Advantages Of Decision Trees Using Data Mining In Indian Retail Industry

Jayanthi Ranjan [1], Ruchi Agarwal [2]

Institute of Management Technology [1], Birla Institute of Technology [2], India

ABSTRACT

Indian Retail industry has emerged as one of the most dynamic and fast paced industries with several players entering the market. The data that retail industry collect about their customers is one of the greatest assets of it. Data mining (DM) helps in extracting the buried valuable information within the vast amount of data. The decision trees using DM could make a significant difference to the way in which a retail industry run their business, and interact with their current and prospective customers. The derived information can be utilized in predicting, forecasting and estimating the important business decisions, which can help in giving a retailer the competitive edge over their competitors. The paper demonstrates the advantages of decision trees using DM in Indian retail industry with the help of an empirical study.

Keywords: *Data mining, Decision trees, Retail industry, Customers*

1. Introduction

In the recent years the significant changes are done in the retail industry which has important implications on DM. Retail industry is using information technology (IT) for generating, storing and analyzing mass produced data not only for operational purposes but also for enabling strategic decision making to survive in a competitive and dynamic environment. DM helps in reducing information overload along with the improved decision-making by searching for relationships and patterns from the huge dataset collected by organizations. It enables a retail industry to focus on the most important information in the database and allows retailers to make more knowledgeable decisions by predicting future trends and behaviors. The DM uses the business data as raw material using a predefined algorithm to search through the vast quantities of raw data, and group the data according to the desired criteria that can be useful for the future target marketing (Ahmed, 2004). Through DM and the new knowledge it provides, individuals are able to leverage the data to create new opportunities or value for their organizations (Wu, 2002). DM helps in extracting diamonds of knowledge from the historical data, and predicts future outcomes. Ranjan et al. (2008) demonstrated the effect of DM in better decision making in human resource management system. DM helps in optimizing business decisions. Berman and Evans (2008) opinioned that data mining is used by retail executives and other employees and sometimes channel partners- to analyze information by customer type, product category, and so forth in order to determine opportunities for tailored marketing efforts that would lead to better retailer performance.

Decision trees are well known methods of predictive modeling used for DM purposes since they provide interpretable rules and logic statements which enable more intelligent decision making. Decision trees create a segmentation of the original data set. The predictive segments that are derived from the decision tree come with a description of the characteristics that define the predictive segment. Thus the decision trees and the algorithms that create them may be complex, but the results can be presented in an easy-to-understand way that can be quite useful to the business user (Berson and Smith, 2008). Gearj et al. (2007) demonstrated that decision tree diagramming is a demanding yet flexible technique which allows the representation of sequential decisions and subjectively based data in a readily understood form. Sheu et al. (2008) found that the consumers' past online shopping experience would directly affect their decision-making. Yang et al. (2008) use decision tree and association rules to predict cross selling opportunities.

The arrival of retail boom caused the global technology vendors to quickly get into the marketplace with solutions that claim to make retailers' lives simpler. Retailers have to put in great efforts to really know their customers. Retail industry emphasized on quick delivery of customer focused services (offers, promos, etc) since adapting to customer needs in a very limited period of time is also very important. Retailers continuously get the advantage from information collected from customers' transactions. Hence requirements of retail, technology wise would encompass business intelligence, data mining/warehousing, and other similar technologies since using these, retailers can constantly benefit from newly observed trends based on user purchases (Sohoni, 2007). The changing consumption patterns trigger changes in shopping styles of consumers and also the factors that drive people into stores (Kaur and Singh, 2007). Hou and Tu (2008) addressed that the managers in the contemporary marketing must importantly identify potential customer relationships to positively affect corporate performance. Ranjan and Bhatnagar (2008) opinioned that the optimization of revenue can be accomplished by a better understanding of customers, based on their purchasing patterns and demographics, and better information empowerment at all customers touch points, whether with employees or other media interfaces. With the retail boom, companies are likely to deploy IT tools that help them enhance the end-customer's experience. Jones and Ranchhod (2007) expressed that the strategic focus is required on the real complexity of the relationship that organizations are initially able to establish with customers. Sangle and Verma (2008) opinioned that the customer relationship management unites the potential of marketing strategies and IT to create profitable, long-term relationships with customers and helps in enhancing the opportunities to use data and information to both understand customers and co-create value with them.

The paper proceeds as follows: Section 2 presents Literature Review. Section 3 explains Research Methodology. Section 4 discusses about Indian Retail Industry. Section 5 explains the concept of Data Mining. Section 6 presents advantages of Decision trees in retail industry. Section 7 concludes the paper.

