Enhanced Rule Induction Algorithm for Customer Relationship Management

Prabha Dhandayudam* and Ilango Krishnamurthi

Department of CSE, Sri Krishna College of Engineering and Technology, Coimbatore, India

Received: 5 Dec. 2012, Revised: 22 Jan. 2013, Accepted: 24 Feb. 2013
Published online: 1 Jul. 2013

Abstract: Customer Relationship Management (CRM) helps businesses to gain insight into the behavior of customers and their value so that the company can increase their profit by acting according to the customer characteristics. In order to analyze the customer needs and behavior, data mining is used to extract information from the customer database. For analyzing the customer behavior, the important attributes in the customer database are first chosen and then they are segmented into groups using clustering algorithm based on those attribute values. The rules are then generated to describe the customers in each group using LEM2 (Learning from Examples Module, version 2) algorithm and the proposed algorithm. These rules can be used by the business people to predict the behavior and to vary their promotional activities like coupon distribution and special discounts. It is observed that the proposed algorithm is better in terms of time complexity and performance measures.

Keywords: Customer Relationship Management, Clustering, Rule Induction, Rough set theory.

1. Introduction

Customer Relationship Management (CRM) technology is a mediator between customer management activities in all stages of a relationship (initiation, maintenance and termination) and business performance. It consists of customer identification, customer attraction, customer retention and customer development. Customer identification is the process of making the customers to buy for the first time by segmenting customer into groups and analyzing customer characteristics. The customer identification is followed by customer attraction which motivates each segment of customers in different way. Customer retention and customer development deals with retaining the existing customers and maximizing the customer purchase value respectively. Customer identification is the most important phase in CRM because once the customer is identified correctly only we can attract, retain and develop them further. In the customer identification phase, customer segmentation involves segmenting customers into predefined number of customers and target customer analysis involves analyzing customer behavior or characteristics. With the help of target customer analysis, the company can predict new customer’s behaviors and also differentiate their attraction process for existing customers [1].

In order to analyze the customer needs and behavior, data mining is used to extract information from the customer database. Data mining is a collection of techniques for efficient automated discovery of previously unknown, valid, novel, useful and understandable patterns in large databases. These patterns are used in an enterprise’s decision making process [2].

Clustering and rule induction are data mining techniques for dividing the data into required number of groups and discovering all possible patterns from the database respectively [3]. So, clustering and rule induction can be used for customer segmentation and target customer analysis of customer identification phase in CRM.

The clustering technique produces clusters in which the data inside a cluster has high intra class similarity and low inter class similarity. The traditional approaches for clustering like partitioning and hierarchical algorithm deals with numerical data whose inherent geometric properties can be exploited to naturally define distance functions between data points. Distance functions like Manhattan, Euclidean and Minkowski are used for allocating a data point to the appropriate clustering [3]. The customer data available for clustering is categorical so the above procedure is infeasible. Rule induction algorithms can be

* Corresponding author e-mail: prabhadhandayudam@gmail.com
used to generate rules or patterns for each cluster produced by clustering algorithm. Rules are in the form of If-Then condition. If part is called as antecedent and Then part is called as consequent. Antecedent contains conditional variables and consequent contains single decision variable. The conditional variables are the attributes in the given data and the decision variable is the cluster number assigned to the data using clustering algorithm. Rules generated for a cluster should cover all the data within that cluster and no rule should be satisfied by any data in other clusters. These properties are called as completeness and consistency of the rule induction algorithm. Usually, two-third of data is used as training data and remaining one-third of the data is used as test data. Rules are generated for training data and these rules are used to predict the decision variable of test data [4].

