ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

International Journal of Mechanical Engineering

Customer segmentation in intelligent learning mechanism in E- banking using Frequent Item-set Hierarchical Clustering

M.Rajeswari[1], LivyaGeorge[2], Dr. D.Brindha[3], D.Vaduganathan[4], A. Suresh Kumar[5]

[1]Associate Professor, Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences,

[2]Assistant Professor, Department of Computer Science and Engineering, Sahrdaya College of Engineering and Technology,

[3]Assistant Professor in the Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, brindha@karunya.edu,

[4]Assistant Professor, Department of Computer Science and Engineering, Karpagam Institute of Technology, Coimbatore,

[5]Assistant Professor, Department of Computer Science and Engineering, Excel Engineering College (Autonomous), Erode,

Abstract

Direct marketing is a business model that uses data mining techniques and marketing databases for personalization and business intelligence. It is a new approach that uses interactive one-to-one communication between marketer and customer aimed at specific customers with personalized advertising and promotional campaigns. There are few attempts of using Data Mining for Direct Marketing. The major problems encountered are mining of the huge volume of customer topographies for marketing purpose, churn management and deficiency of binary classification algorithms. In this paper, we propose a novel and efficient learning algorithm called CSBC for intelligent Learning based Direct Marketing using Frequent Item-set Hierarchical Clustering which can be used by the marketers to personalize the next campaign. Frequent-Itemset Hierarchical clustering and ranking method is incorporated in CSBC to grade/segment the customers into more than two classes. A classifier produced rules would be used to predict the new customer data. Also, the marketers will be able to predict the churners by evaluating the predicted classes/ranks along with the relationship on the client information attributes of dataset. Prediction of respondents and non-respondents from new data which will be useful for direct marketing is done using fully automated knowledge extraction procedure. Customer Segmentation is evaluated using lift measure and ROC curve measure in experiments.

Keywords: Data Mining; FIHC; Direct Marketing; Binary Classification algorithms; Customer Segmentation; Banking

1. INTRODUCTION

Banking has innumerable types of data with numerous data records [20]. To use the available massive data effectively for the marketing personnel, it is mandatory to arrange the data so that it is helpful for the person who has no knowledge about the data mining or the programming languages. The non-technical person in banking domain could be benefited by the concealed knowledge. Managing the actual execution of knowledge extracted by the target marketing (Direct marketing) prediction model. There are two perspectives to advertisement and promotion in general which includes mass marketing and direct marketing. It is observed that in the competitive market, direct marketing has become highly effective which examines the customer characteristic needs and chooses decisive customers as the goal for promotion.

The main aim of the paper is for Direct Marketing to make the marketing effective even though it has a few downsides, for example it may lead to interruption of privacy [13]. Direct marketing is a business model that uses data mining techniques and marketing databases for personalization and business intelligence. It is a new approach that uses interactive one-toone communication between marketer and customer. It is aimed at specific customers with personalized advertising and promotional campaigns. It involves investigating the customers' characteristics and needs to identify customer market value and select customers most likely to respond to promotions. This method is progressively establishing itself as the preferred option for companies because of its low cost and high profitability. If we give more data for training, it may lead to generalize the new unseen data.

Data mining techniques are useful for portraying customers based on their response in campaigning surveys. Various techniques like clustering and classification can be applied for finding characteristics of people in banking domain like great indicators of buying practices on banks products such as term deposits, savings accounts, current accounts etc. The customers can be segmented based on their behaviours during campaigning. This must be done with efficiency in time. An examination announces that subsequent to applying affidavit of vast yearly salary rewards, a few clients need to trust that the cash will be credited into their shared reserve accounts outside the bank. This speaks to loss of trade for the bank. These clients should feel advantageous to keep their cash in the bank, advertising experts can utilize campaign management software to quickly distinguish substantial deposits and set off a reaction. The marketer requires an automatic mail to be sent to the customer as soon as it reaches the threshold level [9].

There are few attempts of using Data Mining for Direct Marketing. Varshney and Mojsilovic [1] demonstrated the need of knowledge discovery in a business organization. They presented various business analytical applications which include revenue opportunity estimation, sales recommender system, skills capacity planning, workflow evaluation and optimization. Albers and Clement [2] analyzed success drivers for e-business systems. Indices and indicators were tabulated for the organization of success. Operationalization of success indices and market strategy indices were defined along with business model and covariates. Model validation was done using the impact of market strategy indices and the impact of business model. Ling and Li [3] demonstrated data mining for direct marketing with the problems and solutions. The process of data mining for direct marketing, dataset collection and processing and specific problems which need to be focused in data mining were elaborated. Flici [4] proposed a dependable framework for direct marketing which uses data mining technique for marketers. They constructed a Direct Marketing Framework (DMF) which makes use of combination of different data mining methodologies and direct marketing concepts.

Moro and Raul [5] stated the effect of using CRISP-DM model consisting of business understanding, data understanding, and data pre-processing which is defined by preprocessed dataset, modelling, evaluation and deployment. Tsiptsi and Chorianopoulos [6] discussed the importance of customer segmentation along with the processing of applying segmentation in Retail Banking. They explained the data mining techniques used for customer segmentation in Customer Relationship Management (CRM). Setnes and Kaymak [7] stated the concept of fuzzy clustering technique which was utilized in target determination from substantial databases for direct advertising purposes. A general cycle of information mining venture utilizing CI technique which includes requirement and feasibility analysis, application domain, data access, data preparation, data exploration, Applying CI techniques and evaluation were elaborated by the authors. Fuzzy data mining along with modelling for target selection was determined using data access, preparation and so on. Au and Chan [8] considered the problem of discovering fuzzy association rules using objective interestingness measures for which the rules were presented to the marketers. The authors portrayed a fuzzy method along with its application to an information mining which includes a huge database that was furnished by a global bank with workplaces in Hong Kong.