2. Literature Review

With the retail boom and the dynamic competitive environment, every retailer must make decisions in the face of uncertainty, and live with the consequences. Before making a decision, retailer should analyze the outcomes of a few alternative actions which help in determining whether a decision will produce the favorable consequences or not. The consequences of a decision in the retail industry are analyzed by using a decision tree to gain competitive edge over the competitors. DM is being used widely in the context of business but the advantages of decision trees using DM are not explored. This is the motivation of our paper.

Sheu et al. (2008) found that the consumers' past online shopping experience would directly affect their decision-making. Ranjan et al. (2008) demonstrated the effect of DM in better decision making in human resource management system Yang et al. (2008) use decision tree and association rules to predict cross selling opportunities. Gearj et al. (2007) demonstrated that decision tree diagramming is a demanding yet flexible technique which allows the representation of sequential decisions and subjectively based data in a readily understood form. Wang et al. (2008) found the application of Decision Trees in Mining High-Value Credit Card Customers.

Sarantopoulos (2003) described the development and the validation of a decision tree, which aims to discriminate between good and bad accounts of the customers of a particular retailer based on a sample of orders placed between certain periods of time. Lemmens and Croux (2006) explored the bagging and boosting classification techniques which significantly improved the accuracy in predicting churn. Lima et al. (2009) showed how the domain knowledge can be incorporated in the data mining process for churn prediction by analysing a decision table extracted from a decision tree or rule-based classifier. Velikova and Daniels (2004) presented methods to enforce monotonicity of decision trees for price prediction. Chen and Hung (2009) used decision trees to summarize associative classification rules. Lee and Siau (2001) reviewed data mining techniques. Hou and Tu (2008) found that business with customer relationship management practices is linked to better performance outcomes, including perceptual and financial performance. Jones and Ranchhod (2007) augmented the concepts from technology-enabled customer relationship management towards an exploratory framework, designed to explore the nature of customer attention. Sangle and Verma (2008) identified and analyzed the determinants of adoption of customer relationship management in Indian service sector. Ranjan and Bhatnagar (2008) presented the benefit and application of the data mining tools through which the firm achieves competitive advantage by selecting the best suited data mining tool according to their need.

3. Research Methodology

Decision trees are used for representing a set of decisions by their tree-shaped structure and can generate rules for the classification of the dataset. They are very important for a retailer since it helps in strategic decision making.

The customer transaction data is very valuable asset for any company hence the need for research design was felt. So, the data for this paper was collected in two phase. First the primary data is collected through various sources which include personal interviews, surveys and filled questionnaire, review the available online software packages, attending conferences and seminars, etc. Secondary data is collected through studying the literature related to research that is available in various journals, books, magazine, websites, established doctoral thesis, etc.

The authors got the customer transaction database of one retail firm (name masked) which is analyzed with the help of data mining tool SPSS' Clementine. The basic objective is to study the advantages of decision trees using DM in Indian retail industry with the help of an empirical study.

4. Indian Retail Industry

The increased globalization, market saturation, and increased competitiveness give rise to mergers and acquisitions. Indian retailers are seeking competitive advantages by better improving relationships with customers which has taken on new life. Rogers (2005) addressed that the companies recognize that customer relationships are the underlying tool for building customer value, and they are finally realizing that growing customer value is the key to increasing enterprise value.

The retail sector is growing rapidly in the Indian scenario as well as globally. With the Indian retail sector booming, it brings immense opportunities for foreign as well as domestic players. The changing lifestyle of the Indian consumer makes it essential for the retailers to understand the patterns of consumption. The changing consumption patterns trigger changes in the shopping styles of consumers and also in the factors that drive people into stores (Kaur and Singh, 2007). The Indian retail has been transformed due to the attitudinal shift of the Indian consumer in terms of choice preference, value for money and the emergence of organized retail formats. Rising incomes, increased advertising, and a jump in the number of women working in the country's urban centers have made goods more attainable and enticing to a larger portion of the population. At the same time, trade liberalization and more sophisticated manufacturing techniques create goods that are less expensive and higher quality (Hanna, 2004). Pande and Collins (2007) explored to centralize the retail supply chain in India with the goal to improve overall retail business in India.

Vector (2007) explored that the Retail is India's largest industry with the market size of around US $312 billion in which organized retailing comprises only 2.8 per cent of the total retailing market and is estimated at around US$ 8.7 billion. The organized retail sector is expected to grow to US $ 70 billion by 2010. FICCI Retail Report (20007) reported that the estimates predict that the overall size of the retail sector in India is expected to touch US$427 billion by 2010 and US$637 billion by 2015 with the modern segment expected to account for 22 per cent by 2010, up from the present four per cent.