Rough set theory, proposed by Pawlak in 1982 can be seen as a new mathematical approach to vagueness [5]. It is capable of discovering important facts hidden in that data [6,7]. It has been applied successfully in many practical real-world problems [8,9,10,11]. It can also be used to analyze the underlying market demand in any industry slated for deregulation [12]. As a powerful technology in artificial intelligence (AI) and knowledge discovery in database (KDD), rough set theory has been playing an important role in categorical clustering and rule induction. The experimental results in [13] shows that MMR (Min-Min-Roughness), a rough set based categorical clustering algorithm achieves better result compared to bi-clustering, total roughness, fuzzy set based algorithms and other dissimilarity approaches. There are a number of algorithms of rule induction based on rough set theory been proposed [14,15]. LEM2 (Learning from Examples Module, version 2) algorithm based on rough set theory is frequently used for rule induction; it explores the search space of attribute-value pairs [16].

In this paper, an enhanced rule induction algorithm has been developed to generate rules for clustered customer’s data produced by MMR. The performance of the algorithm is compared with LEM2 algorithm.

The rest of the paper is organized in the following: In Section 2 we describe the overview of customer identification, MMR algorithm and LEM2 algorithm. In Section 3 we propose an enhanced rule induction algorithm. In Section 4 we compare the prediction results obtained using rule induction algorithms. Finally in Section 5 we conclude the best rule induction algorithm according to the criteria chosen for comparison.

2. Literature Review

2.1. Customer Identification

Customer identification, the first phase of CRM consists of customer segmentation and target customer analysis. Customer Segmentation gives a quantifiable way to analyze the customer data and distinguish the customers based on their purchase behavior [17]. It is the process of dividing customers into homogeneous groups on the basis of common attributes [18]. It is typically done by applying some form of cluster analysis to obtain a set of segments [19]. In this way the customers can be grouped into different categories for which the marketing people can employ targeted marketing and thus retain the customers. Target customer analysis is used to analyze the customers in each cluster or segment so as to predict the new customer to the appropriate cluster. Once the customers are segmented, rules can be generated to describe the customers in each group based on their purchase behavior. These rules can be used to classify the new customers to the appropriate group who have similar purchase characteristics. RFM model is used to identify and represent the customer characteristics by three attributes namely Recency ($R$), Frequency ($F$) and Monetary ($M$). $R$ indicates the interval between the time that the latest consuming behavior happens and present. $F$ indicates the number of transactions that the customer has done in a particular interval of time. $M$ indicates the total value of the customer’s transaction amount in a particular interval of time [17]. The RFMP model is the modified RFM model where the payment details of the customers are considered. $F$ indicates the average time interval between payment and purchase date. Payment detail of the customer is an important attribute because any two customers with same $R,F,M$ value but different $P$ value cannot be treated equally by the company. The customers are segmented using their consuming behavior via RFMP attributes. This ensures that the customer segmentation is done objectively [20]. The values for $R,F,M$ and $P$ attributes are continuous. The continuous values are normalized to categorical values as very low, low, middle, high and very high for effective analysis.

2.2. MMR algorithm

Once the important attributes ($R,F,M$ and $P$) of the customer are chosen, the customers are segmented using MMR algorithm. It proposes a new way to measure data similarities based on the roughness concept. The measure roughness allows an object to belong to a cluster with different degrees of belonging. Let $A$ is a set of attributes, $a_i \in A, V(a_i)$ refers to the set of values of attribute $a_i$. $X$ is a subset of objects having one specific value $\alpha$ of attribute $a_i$ that is $X(a_i = \alpha)$. $|X_{a_j}(a_i = \alpha)|_{LB}$ refers to the upper approximation and $|X_{a_j}(a_i = \alpha)|_{UB}$ refers to the upper approximation with respect to $\{a_j\}$, then $R_{a_j}(X)$ is defined as the roughness of $X$ with respect to $\{a_j\}$, that is

$$R_{a_j}(X \mid a_i = \alpha) = 1 - \frac{|X_{a_j}(a_i = \alpha)|_{LB}}{|X_{a_j}(a_i = \alpha)|_{UB}} \tag{1}$$

where $a_i, a_j \in A$ and $a_i \neq a_j$. 

