Volume: 08 Issue: 04 Pages: 3161-3168 (2017) ISSN: 0975-0290

Optimized Feature Extraction and Actionable Knowledge Discovery for Customer Relationship Management

Dr P.Senthil Vadivu
Assc Professor & Head, Dept of Computer Applications(UG), Hindusthan College of Arts and Science, Coimbatore-641028
Email:sowjupavan@gmail.com

ABSTRACT

In today's dynamic marketplace, telecommunication organizations, both private and public, are increasingly leaving antiquated marketing philosophies and strategies to the adoption of more customer-driven initiatives that seek to understand, attract, retain and build intimate long term relationship with profitable customers . This paradigm shift has undauntedly led to the growing interest in Customer Relationship Management (CRM) initiatives that aim at ensuring customer identification and interactions. The urgent market requirement is to identify automated methods that can assist businesses in the complex task of predicting customer churning.

The immediate requirement of the market is to have systems that can perform accurate

(i) identification of loyal customers (so that companies can offer more services to retain them)

(ii) **prediction of churners to ensure that only the customers who are planning to switch their service providers are being targeted for retention**

Keywords: CRM, CLA-AKD, LOF, ACO, UDA

Date of Submission: Feb 09, 2017 Date of Acceptance: Feb 17, 2017

1. INTRODUCTION

To design and develop such systems, this work combines data mining algorithms and decision making by formulating the decision making problems directly on top of the data mining results in a post processing step. These systems can discover loyal customers, churning customers and can produce a set of actions that can be applied to transform churning customers to loyal customers. These systems can effectively produce intelligent CRM for telecommunication industries. This introduction is to present the various topics related to the topic of research.

2. CUSTOMER RELATIONSHIP MANAGEMENT (CRM)

The main goal of CRM is to aid organizations in better understanding of each customer's value to the company, while improving the efficiency and effectiveness of communication. . It is one of the most helpful analysis task used by companies to maintain loyal customers.

Figure 2 : Work Flow of Customer Relationship Management (CRM)
(Source: http://www.zoho.com/crm/how-crm-works.html)

There are three aspects of CRM that can be used. They are operational CRM, Analytical CRM and Collaborative CRM. Operational CRM is concerned with analyzing and improving operations like (i) marketing automation, (ii) sales force automation and (iii) customer service and self-service. Analytical CRM, uses data warehousing and data mining techniques, to improve company-customer relationships. Collaborative CRM, as the name suggests, integrates operation CRM with analytical CRM

3. CRM AND TELECOM INDUSTRY

Telecom is a huge and varied bastion of technologies, companies, services and politics that is truly global in nature. Telecom industries customer service is key to sales and loyalty and customer service has become the differentiator between its competitors.

Telecom industries use data mining techniques with the aim of constructing customer profiles, which predict the characteristics of customers of certain classes. Examples of these classes are:

- What kind of customers (described by their attributes such as age, income, etc.) are likely attritors (who will go to competitors), and
- what kind are loyal customers?

4. CUSTOMER CHURNING

Customer churns also known as customer attrition, customer turnover, or customer defection, is the loss of clients or customers (http://en.wikipedia.org/wiki/Customer_attrition).

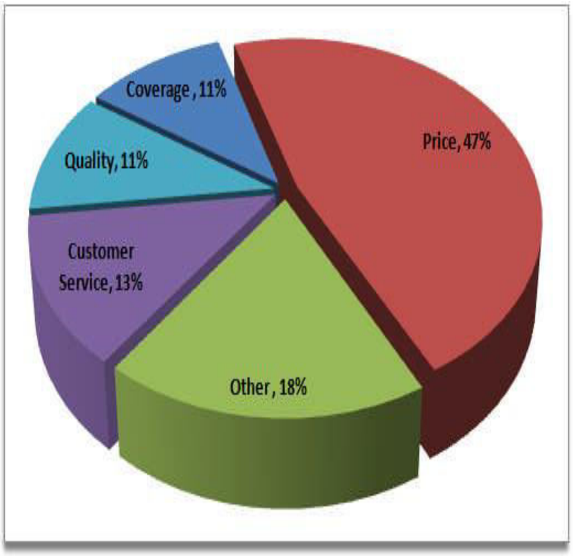

Figure 4 : Causes for Churn(Source : Chu *et al.*, 20007)

5. ACTIONABLE KNOWLEDGE

Actionable knowledge discovery is critical in promoting and releasing the productivity of data mining and knowledge discovery for smart business operations and decision making. In general, telecom industries use algorithms and tools which focus only on discovering patterns that identify loyalty and churners. In most of the companies, the processes of identifying loyal customers and predicting churners are performed by human experts. These human experts use the visualization results and interestingness ranking to discover knowledge manually.