5. Data Mining

Data Mining is a process of analyzing the data from different perspectives and presenting it in a summarized way into useful information. It extracts patterns and trends that are hidden among the data. It is often viewed as a process of extracting valid, previously unknown, non-trivial and useful information from large databases (Rao, 2003). Han and Kamber (2007) expressed that the DM is extracting or mining knowledge from large amount of data. Feelders et al. (2000) opinioned that the DM is the process of extracting information from large data sets through the use of algorithms and techniques drawn from the field of statistics, machine learning and database management systems. Noonan (2000) explained that DM is a process for sifting through lots of data to find information useful for decision making. It helps in predicting the future of the business. It can make the improvement in every industry throughout the world. The data can be mined and the results can be used to determine not only what the customers wants, but to also predict what they will do. West (2005) addressed that by relying on the power of data mining, retailers can maintain the consistency and accuracy of their underwriting decisions; they can significantly reduce the impact of fraudulent claims; and can have a better understanding of their customer's wants and needs. It can be used to control costs as well as contribute to revenue increases (Two Crows Corporation, 2005).

The DM software uses the business data as raw material using a predefined algorithm to search through the vast quantities of raw data, and group the data according to the desired criteria that can be useful for the future target marketing (Ahmed, 2004). DM involves the use of predictive modeling, forecasting and descriptive modeling techniques. By using these techniques, a retail firm can proactively manage customer retention, identify cross-sell and up-sell opportunities, profile and segment customers, set optimal pricing policies, and objectively measure and rank which suppliers are best suited for their needs (Bhasin, 2006). DM applications automate the process of searching the huge amount of data to find patterns that are good predictors of purchasing behaviors. After mining the data, marketers must feed the results into campaign management software that manages the campaign directed at the defined market segments (Thearling, 2007).

Wang and Wang (2007) pointed out that the DM techniques for the online customer segmentation helps in clustering the customers on the basis of the characteristic that they show while purchasing the product online or surfing the net. Chen, Wu and Chen (2005) effectively discovered the current spending pattern of customers and trends of behavioral change by using DM tools, which would allow management to detect in a large database potential change of customer preference, and provide products and services faster as desired by the customers to expand the client base and prevent customer attrition. Pan et al. (2007) found that the problem of classification of the customer is cost sensitive in nature. Consumer-focused companies with sizable caches of information on current and potential customers such as retailers are ideal for data mining technology (Cowley, 2005).

Chen and Liu (2005) focused on enhancing the functionality of current applications of DM. Berry and Linoff (2001) expressed that only through the application of DM techniques can a large enterprise hope to turn the myriad records in its customer databases into some sort of coherent picture of its customers. It can also be used to locate individual customers with specific interests or determine the interests of a specific group of customers (Guzman, 2002). Berman and Evans (2008) opinioned that DM is used by retail executives and other employees-and sometimes channel partnersto analyze information by customer type, product category, and so forth in order to determine opportunities for tailored marketing efforts that would lead to better retailer performance.

6. Advantages of Decision trees in Retail Industry

Decision trees are an excellent tool in decision-making and DM systems in retail industry. They provide good service to any analyst or manager. This is further explained in the following subsections:

6.1. Decision Trees

Decision trees provide an effective method of decision making in retail industry. Savage (2003) opinioned that the decision trees can sharpen and formalize the decision-making process. It helps in making the best decisions on the basis of existing information. Decision trees helps in choosing between several courses of action. They define a tree structure in which leaves represent classifications and branches represent conjunctions of features that lead to those classifications. This is a very effective structure in which options can be laid and the possible outcomes of choosing those options can be investigated. They also help in forming a balanced picture of the risks and rewards associated with each possible course of action. D'Souza (2007) expressed that a decision tree can be learned by splitting the source data set into subsets based on an attribute value test in which the process is repeated on each derived subset in a recursive manner and the recursion is completed when either splitting is non-feasible or a singular classification can be applied to each element of the derived subset. A decision tree helps in partitioning the data into smaller segments called terminal nodes or leaves which are homogeneous with respect to a target variable. Partitions are defined in terms of input variables which define a predictive relationship between the inputs and the target. This partitioning continues until the subsets cannot be partitioned any further using user-defined stopping criteria. By creating homogeneous groups, retailers can predict with greater certainty how customers in each group will behave.