Fung et al. [9] created hierarchical document clustering strategy with the best in class agglomerative, partitioning and successive itemset based techniques. They proposed utilizing the idea of regular itemsets, which originates from association rule mining, for document clustering. The instinct of that grouping foundation is that each cluster is distinguished by some basic words, called frequent itemsets, for the records in the cluster. The authors clarified that data mining and association mining, analysis of cluster and various hierarchical clustering. Different regular itemset-based strategies were also expounded by the authors. Other clustering techniques such as density-based and grid-based methods were also defined. Yin et al [10] proposed a greedy heuristic algorithm with ensemble-based decision-tree algorithm which was used to apply in a collection of decision trees for producing appropriate rules. Using different functionalities of data mining, the positives of the client profile can be detected. The combination of DM functionalities may result in good accuracy as well as cost (Time) effectiveness. It is viable to improve the quality of the knowledge discovered from the business data by adapting the Frequent Itemset Hierarchical Clustering (FIHC) with the decision tree techniques [5].

The main aim of this paper is to construct a novel and efficient learning algorithm for Direct Marketing using Data Mining Techniques (DMT) in the Banking domain which can be used by the marketers to personalize the next campaign. It discovers the predictive insights in the databases and turns them into operational reality. The prediction provided by a model is usually called a score. A score (typically a numerical value) is assigned to each record in the database and indicates the likelihood that the customer with scored records will exhibit a particular behavior. For example, if a model predicts customer attrition, a high score indicates that a customer is likely to leave, whereas a low score indicates the opposite.

The related methodologies used binary classification algorithms; however, binary classification can only predict two classes such as responders or non-responders based on the attribute values. Thus, there is a chance of neglecting valuable customers which could be considered for different methods of marketing. Hence, it is advisable to use direct marketing technique which grades the customers in different classes instead of two classes. Frequent Item-set Hierarchical Clustering (FIHC) [6] would be used for grading the dataset i.e. each transaction will be categorized as good or bad for which the decision tree technique could be used for predicting the client profile. Thus, the mixture of the three functionalities leads to the knowledge discovery in the client datasets or profiles available.

The proposed CSBC method attempts to rank the customer transactions so that different methods of campaigning will be done according to the rank. It tries to automate the processing of the large dataset with no or minimal human intervention. It is made to provide churn management.

The rest of the paper is organized as follows: Section 2 details the proposed system design, the evaluation is encompassed in section 3, and section 4 has the conclusion.

2. THE PROPOSED SYSTEM DESIGN

Fig. 1 explains the general overview of the proposed system. The input to the system is the binary classified data. The data pre-processing step involves preparation of data before giving to the FIHC algorithm. It also uses the sub-attribute utilization explained in [12] for improving the association rules obtained for the creation of clusters. Clustering is applied in FIHC module in order to produce the disjoint clusters which will be given as the input for ranking module. These disjoint clusters are ranked with respect to the quality of discrimination based on alpha, beta and gamma values. Also, decision tree module will be given the ranked clusters as input in order to generate the rules. Hence, the proposed system will be tested using the datasets available which is elaborated in the next subsections.

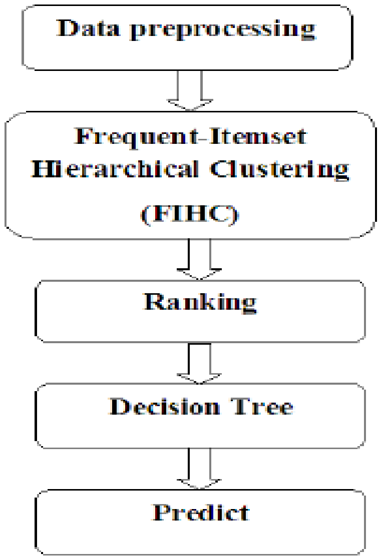

Fig. 1 CSBC -Design Overview for Learning based E-banking

Data Pre-processing

Fig. 2 presents the Pre-processing module with expected input and output. The training data with respect to bank deposit subscription were extracted from [19], which consists of ".csv file" that includes all the required attributes. The dataset gives information about the previous campaign and the corresponding results. The dataset is pre-processed to convert the discrete values of duration, balance etc into different subjective values such as low, medium and high.

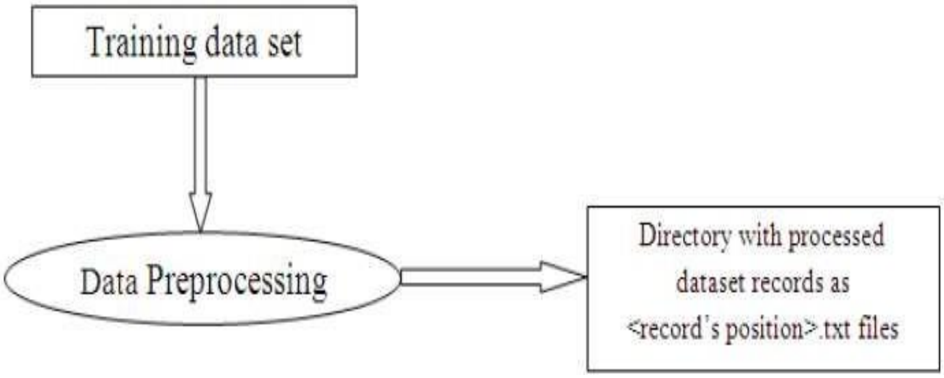

Fig. 2 Detailed Data Processing Step

Each transaction is further translated into a file and saved in a folder which is to be utilized by FHIC later. Each file in the folder contains the attribute values with their name. For Example: "Job" = "$Entrepreneur$","$marital$" = "$married$"

Data have their own characteristics which are considered to be valuable. Hence, it is important to preserve the values of these attributes in an effective manner. Also, the subattribute utilization method which was elaborated by [12] includes binary, symbolic, and continuous values to maintain the characteristics of data. It is desired to use α, β, γ for binning the values of attributes. For ex, duration, balance which needs continuous analysis in nature, the values are binned into "low","$medium1$","$medium2$","$high$"

Frequent Itemset Hierarchical Clustering(FIHC)