\[ \text{Section identification, MMR algorithm and LEM2 algorithm. In Section } 2 \text{ we describe the overview of customer identification, MMR algorithm and LEM2 algorithm. In Section 3 we propose an enhanced rule induction algorithm. In Section 4 we compare the prediction results obtained using rule induction algorithms. Finally in Section 5 we conclude the best rule induction algorithm according to the criteria chosen for comparison.} \]

\[ \text{2. Literature Review} \]

\[ \text{2.1. Customer Identification} \]

\[ \text{Customer identification, the first phase of CRM consists of customer segmentation and target customer analysis. Customer Segmentation gives a quantifiable way to analyze the customer data and distinguish the customers based on their purchase behavior [17]. It is the process of dividing customers into homogeneous groups on the basis of common attributes [18]. It is typically done by applying some form of cluster analysis to obtain a set of segments [19]. In this way the customers can be grouped into different categories for which the marketing people can employ targeted marketing and thus retain the customers. Target customer analysis is used to analyze the customers in each cluster or segment so as to predict the new customer to the appropriate cluster. Once the customers are segmented, rules can be generated to describe the customers in each group based on their purchase behavior. These rules can be used to classify the new customers to the appropriate group who have similar purchase characteristics. RFM model is used to identify and represent the customer characteristics by three attributes namely Recency ($R$), Frequency ($F$) and Monetary ($M$). $R$ indicates the interval between the time that the latest consuming behavior happens and present. $F$ indicates the number of transactions that the customer has done in a particular interval of time. $M$ indicates the total value of the customer’s transaction amount in a particular interval of time [17]. The RFMP model is the modified RFM model where the payment details of the customers are considered. $F$ indicates the average time interval between payment and purchase date. Payment detail of the customer is an important attribute because any two customers with same $R,F,M$ value but different $P$ value cannot be treated equally by the company. The customers are segmented using their consuming behavior via RFMP attributes. This ensures that the customer segmentation is done objectively [20]. The values for $R,F,M$ and $P$ attributes are continuous. The continuous values are normalized to categorical values as very low, low, middle, high and very high for effective analysis.} \]

\[ \text{2.2. MMR algorithm} \]

\[ \text{Once the important attributes ($R,F,M$ and $P$) of the customer are chosen, the customers are segmented using MMR algorithm. It proposes a new way to measure data similarities based on the roughness concept. The measure roughness allows an object to belong to a cluster with different degrees of belonging. Let $A$ is a set of attributes, $a_i \in A, V(a_i)$ refers to the set of values of attribute $a_i$. $X$ is a subset of objects having one specific value $\alpha$ of attribute $a_i$ that is $X(a_i = \alpha)$. $|X_{a_j}(a_i = \alpha)|_{LB}$ refers to the upper approximation and $|X_{a_j}(a_i = \alpha)|_{UB}$ refers to the upper approximation with respect to $\{a_j\}$, then $R_{a_j}(X)$ is defined as the roughness of $X$ with respect to $\{a_j\}$, that is} \]

$$R_{a_j}(X \mid a_i = \alpha) = 1 - \frac{|X_{a_j}(a_i = \alpha)|_{LB}}{|X_{a_j}(a_i = \alpha)|_{UB}} \tag{1}$$

\[ \text{where $a_i, a_j \in A$ and $a_i \neq a_j$.} \]
Let \(|V(a_i)|\) be the number of values of attributes \(a_i\), the mean roughness on attribute \(a_i\) with respect to \(\{a_j\}\) is defined as

\[
\text{Rough}_{a_j}(a_i) = \frac{R_{a_j}(X \mid a_i = \alpha_1) + \cdots + R_{a_j}(X \mid a_i = \alpha_i V(a_i))}{|V(a_i)|} \tag{2}
\]

where \(a_i, a_j \in A\) and \(a_i \neq a_j\).