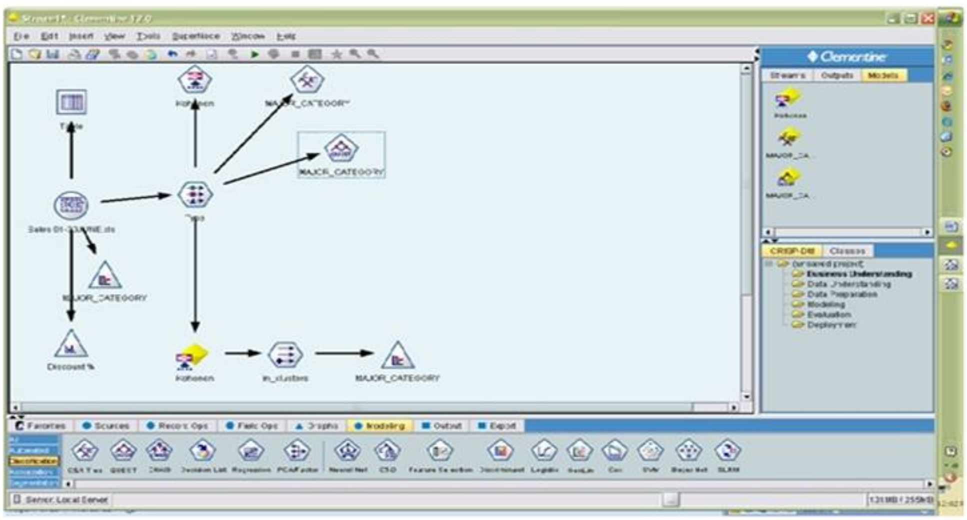
Decision trees are used in segmenting groups of customers and developing customer profiles which helps marketers to produce targeted promotions and achieve higher response rates. The main goals of data analysis and data mining are to predict future outcomes and identify factors that can produce desired effect. Sarantopoulos (2003) described the development and the validation of a decision tree, which aims to discriminate between good and bad accounts of the customers of a particular retailer based on a sample of orders placed between certain periods of time. Gearj et al. (2007) demonstrated that decision tree diagramming is a demanding yet flexible technique which allows the representation of sequential decisions and subjectively based data in a readily understood form.

Decision trees are used in either estimating a metric target variable or classifying observations into one category of a non-metric target variable by repeatedly dividing observations into mutually exclusive and exhaustive subsets. So, the algorithm used for constructing decision trees is also referred to as recursive partitioning algorithm. In a decision tree, each observation is eventually assigned to a node (also called leaf) that has a predicted value or classification. The end product can be graphically represented by a tree-like structure (called a decision tree), which is a compact representation of the data. The end product can also be represented by explicit decision rules. The resulting visual representation and explicit rules make decision trees easy to interpret and use. Decision trees can also be used in modeling complex non-linear and interaction relationships reasonably well. Many algorithms are available to construct decision trees. The more common ones are CHAID (Chi-square Automatic Interaction Detection), C5.O (a proprietary algorithm) and CART (Classification and Regression Trees). Some algorithms are used for metric target variables only, some for non-metric target variables only and some for both. Decision tree algorithms are very intensive (i.e. a lot of computations are performed to construct the tree).

6.1.1. Classification And Regression Trees: Empirical Study

Classification and Regression Trees (CART) is a data exploration and prediction algorithm developed by Leo Breiman, Jerome Friedman, Richard Olshen, and Charles Stone (Berson and Smith, 2008). It is a tree- based classification and prediction method that uses recursive partitioning to split the training records into segments with similar output field values. It is a robust, easy-to-use decision tree that automatically sifts large, complex databases, searching for and isolating significant patterns and relationships which is then used to generate reliable, easy-to-grasp predictive models for applications such as finding best prospects and customers, targeted marketing, etc. (Salford System, 2009). Behaviour of purchased product by using Classification & Regression Modeling with the help of data mining tool SPSS' Clementine. The analysis is done on the database of a retail firm (name masked) with the help of SPSS' Clementine tool which is shown in the following figure 1:

Figure 1: Analysis On The Database Of A Retail Firm Using SPSS' Clementine Tool



The results of the analysis are shown in the following Figure 2 & 3

Figure 2: Results Of The Above Analysis

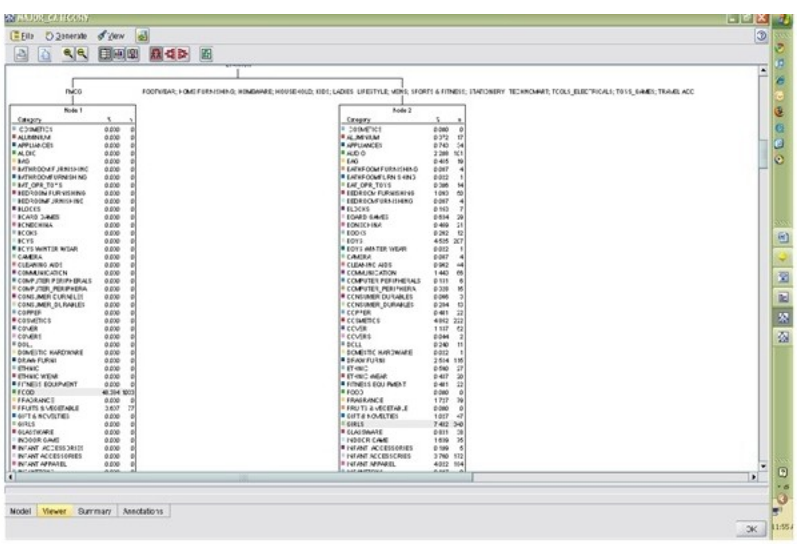
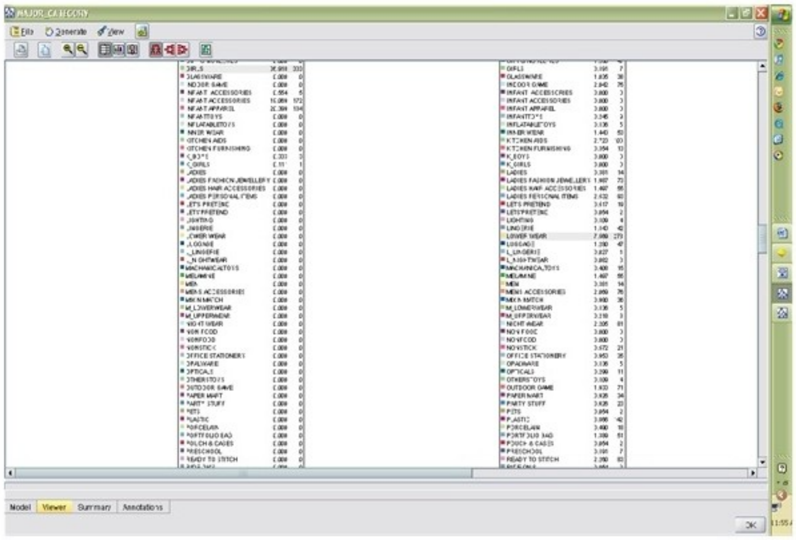


Figure 3: Results of the above analysis (Contd.)



In the above figures, n is the number of records and % represents the percentage of n. Here category of products has been sub divided into two groups FMCG & combinations of other products which will be further sub divided into sub parts. From the above results we see that girl's items are sold more. Further under girls section the lower wears are more sold. Likewise we can see more results and accordingly make decisions.

6.2. Advantages Of Data Mining Enabled Decision Trees In Retail Industry

Data mining enabled decision trees are widely used in retail industry. Its advantages are endless. It collects huge amounts of data on sales, customer shopping history, goods transportation, consumption, and service. The data quantity is continuously expanding exponentially, mainly due to increasing ease, availability, and popularity of business conducted on the web or e-commerce. For DM, retail data is a rich source. Han and Kamber (2007) expressed that the retail DM can help identify customer buying behaviors, discover customer shopping patterns and trends, improve the quality of customer service, achieve better customer retention and satisfaction, enhance goods consumption ratios, design more effective goods transportation and distribution policies, and reduce the cost of business. James et al. (2007) opinioned that many Indian firms have been heavily investing in IT for the transformation of their terabytes of data to help them to manage their business decisions more effectively and gain a competitive advantage. With the help of DM techniques, retailers can improve their inventory logistics and reduce their cost in handling inventory. They can identify the demographics of their customers such as gender, martial status, number of children, etc. and the products that they buy. This information can be extremely useful in stocking merchandise in new store locations as well as identifying more selling products in one demographic market that should also be displayed in stores with similar demographic characteristics. For nationwide retailers, this information can have a tremendous positive impact on their operations by decreasing inventory movement as well as placing inventory in locations where it is likely to sell (Wu, 2002). DM can also be used to locate individual customers with specific interests or determine the interests of a specific group of customers (Guzman, 2002). Only through the application of DM techniques can a large enterprise hope to turn the myriad records in its customer databases into some sort of coherent picture of its customers (Berry and Linoff, 2001). Baesens et al. (2009) expressed that the DM is increasingly playing a key role in decision making. Most retailers collect and have access to huge amount of data, collected from day to day operations e.g. customer loyalty data, retail store sales and merchandise data, demographic data etc. There is a great potential to develop systems that enable retailers to manage, explore, analyze, synthesize and present data in a meaningful manner for strategic decisions. Retail managers are in a constant need for right kind of information for making effective decisions (Sharma and Vyas, 2007). Retailers are making more use of data mining to decide which products to stock in particular stores(and even how to place them within a store), as well as to assess the effectiveness of promotions and coupons (Two Crows Corporation, 2005).