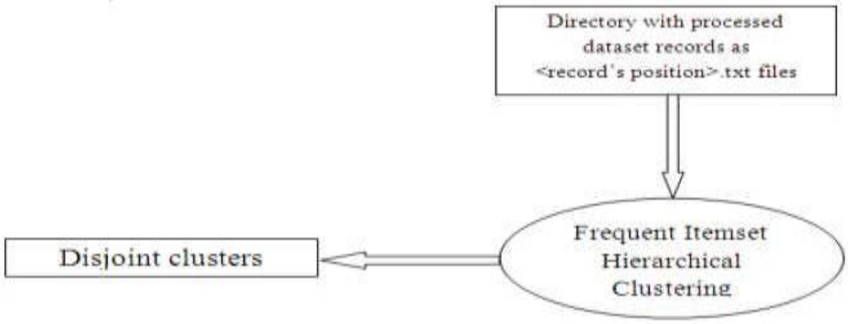

Fig.3 Detailed Frequent Itemset Hierarchical Clustering

Fig. 3 represents the FIHC module with required input and the output and the result produced from the dataset. An experimental result which was given by [9] reveals that this proposed approach FIHC performs well with respect to accuracy. It has been observed that this FIHC is robust when the dataset is getting enlarged. It performs clustering and place various input .txt files into the respective folders.

Algorithm:

The parameters α, β, and γ of attribute Ai are set as follows:

α = average value of attribute Ai in the database; β = the largest value of attribute Ai in the database; α + β = 2 γ.

For continuous attributes like duration, balance, etc., the values are binned into "low", "medium1", "medium2", "high". This is done using the following formulas.

Ai=Attribute

```
α(Ai)=b(Ai) = average(values of Ai);

β(Ai)=c(Ai)=max(values of Ai);

γ(Ai)=a(Ai)=2*b-c;

if(value of A<a(Ai))

value of Ai=low;

else if( a(Ai)<value of A<b(Ai) )

value of Ai=medium1;

else if(b(Ai)<value of Ai<c(Ai)

value of Ai=medium2;
```

Ranking

Fig. 4 represents the Ranking module which includes its appropriate input and output. The disjoint clusters obtained from previous module are given as the input to this module. The ranking is done on clusters according to the quality in the records based on the attribute values. The quality of each record is determined based on the score value. The following subsections 2.3.1 and 2.3.2 explain the scoring and the process of assigning ranks respectively.

Scoring

A unique value is assigned to each record in the database which denotes the particular behavior of the customer. The attribute values of call duration, last contact month, number of previous contacts, days since last contact, last contact result and first contact duration maintain the characteristics of that cluster which will be used for ranking. The cluster with highest score is ranked 1 and the subsequent clusters are ranked accordingly. Lift is a number representing the increase in responses from a targeted marketing application using a predictive model over the response rate when no model is used.

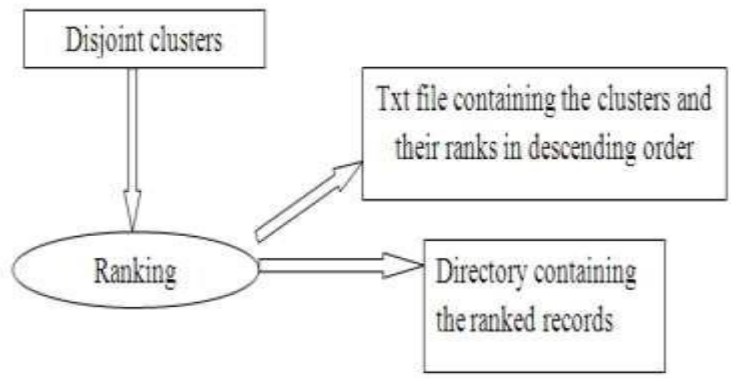

Fig. 4 detailed Ranking

The score for the cluster will be depending upon the history information suggested by [5].

```
For (each cluster)

{

    Score = ( number of transactions with call duration

    =(medium2 || high))+

    (number of transactions with month= largest probability of success)+

    (number of transactions with pdays=(low || medium2))+
```

(number of transactions with job=(employed))+

(number of transactions with housing =yes)+

(number of transactions with education =(secondary||tertiary)+

(number of transactions with balance = (medium2 || high))

}

Assigning Ranking

The clusters are sorted according to the score in the descending order. The clusters which have position higher will have the higher rank. The cluster with highest score will get the rank 1. The second highest will get rank 2 and so on. The cluster with rank 1 will be very good in the sense; can be contacted for promotions directly. For example:

Take (“education”=”secondary”) cluster

266.txt :-

"age"=41,"job"="entrepreneur","marital"="married","education"="primary", "default"="no", "balance"=236,"housing"="yes","loan"="no","contact"="unknown","day"=5,"month"="may","duration"=151, "campaign"=1,"pdays"=-1,"previous"=0, "poutcome"="unknown"

If thresholds set properly according to the attributes the cluster will get a score of 1(one for duration > threshold (assumed to be 60)); threshold value will depend on attribute types.

If thresholds set properly according to the attributes the cluster will get a score of 1(one for duration > threshold (assumed to be 60)); threshold value will depend on attribute types.

Decision Tree

Fig. 5 explains steps to create a decision tree. The ranked text files are converted to .csv file and the .csv file is given as the input to the decision tree. The rules are stored in the .txt file. This module tries to give the predictive insights to the marketers. By taking the ranks of clusters as classes create the decision tree so that it could be tested with the new transactions, so that the prediction and evaluation of the new data can be done. Certainty factor of a leaf in a decision tree is the ratio of the number of records of the majority class with the total records in the leaf.