Given \(n\) attributes, MR, min-roughness of attribute \(a_i\) \((a_i \in A)\) refers to the minimum of the mean roughness, that is, \(\text{MR}(a_i) = \min(\text{Rough}_{a_1}(a_i), \ldots, \text{Rough}_{a_j}(a_i))\) where \(a_i, a_j \in A\) and \(a_i \neq a_j\). The MMR is defined as the minimum of the Min-Roughness of the \(n\) attributes. That is, \(\text{MMR}(\text{MR}(a_1), \ldots, \text{MR}(a_i), \ldots)\) where \(a_i \in A\), \(i\) goes from 1 to cardinality(A). Attribute-value pairs from the data set with minimum roughness is chosen as splitting attribute and the splitting is done based on the value of the attribute which has that minimum roughness. At subsequent iterations, the leaf node having more objects is selected for further partition. The algorithm terminates when it reaches a pre-defined number of clusters [13].

### 2.3. LEM2 algorithm

After segmenting the customers using MMR algorithm, rules can be generated to describe the customers in each group based on their purchase behavior. Rule induction is the technique used to find regularities hidden in the data and expressed in terms of rules. It belongs to supervised learning when data are already grouped. The input for rule induction comes from clustering algorithm. Rules are in the form of

if (attribute-1,value-1) and (attribute-2,value-2) and \ldots and (attribute-n,value-n) then (decision,value)

In the database, each row is called as a case and each column is called as an attribute. Attributes are independent variables and decision is a single dependent variable. Here, Recency, Frequency, Monetary, Payment are attributes and cluster number is the decision variable. The set of all cases labeled by same decision value is called a concept. A case \(x\) is covered by a rule \(r\) if and only if every condition (attribute-value pair) of \(r\) is satisfied by the corresponding attribute value for \(x\). A concept \(C\) is completely covered by a rule set \(R\) if and only if for every case \(x\) from \(C\) there exists a rule \(r\) from \(R\) such that \(r\) covers \(x\). \(R\) contains set of rules for each decision value. \(R\) is complete if and only if every concept from the data set is completely covered by \(R\). A rule \(r\) is consistent if and only if for every case \(x\) covered by \(r\), \(x\) is a member of the concept \(C\) indicated by \(r\). \(R\) is consistent if and only if every rule from \(R\) is consistent with the data set. Rule induction produces complete and consistent rule set [21].

A block of an attribute-value pair \(t = (a, v)\), denoted \([t]\), is the set of all examples that for attribute a have value \(v\). A concept, described by the value \(w\) of decision \(d\), is denoted \([d, w]\), and it is the set of all examples that have value \(w\) for decision \(d\). Let \(B\) be a concept and let \(T\) be a set of attribute-value pairs. Concept \(B\) depends on a set \(T\) if and only if

\[\emptyset \neq [T] = \bigcap_{t \in T} [t] \subseteq B\] \tag{3}

Set \(T\) is a minimal complex of concept \(B\) if and only if \(B\) depends on \(T\) and \(T\) is minimal. Let \(\tau\) be a nonempty collection of nonempty sets of attribute-value pairs. Set \(\tau\) is a local covering of \(B\) if and only if the following three conditions are satisfied:

1. each member of \(\tau\) is a minimal complex of \(B\),
2. \(\bigcup_{T \in \tau} [T] = B\)
3. \(\tau\) is minimal (\(\tau\) has the smallest possible number of members)

For each concept \(B\), the LEM2 algorithm induces production rules by computing a local covering \(\tau\). Any set \(T\), a minimal complex which is a member of \(\tau\), is computed from attribute-value pairs selected from \(T(G)\) of attribute-value pairs relevant with a current goal \(G\), i.e., pairs whose blocks have nonempty interaction with \(G\). The initial goal \(G\) is equal to the concept and then it is iteratively updated by subtracting from \(G\) the set of examples described by the set of minimal complexes computed so far. Attribute-value pairs from \(T\) which are selected as the most relevant, i.e., on the basis of maximum of the cardinality of \([t] \cap G\), if a tie occurs, on the basis of the small cardinality of \([t]\). The last condition is equivalent to the maximal conditional probability of goal \(G\) given attribute-value pair \(t\). For a set \(X\), \(|X|\) denotes the cardinality of \(X\) [22]. The procedure of LEM2 is shown in algorithm 1.

