4), 2005.
3. Badr Hssina et.al, " A Comparative study of Decision tree ID3 and C4.5," (IJACSA), Special Issue on Advances in vehicular Ad Hoc Networking and Applications 2014. http://dx.doi.org/10.14569/special issue.2014.040203.
4. Christry,A.J., et al, "RFM ranking - An effective approach to customer segmentation," Journal of king Saud University Computer and Information Sciences , https://doi.org/10.1016/j.jksuci.2018.09.004, 2018
5. Chawla Nitesh, "C4.5 and Imbalanced datasets: Investigating the effect of sampling method, Probabilistic estimate, and decision tree structure," 2003
6. Domingos P & Pazzani, "M. Machine Learning 29: 103," https://doi.org/10.1023/A:1007413511361, 1997
7. F.Duan, Z.Zhao & X.Zeng," Application of Decision Tree Based on C4.5 in Analysis of Coal Logistics Customer, " 3rd International Symposium on Intelligent Information Technology Application, Shanghai, pp 380-383, 2009. Doi:10.1109/IITA.2009.245.
8. Langley .P & Sage .S, "Induction of Selective Bayesian Classifiers," Uncertainty Proceedings,399-406, 1994 Doi:10.1016/b978-1-55860-332-5.50055-9
9. Ratanamahatana, Chotirat & Gunopulos Dimitrios, "Feature Selection for the Naive Bayesian Using Decision Trees," Applied Artificial Intelligence.17.475-487, 2003. DOI:10.1080/713827175.
10. Rahmadani S et.al," Comparison of Naive Bayes and Decision Tree on Feature Selection Using Genetic Algorithm for Classification Problem, " Published under license by IOP Publishing Ltd, Volume 978, Conference

11. Sona Taheri, Musa Mammadov , Adil M.Bagirov," Improving Naive Bayes classifier using conditional probabilities, ".9th Australasian Data Mining Conference.121.63-68- volume 121, 2009
12. Dhandayudam Prabha & Krishnamuthi Ilango," Customer Behavior Analysis Using Rough Set Approach," Vol 8.Pg 21-33, 2013.DOI:10.4067/S0718-18762013000200003.
13. Zeynab Soltani, et.al, "The impact of the customer relationship management on the organization performance," The Journal of High Technology Management Research, vol 29,Is 2, Pg 237-246, 2018
14. Kononenko I," Semi-naive bayesian classifier, " In: Kodratoff Y. (eds) ML — EWSL-91. EWSL 1991,pp.206-219.(Lecture Notes in AI), vol 482,1991, Springer, Berlin, Heidelberg. DOI: 10.1007/BFb0017015.
15. Michael J.Pazzani, "Searching for Dependencies in Bayesian Classifiers," Learning from Data: AI and Statistics. DOI: 10.1007/978-1-4612-2404-4_23
16. Hall, Mark," A Decision Tree-Based Attribute Weighting Filter for Naive Bayes," 2007, DOI:10.1007/978-1-84628-663-6_5.
17. Abellán, J. & Castellano, Francisco," Improving the Naive Bayes Classifier via a Quick Variable Selection Method Using Maximum of Entropy,". Entropy. 19, 2017 DOI: 10.3390/e19060247.
18. Kohavi, Ron. (1997),"Scaling Up the Accuracy of Naive-Bayes Classifiers: a Decision-Tree Hybrid. KDD,".
19. S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, Elsevier, 62:22-31, June 2014

AUTHORS PROFILE

Dr.D.PRABHA, is an Associate Professor in Department of Computer Science and Engineering at Sri Krishna College of Engineering and Technology, Coimbatore, Tamilnadu, India. She has completed her Under Graduate Studies and Post Graduate Studies in Computer Science and Engineering. She has received her Ph.D in Information and Communication Engineering from Anna University, Chennai in the year 2014. She is presently guiding 8. Ph.D. research scholars registered under Anna University, Chennai. Her research interests include Data Mining, Big Data Analytics and Soft Computing. She has published several papers in premier indexed journals.

Mr.R.SIVA SUBRAMANIAN, is currently working as an Assistant Professor in Jawahar Engineering College; Chennai He is also pursuing his part time Ph.D in Anna University, Chennai. He has completed B.Tech IT in Srinivasa Institute of Engineering and Technology, Chennai. He has obtained M.Tech IT from Bharath University, Chennai. His research domain includes Data Mining, Customer Relationship Management and Big Data