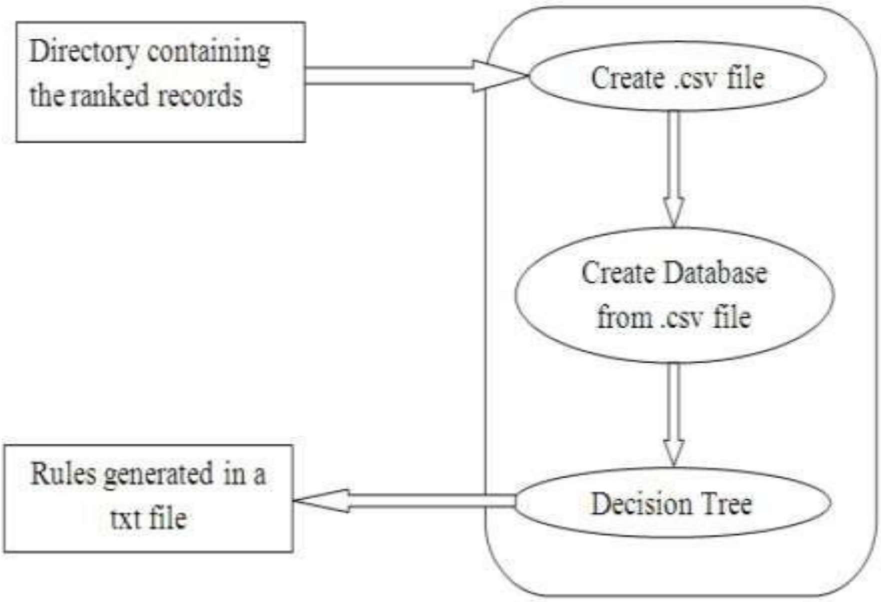

Fig. 5 Decision Tree creation

The proposed model will try to classify new data with classes as ranks (1, 2, 3, etc.). When a new record is given, it can be classified into the ranks (classes) which are a measure of marketing method (telephone/email) that can be applied to the new data and helps in churn management. Yin et al. [10] developed a greedy heuristic algorithm that uses a collection of decision trees instead of a single tree to produce the rules. The multiple Decision Trees are created with the attributes with equal information gain as root nodes. The decision trees which is having largest number of leaves having the ranks (1, 2, 3, etc.) and having the test results high. Comparing the classes/ranks with the bank client information such as annual balance could be applied for churn management. The actionable knowledge could be extracted using the method described in [10].

2.5. Binary Classification Algorithm

Using the decision tree rules which are in txt file, classify the new customer record. Here, the rule which is matching at first time is taken. Fig. 6 demonstrates the Customer segmentation in banking sector.

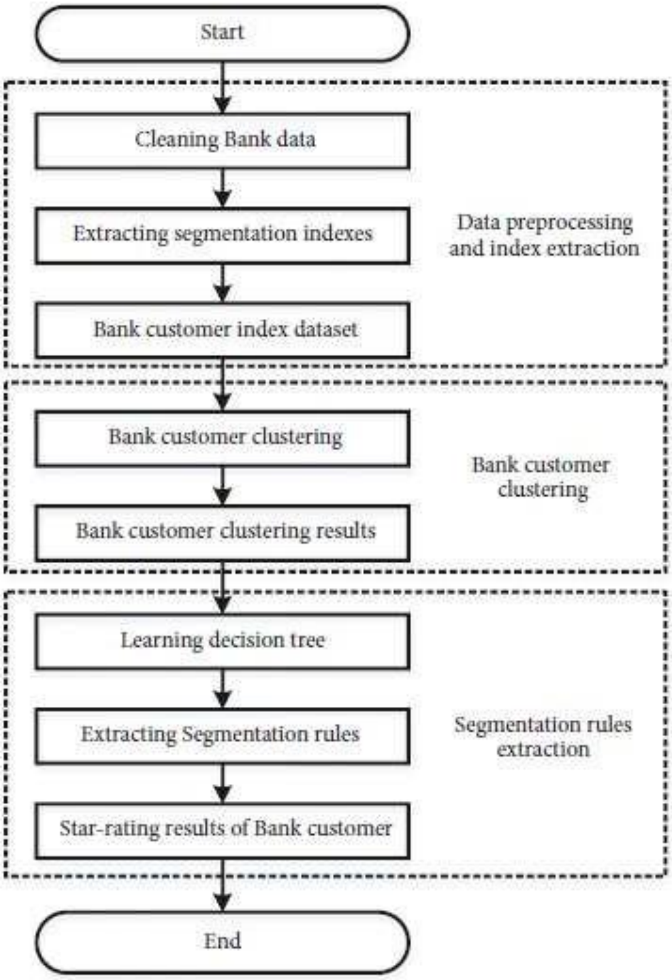

Fig. 6 Intelligent Learning based Customer Segmentation for E-Banking

Input:

BD – Bank customer index dataset
c – Total amount of clusters
S – Total amount of samples in the dataset

Output:

Bank customer details in cluster form

BEGIN:

for i = 1 to S do

Pick the samples from the Bank customer dataset BD to compute the values of clusters $[Ma1, Ma2, \ldots, Mak]^T$

Compute the values of centroid and generate the neighbour in the group of clusters $[Ma1, Ma2, \ldots, Mak]^T$

Reassign every bank customer details to another cluster group

Calculate the mean dissimilar values in the clustering for the current value

If the computed value < latest value then

Replace the latest value and form the suitable values from the

dataset

Repeat the entire process from the beginning of the procedure

Continue the clustering until no change is detected from the group

The computed output will be the result of the bank customer segmentation

The customer segmentation segregates the customers into the groups of similarity index. The index is measured using the behaviour of the customer for bank transaction related purposes. The banking data records with several kinds of data are kept in the automated system to frame the segmentation procedure. The bank customer data manipulation is computed using Eq. (1):

$$Index = Time_{specified} - Time_{availed} \quad (1)$$

where $Time_{specified}$ is used to find the specified time, $Time_{availed}$ is the time availed for the customer satisfaction.

$$Satisfactory_{index} = \sum_{i=1}^{n} Sample_i \quad (2)$$

The actual distance is computed to find the distance within the neighbouring objects is computed in Eq. (3)

$$distance(Object_j, Mean_j) = \sqrt{\sum_{k=1}^{n} (Object_{jk}, Mean_{jk})^2} \quad (3)$$

where $Object_j$ represents t he dimension value for the objects, $Mean_j$ is the mean value for evaluating the customer segmentation.