5.1 METHODOLOGY

The main challenge in the work is to find techniques that bridge the two fields, data mining and telecommunication, for an efficient and successful knowledge discovery. . The eventual goal of this data mining effort is to identify factors that will improve the quality and cost effectiveness of customer care (Yu and Ying, 2009).

Integration of decision support with computer-based CRM can reduce errors, enhance customer care, decrease customer churning and improve customer loyalty (Wu *et al.,* 2002).The main goal of using data mining techniques in CLA-AKD is to identify those customers who share common attributes, which can be interpreted to group them as either loyal or churners.

In this work, two types of operations are performed on customer database. They are,

(i) Customer Loyalty Assessment: To recognize and classify customers with multivariate attributes as loyal customers (non-churners) and churners.

(ii) Actionable Knowledge Discovery – To predict a set of actions that can be used to convert churners to valuable loyal customers.

Both operations can be efficiently performed using cluster analysis and classification.

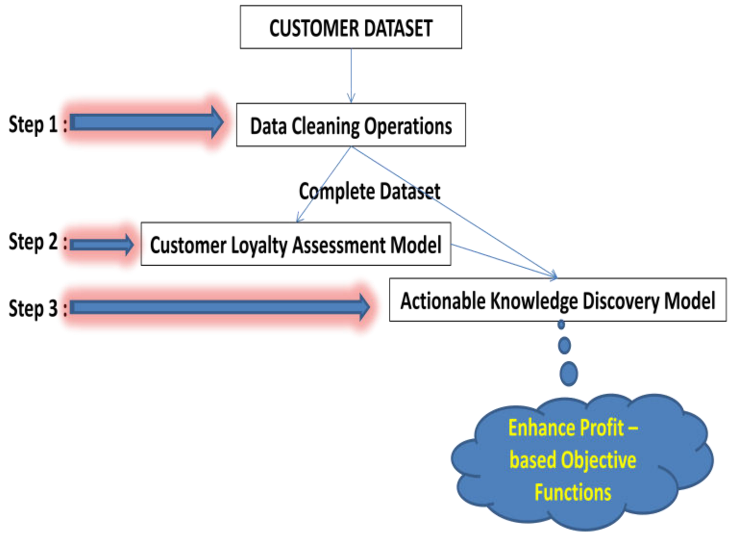

Figure 5.1: Architecture of CLA-AKD Model

The proposed CLA-AKD model performs customer loyalty assessment and actionable knowledge discovery in three steps, namely,

(i) Preprocessing
(ii) Customer Loyalty Assessment
(iii) Actionable Knowledge Discovery.

5.2 PHASE I: DATA CLEANING OPERATIONS

. Data cleaning routines attempt to smooth out noise while identifying outliers, fill-in missing values and correct inconsistencies in data. The amount of time and cost spend on preprocessing has a direct impact on the accuracy (Figure 1.6.1).