The algorithm is run exactly \(|d|\) times, where \(|d|\) is the number of decision classes. The number of decision classes indicates the number of clusters produced by MMR algorithm. The while loop \((G \neq \emptyset)\) is performed at most \(n\) times because we may have the whole set as the upper approximation to every decision class. Here \(n\) is the number of objects in the training set. To select a pair \(t \in T(G)\) as the best one, we have to iterate \(n \times m\) times so that all possible pairs of attributes and values are examined. Here \(m\) is four which indicates the number of attributes in the training set. \(T\) contains \(m\) elements at most and \(\tau\) contains \(n\) elements at most. So the computational complexity of for loop (for each \(t \in T)\) is \(m \times n\). Therefore the total computational complexity of LEM2 is equal to \(O(|d| \times n \times (n \times m) \times (m \times n))\) which is simplified as \(O(|d| \times m^2 \times n^3)\).
3. Proposed Algorithm

In LEM2 algorithm, the rules generated for each cluster is complete and consistent but it doesn’t produce all the consistent rules in a cluster because once a consistent rule is discovered, the objects satisfying that rule is eliminated and rules are discovered for the rest of the objects. Due to this the number of rules produced for a particular cluster becomes less and consequently the chances of predicting the customer to the correct cluster becomes less. In order to overcome this disadvantage the proposed rule induction algorithm produces all the consistent rules and complete rules for the objects in the cluster. Target cluster is the cluster for which rules are generated. Remaining clusters are the clusters other than target cluster. A block of an attribute-value pair \( t = (a, v) \), denoted \( [t] \), is the set of all examples that for attribute \( a \) have value \( v \). A block of \( n \) attribute-value pair \( t_1 = (a_1, v_1), t_2 = (a_2, v_2) \), and so on, \( t_n = (a_n, v_n) \) denoted \( [t_1, t_2, \ldots, t_n] \), is the set of all examples that for attribute \( a_1 \) have value \( v_1 \), for attribute \( a_2 \) have value \( v_2 \), and so on, \( a_n \) have value \( v_n \). A block of size \( 1 \) has one attribute-value pair. A block of size \( n \) has \( n \) attribute-value pairs. For a set \( X \), \(|X|\) denotes the cardinality of \( X \). The procedure for enhanced rule induction algorithm is shown in algorithm 2.

In the while loop of \( G \neq \emptyset \), find \([t]\) having block size 1 and then block size 2 and so on until block size \( m \). Here \( m \) indicates four which denotes the number of attributes in the data set. In each iteration of while loop, \( G \) contains set of objects whose \( T(G) \) contains set of all attribute-value pairs which satisfies the following three conditions:

1. \(|t| \cap B \neq \emptyset\)
2. \(|t| \cap B \neq |t| \cap G|
3. \(|t| \cap C = \emptyset\)

The first condition chooses the pairs whose blocks have non-empty interactions with \( B \). The last condition chooses the pairs whose blocks have empty interactions with \( C \). The first and last conditions are required to satisfy the consistency property of rule generating algorithm. The second condition chooses the pairs whose cardinality of blocks satisfied in \( B \) is not equal to cardinality of blocks satisfied in \( G \). This condition is required so that the rules generated are not redundant. The covering property of rule generating algorithm is satisfied by choosing \( G \neq \emptyset \) as the while loop condition. In each iteration of while loop, rules are generated as the attribute-value pairs of \( t \).

The algorithm is run exactly \(|d|\) times, where \(|d|\) is the number of decision classes. The number of decision classes indicates the number of clusters produced by MMR algorithm. The while loop \((G \neq \emptyset)\) is performed at most \( n \) times because we may have the whole set as the upper approximation to every decision class. Here \( n \) is the number of objects in the training set. To calculate \( T(G) \), all the objects are examined so the computation is \( n \). To select a pair \( t \in T(G) \) as the best one, we have to iterate \( n \times m \) times so that all possible pairs of attributes and values are examined. Here \( m \) is the number of attributes in the training set. Therefore the total computational complexity of proposed algorithm is equal to \( O(|d| \times n \times n \times (n \times m)) \) which is simplified as \( O(|d| \times m \times n^3) \). The proposed algorithm is an enhanced algorithm in terms of time complexity because LEM2 algorithm computation is \( m \) times more than the proposed algorithm.