With the help of distance between the samples, the minimized value is measured for the n distances in Eq. (4).

$$distance(Object_j, Mean) = \min\{distance(Object_j, Mean_j), j \in (1,2,..,n)\} \quad (4)$$

To demonstrate the latest mean values of the clustering dissimilarity for the dataset for the clustering methodology is computed using Eq. (5)

$$Distance_{average} = \frac{1}{M} \sum_{j=1}^{M} distance(Object_j, Mean) \quad (5)$$

3. PERFORMANCE EVALUATION

The rules that are generated by the proposed system are formally evaluated in terms of technical measures as well as business success criteria in the business analysing phase. Choose whether the consequences of a given model appropriately address the underlying business targets. The model is evaluated using lift measure (\%correct prediction vs. \%correct data used) and ROC curve measure. Table 1 and Table 2 values are used to construct the Lift chart and ROC curve which are used to prove the effectiveness of our system. The dataset used is the Portuguese marketing campaign related with bank deposit subscription. Over training is the effect in data analysis, data mining of training too closely on limited available data and building models that do not generalize well to new unseen data. The marketing of products to select groups of consumers that is more likely than average to be interested in the offer. The proposed method is compared to the related methods of DMF [4], CRISP-DM [5], and CRM [6] with the parameters of Accuracy, Sensitivity and Specificity. The simulation results proved that the proposed method has performed well to the related methods.

Table 1. Lift Chart and ROC values

Actual Response	Predicted Response	TPR	FPR
0.04	0.96	0	4.1273665
0.04	0.96	0	4.5485944
0.05	0.95	0	5.0824223
0.06	0.94	0	6.767908
0.07	0.93	0	7.266158
0.1	0.9	0.002229456	11.564186
0.11	0.89	0.002219017	11.734161

Table 2 Lift Chart and ROC values using only first two modules

Actual Response	Predicted Response	TPR	FPR
0.96	0.04	1.0866667	4.113333
0.95	0.05	1.245	4.63
0.95	0.06	1.3667122	5.0668354
0.93	0.07	2.0142858	6.7485714
0.93	0.08	2.4425	7.24
0.91	0.1	5.4844446	11.528889
0.91	0.1	5.4844446	11.698481

The lift measure permits in recognizing how great the class separation is: the higher the better, with the perfect model having an estimation of 1.0. The model which provides a highquality AUC (Area Under Curve) value (maximum value=1.0), is the good model. The lift chart is plotted using the following:

X- axis =% actual response.

Y- axis = %predicted response.

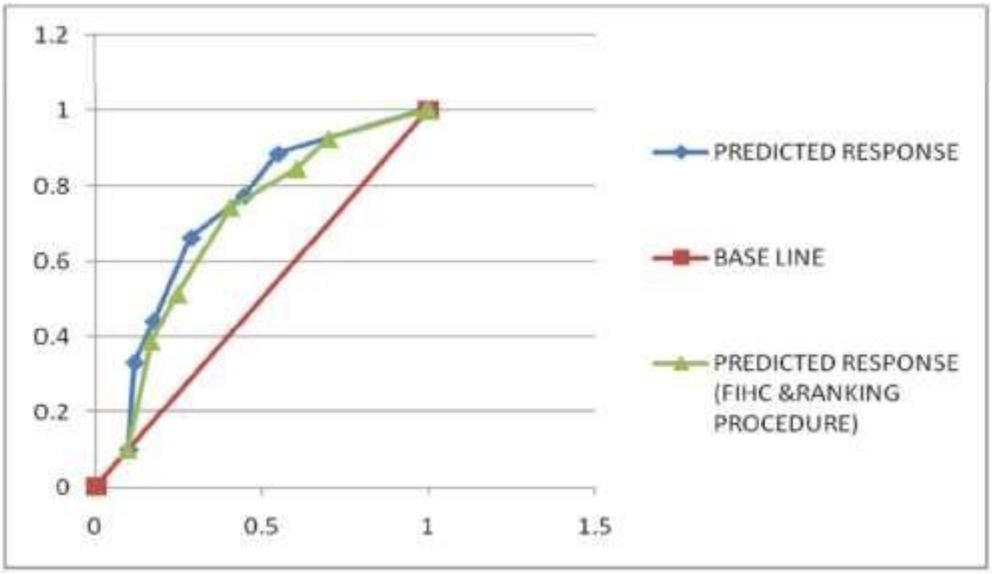

Fig. 6: Comparison of lift chart produced by using only first two modules and decision tree

The lift chart is used to compare the values obtained using the first three modules and the values obtained using the entire system. It is noticed that the lift chart produces nearly the same plot while comparing both of the curves. In Fig. 6, the green curve denotes the lift chart generated by the first three modules and the blue curve represents the response of the entire system. Fig. 7 shows the ROC curve. The green curve gives the FPR of the entire system. The blue curve is the FPR produced using the first three modules. The comparison of the results (LIFT chart and ROC curve) obtained from the rules generated by two modules FIHC and Ranking and the rules produced by the decision tree are done. They provide almost the same results in the case of Lift chart. The AUC (Area Under Curve) produced by the decision tree is higher than the AUC produced by the modules (FIHC and Ranking). Thus, the class discrimination property is good if decision tree is used.

The ROC curve is plotted using the following,

X- axis = FPR(False Positive Rate).

Y- axis = TPR(True Positive Rate).

The customer satisfaction with hundred percentage true positive rate and the related false positives are demonstrated in Fig. 8. The coordination points from 0 to 1 have the value of ROC with corresponding diagonal values. The classification is demonstrated by a point close to the left corner of upper limit indicates the random value of area under curve.