Volume: 08 Issue: 04 Pages: 3161-3168 (2017) ISSN: 0975-0290

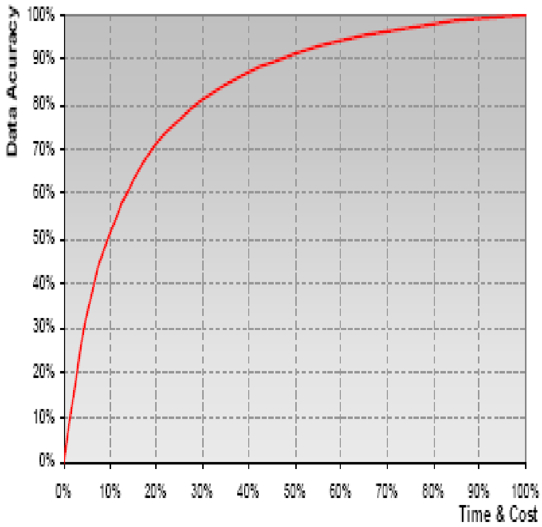

Figure 1.6.1: Preprocessing and Accuracy Tradeoff

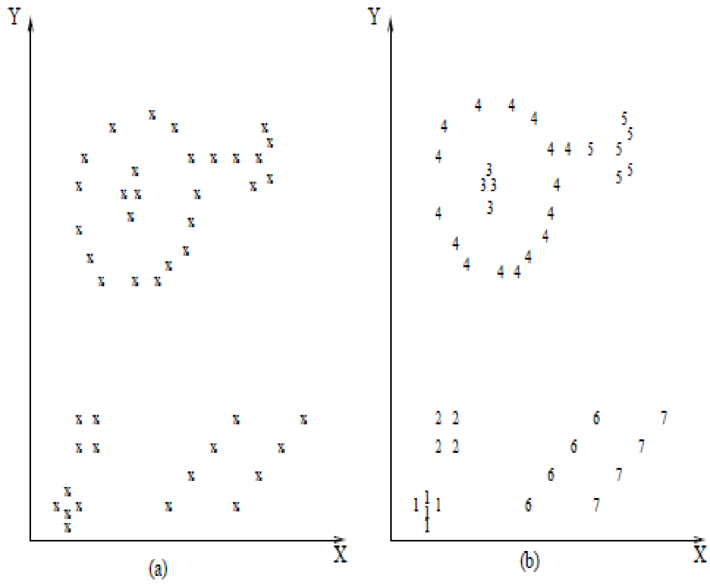

Figure 1.6.2.1: Data Clustering

5.2.1 OUTLIER DETECTION AND REMOVAL

Outlier detection and removal is a process that is considered very challenging and important by all data mining applications. The first task of the first phase concentrates on detecting noise or outliers in the telecom dataset and removes them to obtain a clean dataset. For this purpose, an enhanced density based outlier detection algorithm using Local Outlier Factor (LOF) is proposed.

5.3 PHASE II: CUSTOMER LOYALTY ASSESSMENT MODELS

Frameworks for predicting customer's future behavior has captured the attention of several researchers and academicians in recent years, as early detection of future customer churns is one of the most wanted CRM strategies. Future of company depends on the stable income provided by loyal customers. However, attracting new customers is a difficult and costly task involved with market research, advertisement and promotional expenses.

5.3.1 Step 1: Clustering

Typical pattern clustering activity involves the following steps (Jain and Dubes 1988):

- Pattern representation (optionally including feature extraction and/or selection),
- Definition of a pattern proximity measure appropriate to the data domain,
- Clustering or grouping,
- Data abstraction (if needed), and
- Assessment of output (if needed).

5.3.2 Step 2: Classification

Classification is the process of finding a set of models (or functions) that describe and distinguish data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown. The derived model is based on the analysis of a set of training data.

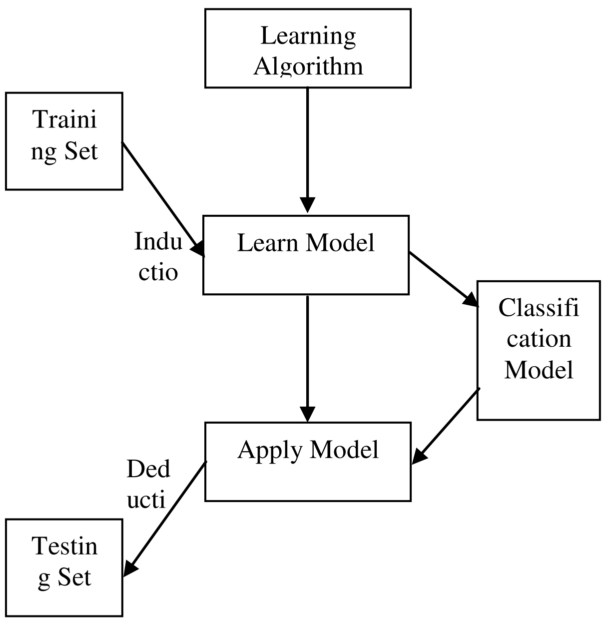

Figure 1.6.2.2: General Process of Classification

5.4 PHASE III: ACTIONABLE KNOWLEDGE DISCOVERY MODELS

The proposed AKD model consists of four steps, The first step implements two dimensionality reduction algorithms, namely, ACO (Ant Colony Optimization) or UDA (Uncorrelated Discriminate Analysis) to reduce the dataset size. The second step builds customer profile from the training data, using two classifiers, a decision tree

Volume: 08 Issue: 04 Pages: 3161-3168 (2017) ISSN: 0975-0290

learning algorithm and Bayesian Network classifier. These two algorithms were selected because of their efficiency provided during churn prediction. The third step searches for optimal actions for each customer. This is a key component of the system's proactive solution. The fourth step produces reports for domain experts to review the actions and selectively deploy the actions.