The retail industry has been shifted its focus from products to customers. Rather than pushing products and making sales, it has now become important to meet customers' needs and keeping customers satisfied. DM applications in the retail industry include applications to obtain insights into customer tastes, purchasing patterns, market share, site locations, patronage and targeting (Peterson, 2003), applications to manage inventory, promotions, margin control and negotiation with suppliers (Reid, 2003) and applications to increase returns from customer interactions, up-/cross-/down-selling efforts and multi-channel customer analysis (Fayyad, 2004).For example, the introduction of bar-code scanners and universal bar-coding has resulted in the accumulation of a wealth of data. Transactional data are now easily gathered at the point- of-sale. The use of credit cards and loyalty card programmes has allowed anonymous transactions to be linked with individual customers' purchases. So, the demographic data of the customer and transactional data can now be analyzed together to yield richer information on customers and their purchasing patterns.

6.2.1. Churn Modelling

Churn is a common phenomenon that occurs in retail industry. By churn we mean those customers, who will be leaving the retailer in the near future. If churn is predicted in advance then corrective actions can be taken so that churning can be minimized. Ju (2008) did the Research on the application of Customer Churn Analysis in Chain Retail Industry.
Customer churn refers to the original customer of companies terminate to purchase products or accept services, and turn to rivals (En, 2007). In churn modelling past data is used to predict future behaviour (i.e., churn). In the modelling stage, past monthly transactional data are available and it is possible to use data in and before a particular month to predict churn behaviour in the next month. In the deployment stage when the churn model is actually applied, it may be the case that for any particular month when churners are to be identified (i.e., predicted) for the month after, the latest data available are those one month before so that preemptive actions can be taken to prevent churn in the coming month. So, a realistic churn model will have to be one that uses data one month before to predict in the current month the potential churners in the next month (Chye, 2005). Hadden et al. (2007) addressed that much research has been invested into ways of identifying those customers who have a high risk of churning.

Retail industry intends to apply the data mining results on existing customers to identify those who exhibit the same behavior as the churners — especially profitable ones — so that actions can be taken to reinforce their loyalty before they are lured away by their competitors. The following predictive modelling tools are used to construct the potential chum models: decision trees (using the C5.0 and CART algorithms), neural networks and logistic regression.

A graphical representation of the decision tree model (using the construction data set is an excellent way to visualize the predictive modelling results and relationships between the input variables and target variable. Generally, input variables appearing higher up in the decision tree have a stronger association with the target variable and hence are more important for predicting churn (i.e. identifying potential churners).

7. Conclusion

Decision trees are the favored technique for building understandable models because of their tree structure and ability to generate rules. This clarity allow for more profit and Return-On-Investment models to be added easily in on top of the predictive models. There is no one model that is superior under all circumstances. This is especially so because different models can lead to different results depending on the actual data being mined. There is no doubt that DM is a very powerful methodology and technology that can be applied in many different commercial and non-commercial contexts. With some imagination and creativity, it can go a long way towards enhancing the competitive advantage of retail industry.

8. References

Ahmed, S. R. (2004), 'Applications of Data Mining in Retail Business', *Proceedings of the International Conference on Information Technology: Coding and Computing*, Vol.2, pp. 455- 459 IEEE.

Baesens, B., Mues C., Martens, D. and Vanthienen, J. (2009) forthcoming, '50 years of data mining and OR: upcoming trends and challenges', *Journal of the Operational Research Society*, Vol.60, pp .S16- S23 (1).

Berry, M. J. A. and Linoff, G. S. (2001*), Mastering Data Mining The art of Customer Relationship Management*, John Wiley & Sons, Inc.

Berson, A. and Smith, S. J. (2008), *Data Warehousing, Data Mining, & OLAP*, Tata McGraw-Hill.

Berman, B. and Evans, J. R. (2008), *Retail Management- A strategic Approach*, Person Publisher.