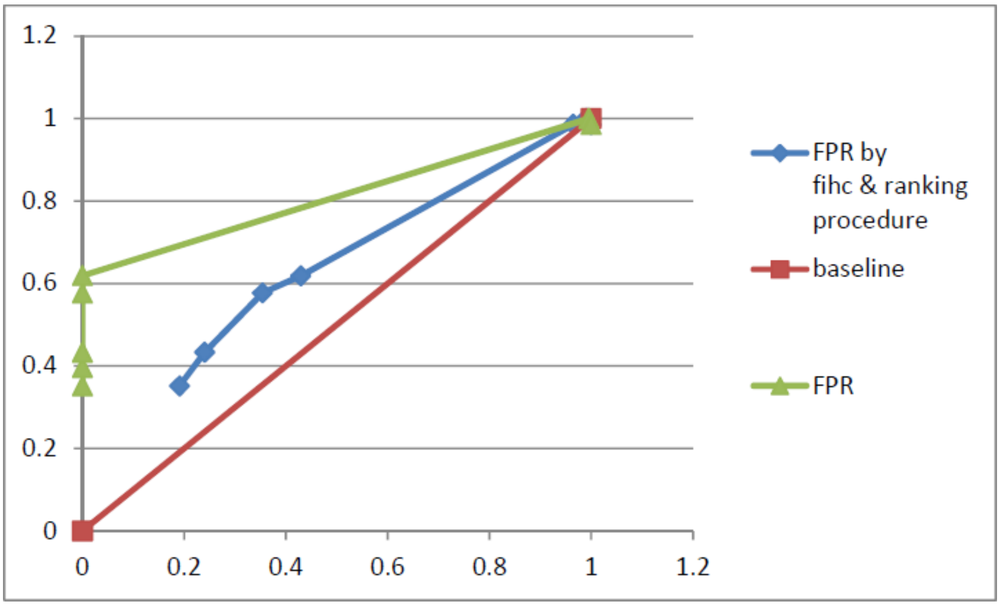

Fig. 7: Comparison of ROC curve produced by using only first three modules and decision tree