Algorithm 1: LEM2 algorithm

\[
\text{begin} \\
\quad G = B \\
\quad \tau = \emptyset \\
\quad \text{while } G \neq \emptyset \text{ do} \\
\quad \quad T = \emptyset \\
\quad \quad \text{while } T = \emptyset \text{ or } |T| \supseteq B \text{ do} \\
\quad \quad \quad \text{Select a pair } t \in T(G) \text{ such that } |t \cap G| \text{ is maximum; if a tie occurs, select a pair } t \in T(G) \text{ with the smallest cardinality of } |t| \\
\quad \quad \quad \text{if another tie occurs, select first pair } t \in T(G) \\
\quad \quad \quad G = |t \cap G| \\
\quad \quad \quad T(G) = \{ t | t \cap G \neq \emptyset \} \\
\quad \quad \quad T(G) = T(G) - T \\
\quad \quad \quad \text{foreach } t \in T \text{ do} \\
\quad \quad \quad \quad \text{if } |T - \{t\}| \subseteq B \text{ then} \\
\quad \quad \quad \quad \quad T = T - \{t\} \\
\quad \quad \quad \text{end} \\
\quad \quad \quad \tau = \tau \cup \{T\} \\
\quad \quad \quad G = B - \bigcup_{t \in \tau} |t| \\
\quad \quad \text{end} \\
\quad \text{foreach } T \in \tau \text{ do} \\
\quad \quad \text{if } \bigcup_{t \in \tau - \{T\}} |t| = B \text{ then} \\
\quad \quad \quad \tau = \tau - \{T\} \\
\quad \quad \text{end} \\
\text{end} \\
\text{Algorithm 1: LEM2 algorithm}
\]

Algorithm 2: Enhanced Rule Induction algorithm

\[
\text{begin} \\
\quad U - \text{Set of all objects in the data set} \\
\quad B - \text{Set of all objects in the target cluster} \\
\quad C = U - B \text{ (set of all objects in } U \text{ but not in } B) \\
\quad G = B \\
\quad \text{temp} = \emptyset \\
\quad \text{while } G \neq \emptyset \text{ do} \\
\quad \quad T(G) = \{ t | t \cap B \neq \emptyset \text{ and } |t \cap B| \neq |t \cap G| \} \\
\quad \quad \text{foreach pair } t \in T(G) \text{ do} \\
\quad \quad \quad \text{temp} = \text{temp} \cup \{t \cap B\} \\
\quad \quad \text{end} \\
\quad \quad G = B - \text{temp} \\
\text{end} \\
\text{Algorithm 2: Enhanced Rule Induction algorithm}
\]
4. Experimental Results

Real data set of the customer transaction details are used for the clustering and rule induction algorithms. The data set consists of 12,028 records of customer transaction for a period of three months for a fertilizer manufacturing company. For each transaction, party id, date of purchase, amount of purchase and payment of purchase are used to define \( R, F, M \) and \( P \) values. For each distinct party id, \( R \) is calculated as the interval between the time that the latest consuming behavior happens and present, \( F \) is calculated as the number of his/her transaction records, \( M \) is calculated as the sum of his/her purchase amount and \( P \) is calculated as the average time interval (in terms of days) between his/her payment date and his/her purchase date for each transaction in the data set. Now the data set has only four attributes namely \( R, F, M \) and \( P \) for 3278 customers. Next the values of \( R, F, M \) and \( P \) are normalized as given below. For normalizing \( R \) or \( P \)

1. The data set is sorted in descending order of the \( R \) or \( P \)
2. Divide the data set into five equal parts of 20% record in each
3. Assign categorical values as very low, low, middle, high and very high to first, second, third, fourth, fifth part of records respectively

For normalizing \( F \) or \( M \)

1. The data set is sorted in ascending order of the \( F \) or \( M \)
2. Divide the data set into five equal parts of 20% record in each
3. Assign categorical values as very low, low, middle, high and very high to first, second, third, fourth, fifth part of records respectively

The normalized data set is now used by MMR clustering algorithm to segment the 3,278 customers into three groups. LEM2 and proposed rule induction algorithms are used to generate rules for training data (two-third in each cluster). The test data (remaining one-third in each cluster) is given as input for LEM2 and proposed rule induction algorithm to predict the cluster value according to their generated rules for training data. The training and testing data are mutually exclusive. The performance criteria for prediction using rule induction algorithms are calculated for all the ten cases. The Table 1 shows the false positive, false negative, true positive, true negative produced by the rule induction algorithms for all the ten cases. The objective of the rule induction algorithm is to maximize true positive and true negative. From the Table 1 it is observed that the proposed rule induction algorithm has minimum \( FP \), minimum \( FN \), maximum \( TP \) and maximum \( TN \) for all the ten cases when compared to LEM2.

Sensitivity, specificity, accuracy, PPV, NPV and \( F \)-measure are calculated using formula 4 to 9 respectively for each algorithm in all the ten cases. The output is shown graphically in Figures 1 to 6. The objective of the rule induction algorithm is to maximize sensitivity, specificity, accuracy, PPV, NPV and \( F \)-measure. From the Figures 1 to 6 it is observed that the proposed rule induction algorithm has equal or maximum value than LEM2 for all the ten cases. The percentage increase in each performance criteria might seem to a smaller value but in real data set where customers are in terms of thousands not in hundreds the proposed algorithm has significant improvement than LEM2.

LEM2 produces only the minimal set of rules whereas proposed rule induction algorithm produces all the
Table 1: FP, FN, TP and TN for rule induction algorithms

<table>
<thead>
<tr>
<th>Case</th>
<th>False Positive</th>
<th>False Negative</th>
<th>True Positive</th>
<th>True Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>1</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>0</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>1</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>11</td>
<td>1</td>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

1—LEM2 and 2—Proposed

possible set of consistent rules to describe the records in the cluster. When the number of rules to describe the cluster is increased then naturally the prediction of an object to its cluster is more accurate.

Figure 1: Sensitivity for rule induction algorithms

Figure 2: Specificity for rule induction algorithms

Figure 3: Accuracy for rule induction algorithms

Figure 4: Positive Predictive Value for rule induction algorithms

Figure 5: Negative Predictive Value for rule induction algorithms

5. Conclusions

Customer Relationship Management is a technology that manages relationship with customers in order to improve the performance of business. In CRM, the customer identification is the important phase because it involves segmenting the customers and analyzing their behavior for further customer attraction, retention and development. In this paper clustering technique in data mining has been used for customer segmentation and rule induction is used for describing customer behavior in each segment. Business people employ different benefit...
schemes for customer in different clusters or segments. So, classifying a customer to the cluster plays an important role in CRM. For a good rule induction algorithm, the customer’s behavior in each cluster should be correctly characterized so that the new customers are predicted to the appropriate cluster. The performance evaluation criteria were chosen based on the prediction accuracy of rule induction algorithm. The experimental result shows that the proposed algorithm is better than LEM2. It has been proved that the time complexity of LEM2 is \( m \) times more than the proposed algorithm where \( m \) indicates the number of attributes chosen for analysis. Thus, it has been evident from the results that the proposed algorithm is an enhanced rule induction algorithm which produces better performance in prediction and has less computation when compared to LEM2 algorithm.

References

Prabha Dhandayudam received B.E and M.E degrees in Computer Science and Engineering in the year 2002 and 2006 respectively. She received gold medal in her post graduate degree. She is currently an Assistant Professor in Sri Krishna College of Engineering and Technology. She has a total of seven year experience in teaching. She is doing research in Applying Data Mining Techniques for Customer Relationship Management.

Ilango Krishnamurthi received his Ph.D in Computer Science and Engineering from the Indian Institute Technology, Chennai. He graduated from BITS, Pilani and received masters’ degree from Iowa State University, USA both in the fields of Computer Science & Engineering. His current research interests are in the areas of semantic web, mobile computing and data mining.