Bhasin, M. L. (2006), 'Data Mining: A Competitive Tool in the Banking and Retail Industries', *The Chartered Accountant*, October.

Chen, S. Y. and Liu, X. (2005), 'Data Mining from 1999 to 2004: an application -oriented review', *International Journal of Business Intelligence and Data Mining*, Vol.1, No.1, pp. 4-21.

Chen R. S., Wu R. C. and Chen J. Y. (2005), 'Data Mining Application in CRM of Credit Card businesses', *Computer Software and Applications Conference*, IEEE , Vol. 2, pp. 39 - 40.

Chen, Y. L. and Hung, L. T. H. (2009), 'Using decision trees to summarize associative classification rules' , *Expert Systems with Applications*, Vol. 36, Issue 2, Part 1, pp.2338-2351.

Chye, K. H. (2005), *Data Mining Applications for Small and Medium Enterprises*, *Centre for research on small enterprise development*, Nanyang technological University, Singapore.

Cowley, S. (2005), 'Data Mining', *IDG News Service*, New York Bureau.

D'Souza, R., Krasnodebski, M. and Abrahams, A. (2007), 'Implementation study: Using decision tree induction to discover profitable locations to sell pet insurance for a startup company', *Journal of Database Marketing & Customer Strategy Management*, Vol. 14, pp.281-288.

En, X.G.,'Study of customer churn based on business intelligence', Dr Thesis, 2007, Shanghai: Fudan University.

Fayyad,U. (2004), 'Optimizing customer insight', *Intelligent Enterprise*, Vol.6 No.8, pp.22-26,33.

Feelders, A., Daniels, H. and Holsheimer, M. (2000), 'Methodological and Practical Aspects of Data Mining', *Information and Management*, Vol. 37, Issue 5, pp.271-281.

FICCI Retail Report 2007, www.ficci.com (accessed on 25 th June 2008).

Gearj, A.E., Gillespiej, S. and Allen, M. (2007), 'Applications of decision trees to the evaluation of applied research projects', *Journal of Management Studies*, Blackwell Publishing Ltd, Vol. 9, Issue 2, pp. 172 – 181.

Guzman, I. (2002), 'A strategic Decision Support Tool for Organizations', *Strategic Management of Information Resources*, Research Paper.

Hanna, J. (2004), 'Ground-Floor Opportunities for Retail in India', *Harvard Business School Newsletter.*

Hadden, J., Tiwari, A., Roy, R. and Ruta, D. (2007), 'Computer Assisted Customer Churn Management: State-Of-The-Art and Future Trends', *Computers & Operations Research*, Vol.34, No.10, pp. 2902-2917.

Han, J. and Kamber, M. (2007), *Data Mining*, Morgan Kaufmann Publishers.

Hou, J-J. and Tu, H.H-J. (2008) 'Customer relationship management strategy and firm performance: an empirical study', *International Journal Electronic Customer Relationship Management*,Vol.2, No.4, pp.364-375.

James, E.R., Peter, C.T. and Sid, L.H. (2007), 'An examination of Customer Relationship Management (CRM) technology adoption and its impact on business-to-business customer relationships', *Total Quality Management and Business Excellence*, Vol. 18, No. 8, pp.927-945.

Jones, S. and Ranchhod, A. (2007) 'Marketing strategies through customer attention: beyond technology-enabled Customer Relationship Management', *International Journal Electronic Customer Relationship Management*, Vol. 1, No. 3, pp.279-286.

Ju, C., Guo, F.(2008), 'Research and Application of Customer Churn Analysis in Chain Retail Industry', International Symposium on Electronic Commerce and Security , IEEE, pp.670 – 673.

Kaur, P. and Singh, R. (2007), 'Uncovering retail shopping motives of Indian youth', *Young Consumers: Insight and Ideas for Responsible Marketers*, Vol. 8, No. 2, and pp.128-138.

Lee, S. J. and Siau, K. (2001), 'A review of data mining techniques', Industrial Management and Data System, Vol.101, No. 1, pp. 41-46.

Lemmens, A. and Croux, C. (2006), 'Bagging and Boosting Classification Trees to Predict Churn', *Journal of Marketing Research*, Vol. 43, Issue: 2, pp: 276-286.

Lima, E., Mues, C. and Baesens, B. (2009), 'Domain knowledge integration in data mining using decision tables: Case studies in churn prediction', *Journal of the Operational Research Society,Vol. 60, pp. 1096-1106.*

Noonan, J. (2000), 'Data Mining Strategies', *DM Review*.

Pan, J., Yang, Q., Yang, Y., Li, L., Li, L., Li, F., T. and Li, G., W. (2007), 'Cost-sensitive-data preprocessing for mining customer relationship management databases', *IEEE Intelligent Systems*, Vol.22, No 1, pp 46-51.

Pande, S. and Collins, T. (2007), 'Strategic implementation of information technology to improve retail supply chain in India', *International Journal of Logistics Systems and Management*, Vol. 3, No. 1, pp. 85-100.

Peterson, K. (2003), 'Mining the data at hand', *Chain Store Age*, Vol. 79, No. 6, pp.36.

Ranjan, J. and Bhatnagar, V. (2008), 'Data Mining tools: a CRM perspective', *International Journal Electronic Customer Relationship Management*, Vol. 2, No. 4, pp.315-331.

Ranjan, J., Goyal, D.P and Ahson, S.I. (2008), 'DM techniques for better decisions in human resource management systems', Int. J. of Business Information Systems, Vol. 3, No. 5, pp.464-481.

Rao, I. K. R. (2003), 'Data Mining and Clustering Techniques', DRTC Workshop on Semantic Web, December.

Reid, K. (2003), 'Digging into data', *National Petroleum New*, Vol. 95, No.8, pp 28-32.

Rogers, M. (2005), 'Customer strategy: observations from the trenches', *Journalof Marketing*, Vol. 69 No.4, pp.262.

Sangle, P.S. and Verma, S. (2008) 'Analysing the adoption of Customer Relationship Management in Indian service sector: an empirical study', *International Jour Relationship Management*, Vol. 2, No.1, pp.85-99.

Salford Systems (2009), Overview of CART,(Retrieved on 08-May-2009).

Sarantopoulos, G. (2003), 'Data mining in retail credit', *Operational Research*, Springer Berlin / Heidelberg, Vol. 3, No. 2, pp. 99-122.

Savage, S. L. (2003), *Decision Making with Insight*, 2nd ed., Brooks/Cole - Thompson Learning, Belmont, CA.

Sharma, A. and Vyas, P. (2007), 'DSS (Decision Support Systems) in Indian organised Retail Sector', Indian Institute of Management Ahmedabad (IIMA) Resear

=Sheu, J. J. Chang, Y. W. and Chu, K.T. (2008), Applying decision tree data mining for online group buying consumers' behaviour, *International Journal of Elect Relationship Management*, Vol. 2, No.2 , pp. 140- 157.

Sohoni, A. (2007), 'Indian Retailers-Ready for Take Off?'available from http://www.tech2.com/biz/india/features/retail/indian-retailers-ready-for-take-o august-2008).

Thearling, K., *Data Mining and Customer Relationships*, (Available from www.thearling. com) (Accessed on 22-August-2008).

Two Crows corporation, 'Introduction to Data Mining and Knowledge Discovery', available at http://www.twocrows.com/ (Accessed on 25/july/2008).

Vector, D. (2007), 'Indian Retail Industry: Strategies, Trends and Opportunities 2007', available at http://www. Marketresearch .com /product (accessed on 4th July

Velikova, M. and Daniels, H. (2004), Decision trees for monotone price models, *Computational Management Science*, Springer Berlin / Heidelberg, Vol. 1, No. 3- 4,

Wang, H. and Wang, S. (2007), 'Mining purchasing sequence data for online customer segmentation', *International Journal of Service Operations and Informatics*

Wang, J., Yuan, B. and Liu, W. (2008), 'Application of Decision Trees in Mining High-Value Credit Card Customers', *Proceedings of eleventh Joint Conference o* Advances in Intelligent System Research.

West, D. (2005), 'Enhancing Value through Data Mining: Insurers can use data mining technology to improve their competitive position', *Insurance Networking N for Technology Management*, October.

Wu, J. (2002), 'Business Intelligence: The Value in Mining Data', *DM Review online*, February.

Yang, X. C., Wu, J., Zhang, X. H. and Lu, T.J. (2008), 'Using decision tree and association rules to predict cross selling opportunities', *International Conference or Cybernetics*, IEEE Conference Proceedings, Vol. 3, Issue :12-15, pp.1807 – 1811.

Contact the Authors:

Jayanthi Ranjan, (Associate Professor-IT), Institute of Management Technology, Ghaziabad (U.P)-India: Tel: 91-120-3002200, Ext.219; Fax: 91-120-3002300; Er jranjan@imt.edu

Ruchi Agarwal, (Research Scholar-PhD), Birla Institute of Technology, Extn Center Noida, Mesra, Ranchi, India, Tel: 09350962983, Fax: 09412621023; Email: r

---