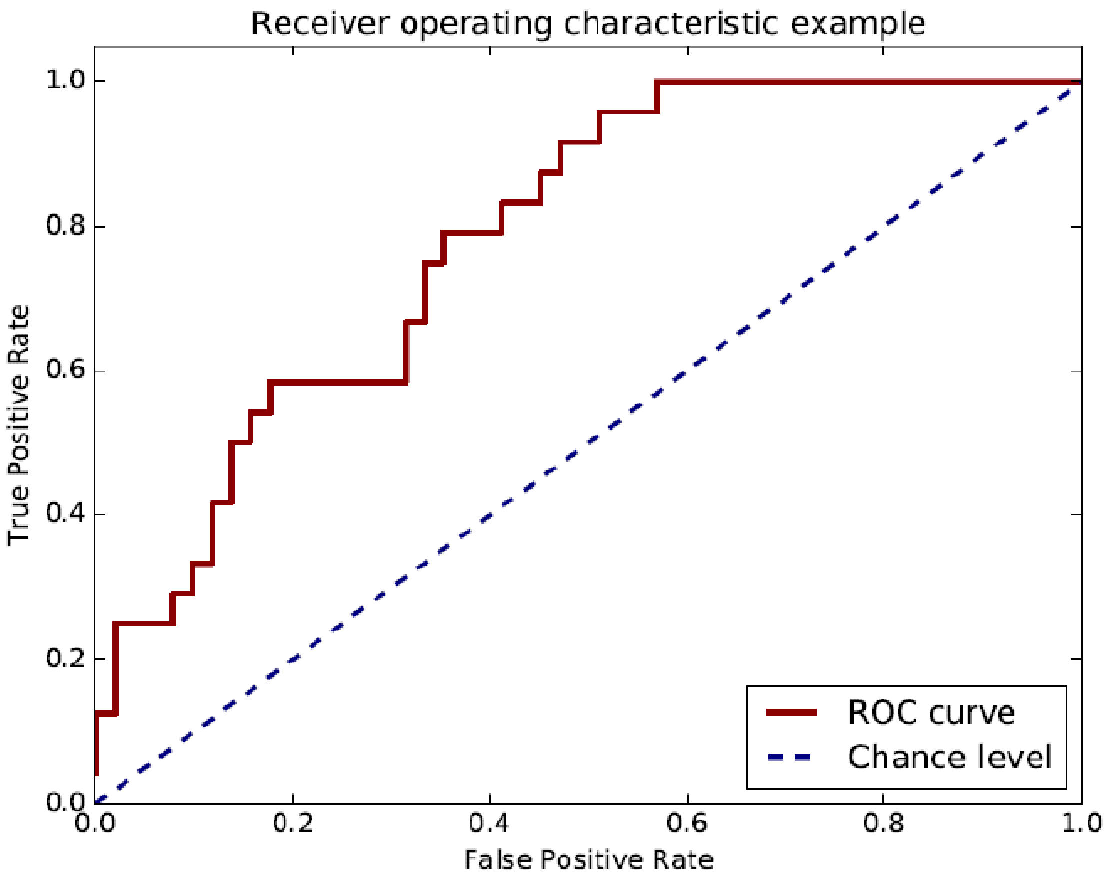

Fig. 8: ROC curve for False Positive rate and true positive rate

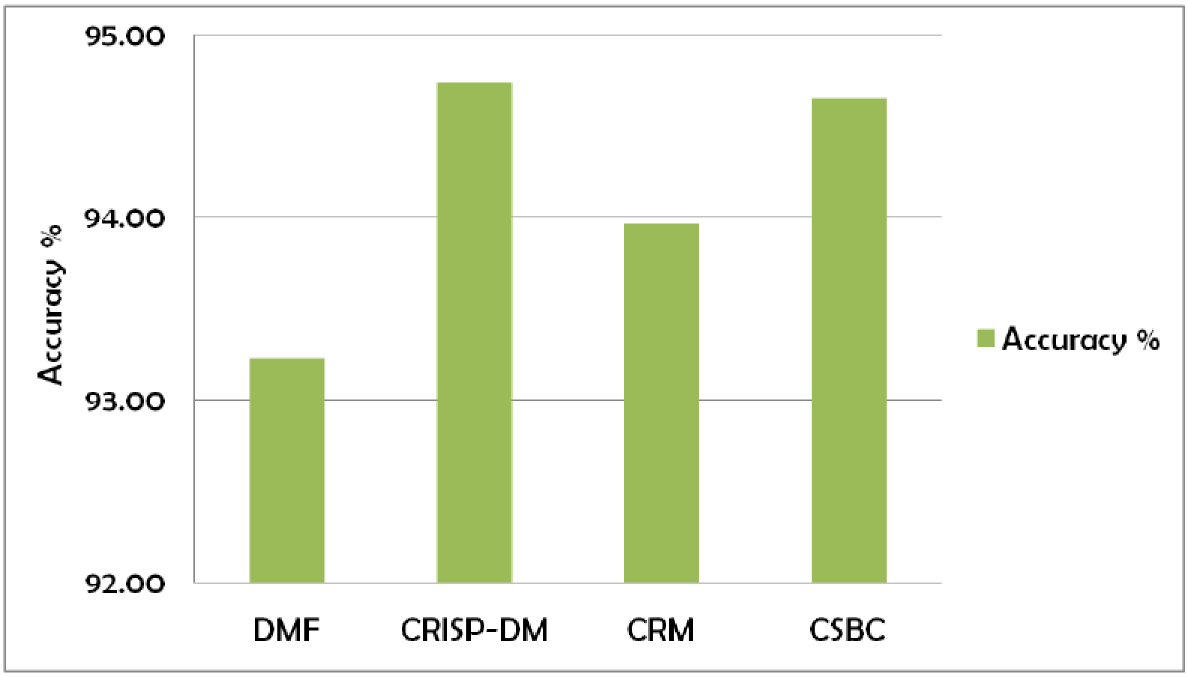

Fig. 9: Accuracy

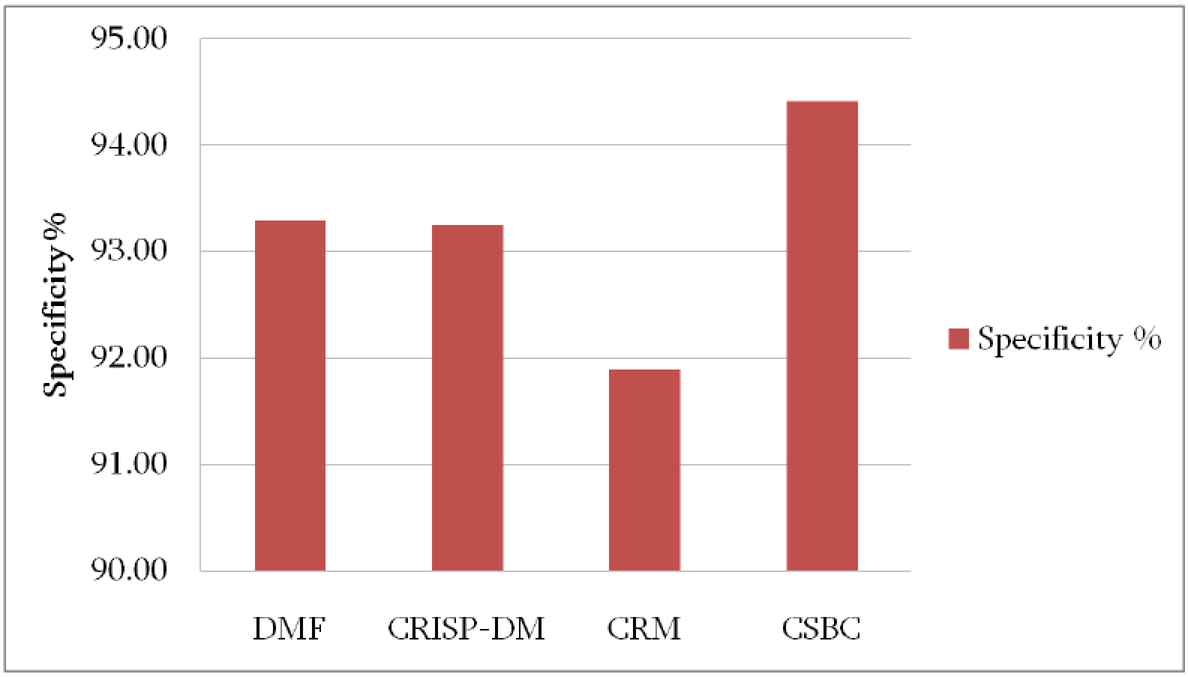

Fig. 10: Specificity

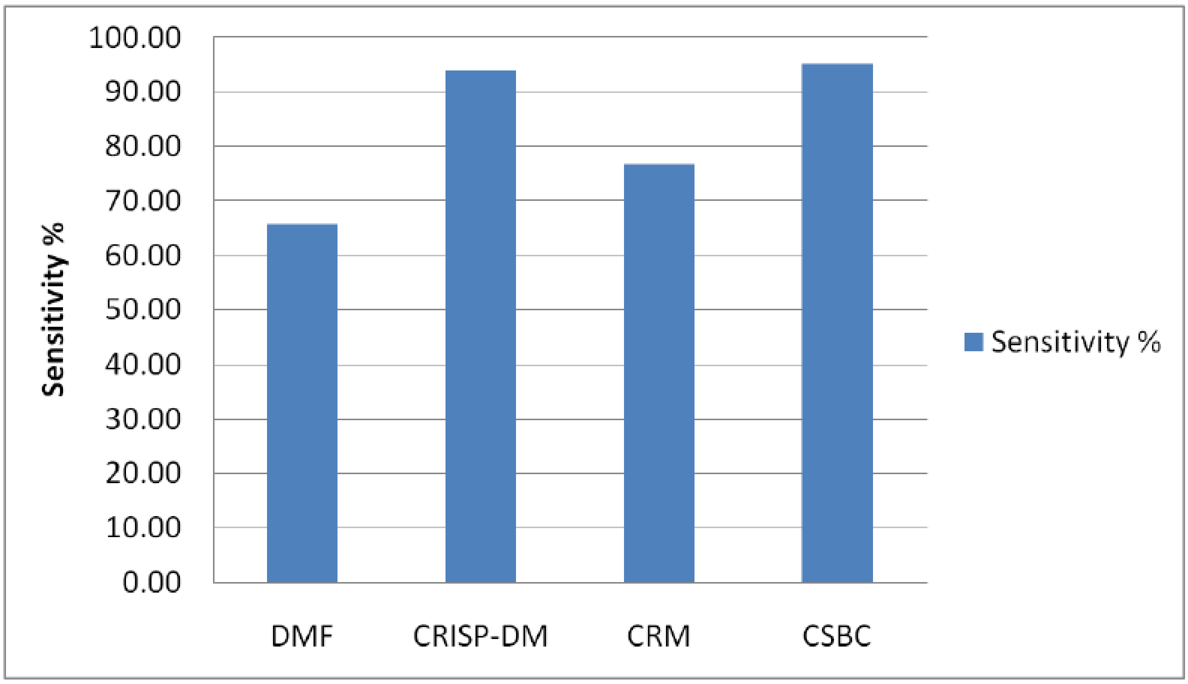

Fig. 11: Sensitivity

The accuracy, Specificity and Sensitivity values are computed using the Eq. (6), Eq. (7) and Eq. (8) respectively.

$$Accuracy = Total\ amount\ of\ correct\ elements\ /\ Total\ amount\ of\ all\ elements \quad (6)$$

$$Specificity = Total\ amount\ of\ true\ negative\ elements\ /\ Total\ amount\ of\ all\ negative\ elements \quad (7)$$

$$Sensitivity = Total\ amount\ of\ true\ positive\ elements\ /\ Total\ amount\ of\ positive\ elements \quad (8)$$

Figures. 9-11 demonstrate that the proposed method has the improved performance compared to the related methods in the specific parameters of accuracy, Specificity and Sensitivity values. 4.

4. CONCLUSION

In this paper, we proposed a novel algorithm called CSBC for Direct Marketing using Frequent Item-set Hierarchical Clustering. Frequent-Itemset Hierarchical clustering and ranking method are incorporated in CSBC to grade/segment the customers into more than two classes. A classifier produced rules would be used to predict the new customer data. Completely automated predictive insights (used for direct marketing) are given to the marketing personnel in banks. It is noticed that our proposed system is not overly dependent on the dataset used. The overtraining problem can be solved if we use the dataset which includes more attributes than the attributes considered in this proposed work.

The customer segmentation is done by two modules Frequent-Item-set Hierarchical Clustering (FIHC) and Ranking based on the Intelligent learning based Banking system. These two modules are used to segment the data into more than two classes/ranks, thus handling the problem of imbalanced class distribution is simplified. The assigned ranks of the customers give an indication to churn management. The customer with higher rank (assigned by proposed system) is withdrawing a large amount, is a clear indication of churning. It would be detected by marketing personnel and can be used for planning strategies for churn management. The rules generated by the decision tree could be used for discriminating any new customer data. Also, ranking can be done on the same data. Thus, the marketer will get the extracted knowledge in a useful form.

The proposed model was verified on benchmark datasets and has better performance than the others. It could be applied to a real bank data and the predictive insights can be given to the marketing personnel or campaign management software. The proposed method can be evaluated based on the response measure of the bank's customers. This leaves in our further studies.