6 CUSTOMER LOYALTY ASSESSMENT MODEL

In telecommunication industry, 'Customer Churn' is used to denote customer movement from one service provider to another. Customer Churn Management (CCM) is the study of algorithms that process customer data and find methods to identify loyal customers (Berson *et al.*, 2000) and take actions to secure these customers to the same company (Kentrias, 2001). CCM implements techniques that have the ability to forecast customer decision to shift from one service provider to another. These techniques use segmentation operations to segment its customers in terms of their profitability and then apply retention management that takes appropriate actions to convert churners to loyal customer segment.

6.1 CLA MODEL

The steps involved in the proposed customer loyalty assessment model are presented in Figure 1.7.1. The proposed customer loyalty assessment model presents a two-step procedure. In the first step, a SOM-(KMeans + DBSCAN) method is proposed. This method combines SOM (Self-Organizing Map) with a hybrid clustering algorithm that combines the partition-based clustering algorithm (KMeans) with a density based clustering algorithm (DBSCAN). This algorithm analyzes the customer database and segments the customers into two distinct groups of customers with similar characteristics and behaviors. The two groups of customers are loyal (non-churners) customers and churners.

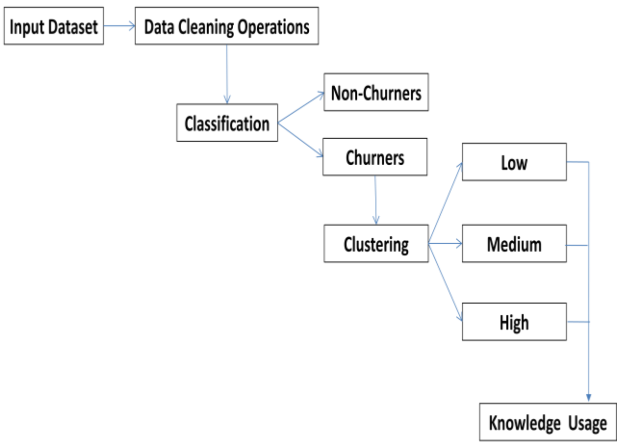

Figure 6.1 : Customer Loyalty Prediction Model

6.2 SELF-ORGANIZING MAPS

Self-Organizing Maps (SOM) are widely used unsupervised neural network that has found several applications that use clustering (Han and Kamber, 2000). Example application includes industry applications (Kaski et *al.*, 1998) and market segmentation (Oja *et al.*, 2002). It has the advantage of projecting the relationships between high-dimensional data onto a two-dimensional display, where similar input records are located close to each other. By adopting an unsupervised learning paradigm, the SOM conducts clustering tasks in a completely data-driven way (Kohonen *et al.*, 1996), that is, it works with little a priori information or assumptions concerning the input data. In addition, the SOM's capability of preserving the topological relationships of the input data and its excellent visualization features motivated this research work to adopt SOM in the design of CLA-AKD.

6.3 CLUSTERING ALGORITHM

The proposed CLA model uses two algorithms, namely, KMeans and DBSCAN to form a hybrid clustering model. In the study, the DBSCAN is enhanced to automatically estimate its parameters. This section presents the working of KMeans algorithm, Enhanced DBSCAN algorithm and proposed Hybrid algorithm.

6.4. KMeans Algorithm

K Means clustering generates a specific number of disjoint, flat (non-hierarchical) clusters. It is well suited for generating globular clusters. The K Means method is an unsupervised, non-deterministic and iterative method with the following properties.

- There are always K clusters.
- The clusters are non-hierarchical and they do not overlap.
- Every member of a cluster is closer to its cluster than any other cluster because closeness does not always involve the 'center' of clusters.

6.5 DBSCAN Algorithm

The second algorithm of the hybrid clustering algorithm is the DBSCAN (Density Based Spatial Clustering of Applications with Noise) algorithm. Density-based clustering, introduced by Ester *et al.* (1996), locates regions of high density that are separated from one another by regions of low density. DBSCAN is a simple and effective density-based clustering algorithm which has been proved to work effectively in the field of clustering.

6.5.1 HYBRID KMEANS-DBSCAN ALGORITHM

The KMeans clustering algorithm is simple and fast, buts its performance degrades in the presence of noise. Alternatively, the DBSCAN algorithm performance is more than satisfactory in the presence of noise but has time complexity, and it is slow. A hybrid model is proposed to take advantage of both KMeans and DBSCAN algorithms to improve the speed and make it robust against noise. For this purpose this study proposes a method to combine KMeans and DBSCAN algorithm. This work enhances the work of El-Sonbaty *et al.* (2004) and consists of the following three steps.

(i) Use a preprocessing step that reduces the dimensionality of the dataset and partition into obtain k parts

(ii) Apply DBSCAN to each partition
(iii) Merge dense regions until the number of clusters is K

6.6 CLASSIFICATION ALGORITHMS

This section presents the three classification algorithms, namely, SVM, BPNN and Decision Trees, that are used to classify churners as low, medium and high retention customers.

6.6.1 Support Vector Machine

A key feature of the proposed ensemble model is the integration of Support Vector Machines with clustering. Among the various classifiers available, SVMs are well known for their generalization performance and ability to handle high dimensional data and a brief general explanation of the same is provided in this section.

6.6.2 Back Propagation Neural Networks

A Back Propagation Neural Network (BPNN) is an artificial neural network where connections between the units do not form a directed cycle. They are the first simplest type of artificial neural network devised where, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network. The Back Propagation Algorithm is a common way of teaching artificial neural networks how to perform a given task. It requires training which has knowledge or can calculate the desired output for any given input. Backpropagation algorithm learns the weights for a multilayer network, given a network with a fixed set of units and interconnections. It employs gradient descendent rule to attempt to minimize the squared error between the network output values and the target values for these outputs.

6.6.3 Decision Tree Classifier

Data classification, one of the important task of data mining, is the process of finding the common properties among a set of objects in a database and classifies them into different classes (Han *et al.*, 1993). One of the well-accepted classification methods is the induction of decision trees. A decision tree is a flow-chart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test and the leaf nodes represent classes or class distributions. A typical decision tree consists of three types of nodes :

a. A root node
b. Internal nodes
c. Leaf or terminal nodes

A root node is a node that has no. incoming edges and zero or more outgoing edges. Each of the internal nodes has exactly one incoming edge and two or more outgoing edges. Nodes having exactly one incoming edge with no outgoing edges are termed as leaf or terminal nodes. An example decision tree is shown in Figure 1.7.5.3

In a decision tree, each leaf node is assigned a class label. The non-terminal nodes, which include the root and other internal nodes, contain attribute test conditions to separate records that have different characteristics. Classifying a test record is straightforward once a decision tree has been constructed. Starting from the root node, the test condition is applied to the record and based on the output, the appropriate branch is followed. This will lead to either another node, for which a new test condition can be applied, or to a leaf node. The class label associated with the leaf node is then assigned to the record.

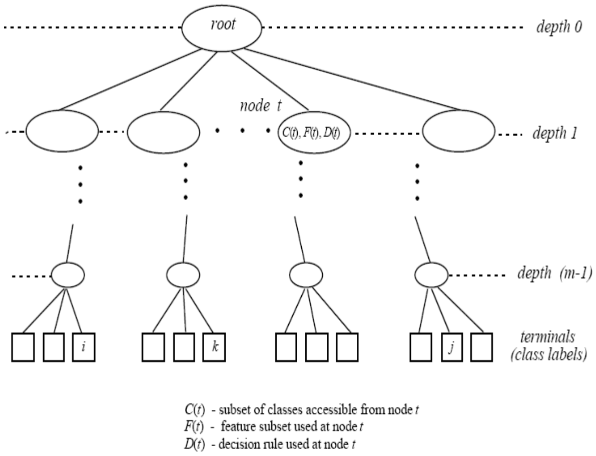

Fig. 6.6.3: Example of a general decision tree

7. RESULTS AND DISCUSSION

Churn represents the loss of an existing customer to a competitor. It is a prevalent problem faced by service provider of a subscription service or recurring purchasable in the telecommunication sector. It is well-known fact that the cost of customer acquisition and win-back is very high, when compared to retaining existing customer. Thus, customer loyalty assessment for loyal customer and churn customer identification and prediction are very important in telecommunication industry. Another problem faced after identifying churners is the decision of what actions to take to convert them to loyal customers. This work has proposed enhanced techniques to perform customer loyalty assessment for identifying loyal customers and churners and also to perform actionable knowledge discovery that presents a set of actions that can be used to improve customer loyalty.

7.1 EVALUATION OF OUTLIER DETECTION

The results show that the enhanced LOF algorithm is efficient in detecting outliers both with respect to outlier detection rate and speed. The outlier detection rate decreased (marginally) with the amount of outliers present. The outlier detection rate is consistent with the different level of randomness proving that the outlier removal or data cleaning process is efficient. On average, 1.89% efficiency gain was envisaged while using ELOF when compared to LOF with respect to detection rate.

Volume: 08 Issue: 04 Pages: 3161-3168 (2017) ISSN: 0975-0290

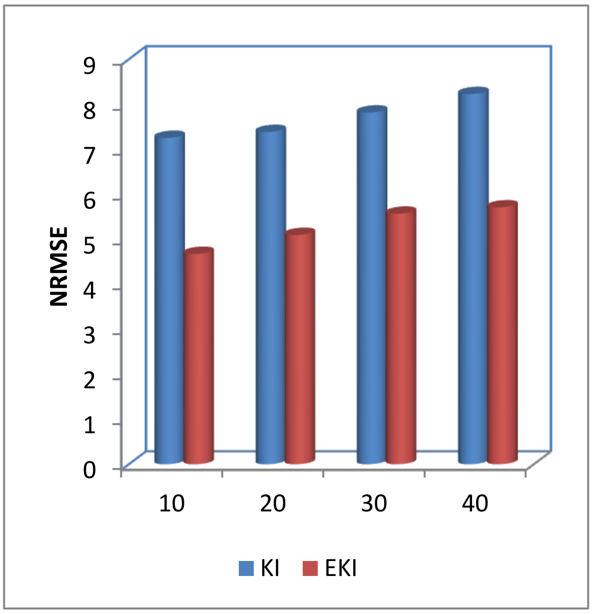

Figure 7.1: NRMSE of Missing Value Handling Algorithms

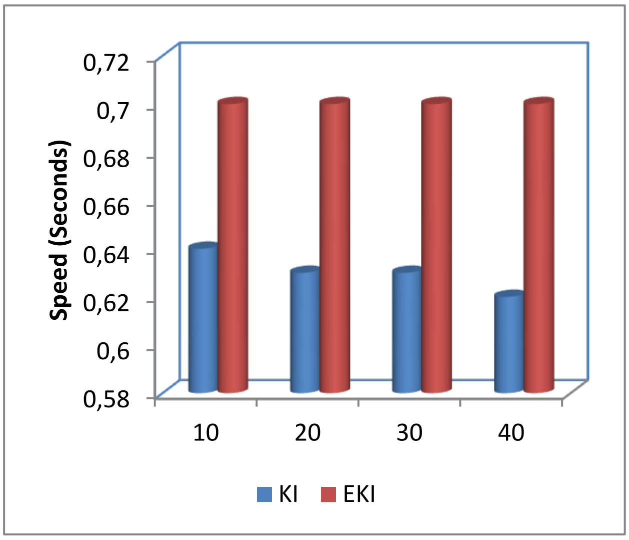

Figure7.2: Speed of Missing Value Handling Algorithms

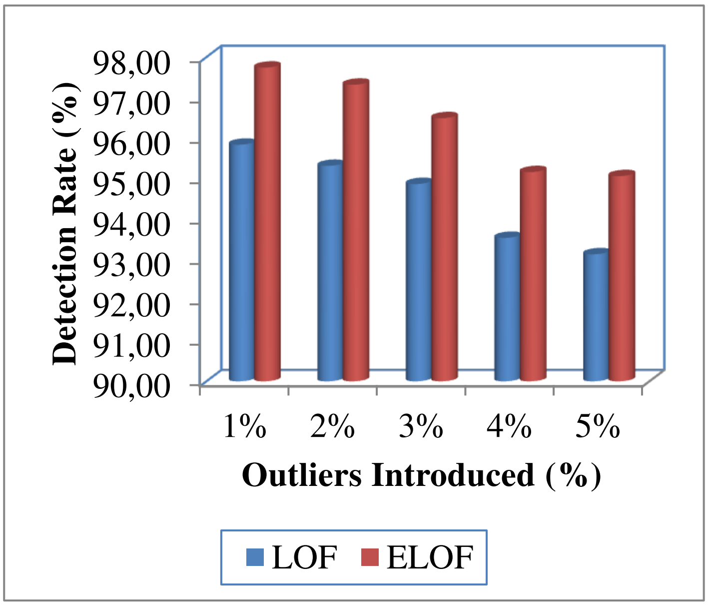

Figure 7.3 : Outlier Detection Rate

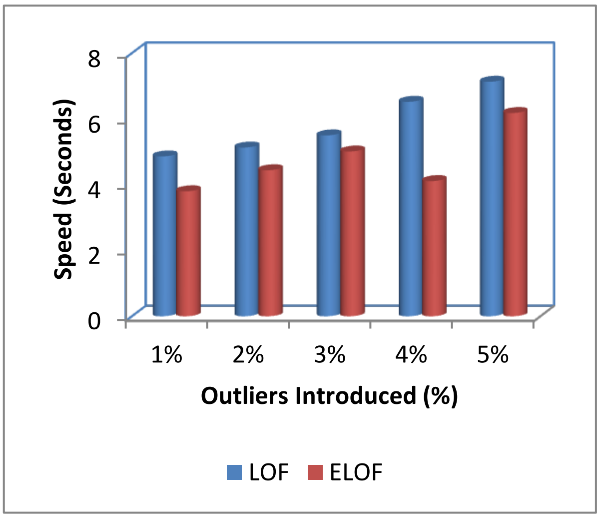

Figure 7.4: Speed (Seconds) of Outlier Detection

8. EVALUATION OF ACTIONABLE KNOWLEDGE DISCOVERY MODELS

The bounded segmentation problem for actionable knowledge discovery along with the proposed AKD systems was applied to calculate the pre-specified k action sets with maximum profit. Figure 8 and 8.1 shows the net profit obtained by the existing and proposed AKD models.

Volume: 08 Issue: 04 Pages: 3161-3168 (2017) ISSN: 0975-0290

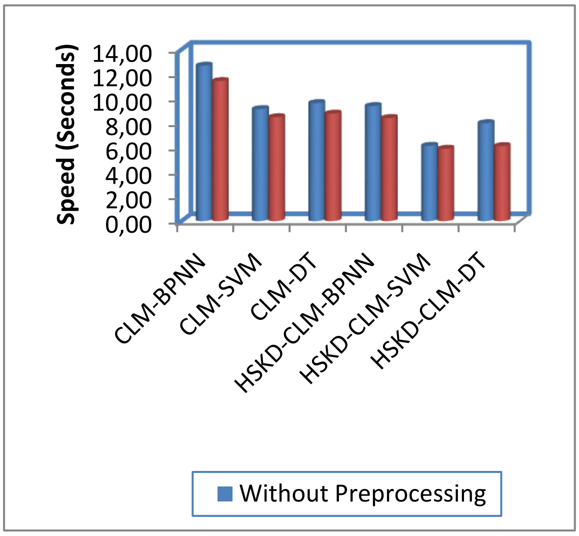

Figure 8 : Speed (Seconds) of Loyalty Assessment Models

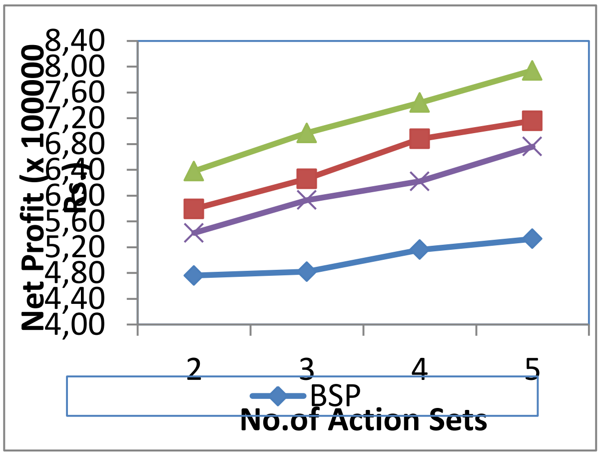

Figure8.1: Net Profit Obtained by Actionable Knowledge Discovery Systems

9 CONCLUSIONS

The performance evaluation stage was conducted in four stages, each analyzing the performance of the missing value handling algorithm, outlier detection algorithm, customer loyalty assessment model and actionable knowledge discovery system. A telecom customer dataset from IDEA customer care was used during experiments. The missing value handling algorithm was evaluated using two performance metrics, namely, Normalized Root Mean Square and speed. Similarly, the outlier detection algorithm was evaluated using outlier detection rate and speed. The customer loyalty assessment procedure used accuracy, error rate and speed to analyze the algorithms. The actionable knowledge discovery system analysis was based on net profit achieved and speed of discovery.

The various experiments conducted proved that the proposed algorithms and the proposed CRM system for customer loyalty assessment and actionable knowledge discovery are efficient. Experimental results showed that the system is effective in terms of analysis accuracy and speed in identifying common customer behaviour patterns and future churn prediction. The developed system has promising value in the current constantly changing telecommunication industry and can be used as effectively by companies to improve customer relationship and improve business opportunities.

10 FUTURE RESEARCH DIRECTIONS

The following can be considered to improve the proposed customer loyalty assessment model and actionable knowledge discovery system.

- The proposed models can be further enhanced, if the processes can be parallelizing. This is feasible, by identifying operations that are independent to each other and propose a parallel architecture to improve the performance.
- Amount of memory used loyalty assessment and action discovery is another area which can be analyzed in future.
- Classification process can be improved by using advanced techniques like ensemble clustering or ensemble classification.

REFERENCES

1. Berson, A., Smith, S. and Thearling, K. (2000, 1999) Building data mining applications for CRM, New York, McGraw-Hill.
2. Chu, B., Tsai, M. and Ho, C. (2007) Toward a hybrid data mining model for customer retention, Knowledge-Based Systems, Vol. 20, No. 8, Pp. 703-718.
3. Ester, M., Kriegel, H-P., Sander, J. and Xu, X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise, Proceedings of the 2nd ACM SIGKDD, Portland, Oregon, Pp. 226-231
4. El-Sonbaty, Y., Ismail, M.A. and arouk, M. (2004) An Efficient Density Based Clustering Algorithm for Large Databases, 16th IEEE International - Conference on Tools with Artificial Intelligence, Pp. 673-677.
5. Han, J. and Kamber, M. (2000) Data Mining: Concepts and Techniques, The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann.
6. Han, J., Cai, Y., Cercone, N. (1993) Data driven discovery of quantitative rules in relational databases, IEEE Trans. Knowledge and Data Engineering, Vol. 5, Pp. 29-40.

Volume: 08 Issue: 04 Pages: 3161-3168 (2017) ISSN: 0975-0290

7. http://en.wikipedia.org/wiki/Customer_attrition, Last Accessed during March, 2014.

8. Jain, A.K. and Dubes, R.C. (1988) Algorithms for Clustering Data, Prentice-Hall.

9. Kaski, S., Kangas, J. and Kohonen, T. (1998) Bibliography of Self- Organizing Map (SOM) Papers 1981-1997, Neural Computing Surveys, Vol. 1, Pp. 102-350.

10. Kentrias, S. (2001) Customer relationship management: The SAS perspective, www.cm2day.com, Last Accessed during September 2013.

11. Kohonen, T., Hynninen, J., Kangas, J. and Laaksonen, V. (1996) SOM PAK: The Self-Organizing Map program package, Helsinki University of Technology, Finland, Report A31, Pp. 1-27.

12. Oja, M., Kaski, S. and Kohonen, T. (2002) Bibliography of Self-Organizing Map (SOM) Papers: 1998-2001 Addendum, Neural Computing Surveys, Vol. 1, Pp. 1-176.

13. Wu, R., Peters, W. and Morgan, M.W. (2002) The Next Generation Clinical Decision Support: Linking Evidence to Best Practice, Journal of Healthcare Information Management, Vol. 16, No.4, Pp. 50-55.

14. Yu, C. and Ying, X. (2009) Application of Data Mining Technology in E-Commerce, International Forum on Computer Science-Technology and Applications, Vol. 1, Pp. 291-293.

Author Biographies

P.Senthil Vadivu received Ph.D(Computer Science) from Avinashilingam University for Women in 2015, completed M.Phil from Bharathiar university in 2006.Currently working as the Head, Department of Computer Applications, Hindusthan College of Arts and Science, Coimbatore-28. Her Research area is decision trees with neural networks.