References

[1]. Business analytics based on financial Time series,by by KR Varshney, IEEE, 2011
[2]. Analysing success drivers of E-Business companies, by S Albers ,IEEE, 2007
[3]. Data mining for Direct Marketing: problems and solutions by by CX Ling.
[4]. A Direct marketing framework to facilitate data mining usage for Marketers by Adel Flici
[5]. Using data mining for bank direct marketing: an application of the CRISP-DM methodology, by Paulo Cortez , Sérgio Moro and Raul M. S. Laureano, IEEE
[6]. Data Mining Techniques in CRM: Inside Customer Segmentation by KonstantinosTsiptsis and AntoniosChorianopoulos
[7]. Fuzzy Modeling of Client Preference from Large Data Sets: An Application to Target Selection in Direct Marketing, by M Setnes, IEEE
[8]. Mining Fuzzy Association Rules in a Bank-Account Database by Wai-Ho Au and Keith C. C. Chan, IEEE
[9]. Hierarchical Document Clustering Using Frequent Itemsets by Benjamin C.M. Fung, Ke Wang *and* Martin Ester
[10]. Extracting Actionable Knowledge from Decision Trees, Qiang Yang, Senior Member, IEEE
[11]. Customer Segmentation of Bank based on Data Mining – Security Value based Heuristic Approach as a Replacement to K-means Segmentation by Shashidhar HV and SubramanianVaradarajan
[12]. An Intrusion-Detection Model Based on Fuzzy Class-Association-Rule Mining Using Genetic Network Programming, by S Mabu ,IEEE,2011.
[13]. Page C. and Luding Y., 2003. "Bank manager's direct marketing dilemmas – customer's attitudes and purchase intention". *International Journal of Bank Marketing* 21, No.3, 147–163.
[14]. Chen M., Chiu A. & Chang H. 2005, "Mining changes in customer behaviour in retail marketing", Expert Systems with Applications, vol. 28, no. 4, pp. 773-781.
[15]. Wang H. & Hong W. 2006, "Managing customer profitability in a competitive market by continuous data mining", Industrial Marketing Management, vol. 35,no. 6, pp.715-723.
[16]. R. C. Dubes and A. K. Jain.Algorithms for ClusteringData. Prentice Hall College Div, Englewood Cliffs, NJ,March 1998.
[17]. Kim Y. 2006, "Toward a successful CRM: variable selection, sampling, and ensemble", Decision Support Systems, vol. 41, no. 2, pp. 542-553.
[18]. Bose I. & Chen X. 2009, "Quantitative models for direct marketing: A review from systems perspective", European Journal of Operational Research, vol. 195, no.1, pp. 1-16
[19]. http://archive.ics.uci.edu/ml/datasets/Bank+Marketing
[20]. Muhammet Sinan Başarslan & İrem Düzdar Argun, "Classification Of a bank data set on various data mining platforms", 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT).