Critical Review of Data Mining Techniques for Insurance Service Operations

Girijesh Pathak
National Insurance Academy, Pune
(girijesh_kgp@hotmail.com)

Amar Nath Jha
Birla Institute of Technology, Mesra
(anjha06@rediffmail.com)

Data Mining Techniques have the potential of finding valuable patterns from the data even if they are hidden. These techniques are of high importance for Insurance Industry as they require data from almost all possible sources in their operations. At the data analysis level it depends on the managers to ask for the appropriate analysis to get valuable information from the data. Some valuable patterns get missed if they are not searched for. This paper discusses various aspects of Insurance Operation and identifies appropriate data mining techniques to reach to hidden valuable patterns to increase the effectiveness of operation.

1. Introduction

1.1 Insurance Service Operations

Insurance is one such service that extends its scope to every section of society – whether it be rich or poor, big corporate house or small individual entrepreneur, space research or agriculture, cosmopolitan city or remote rural area. Insurance service touches everyone.

Insurance service involves several operational activities. Fundamentally, Insurance is about charging premium to cover certain risks and pay claims in those cases where these covered risks happen. There can be several peripheral and central activities in such core insurance operations. Identifying the risks that can be covered, deciding about the premium to be charged, preparing on various clauses of the insurance agreement, executing the agreement, renewal, maintaining relationship with customers, evaluating adequacy of their insurance etc. are some of the activities that are common in both the situations claim or no claim. However, if claim happens, another set of activities related to claim handling like estimation of loss, evaluating the reasons, applying the clauses of the agreement, analysis of claim payment/ claim denial etc. are required not only as a part of claim management process but also as inputs to future insurance activities.

All these activities generate huge volume of data. These are internal data to insurance operations. However, some of the insurance activities also use data from other sources that are not internal to its operations. For example, it needs data from health system regarding various diseases in various lifestyle people, the kind of tests and treatments, costs of these and associated activities while designing a health insurance product. Analyzing the data to generate various reports that are used to keep control over operations and plan for the future are essential part of the operational activities. Additionally, these data are also mined to identify various valuable patterns and make use of them in fraud detection or prevention, generating business leads, identifying untapped potential market segment etc.

Insurance operation related activities also get linked with investment as the customers in many cases expect some financial return even in normal situations. That is why there are such life insurance products that provide death claim in case of death but also provide maturity claim in case of no death during the term plan. As such the Insurance organizations have to manage the finance well and integrate the risk management related activities as well as investment related activities to meet the aspirations of its customers and society.

1.2 Data Mining

Data Mining Techniques are such mathematical techniques that have great capabilities of finding hidden valuable patterns from the data. The main differentiator to differentiate data mining techniques from the data analysis techniques are in the capabilities of these techniques in reaching to the hidden patterns. So, data mining techniques explore all possibilities in the data and fetch the results according to the criterion for being valuable provided by the information seeker.

These techniques involve very high volume of computation even with small dataset. Latest advancements in computation technology has made it possible to do such intense volume of computation with large data that usually a business generates. That is why, these techniques are gaining high importance in business now.

This example explains the core difference between data mining and data analysis. Take for example, an Insurance organization has 15 different insurance products that an insurance seeker can take in various combinations to meet his insurance needs. Insurance company wants to improve its business and also create some kind of customers delight by suggesting the insurance seekers few products depending on their interests shown in various other insurance products. For example, the past data reveals that a large percent of insurers going for property insurance and motor insurance have also gone for health insurance. A customer has taken property insurance and motor insurance, hence the insurer suggests him to think about health insurance as well. This kind of association that a customer going for property and motor insurance is highly likely to go for health insurance as well, can be easily obtained from the data by counting the number of various customers taking property and motor insurance and among them taking the health insurance as well. But in this approach, the insurer has

ISBN: 978-1-943295-06-7

to identify such likely associations based on his intuition or experience and get that valuated from the data before using them. If he doesn't identify some associations but they exist very strongly as per the data, he will miss them and will not be able to take advantage of them in business. With 15 insurance products, 2^15 i.e. 32768 (more than 32 thousands) combinations of insurance products are possible. And, if an insurer looks for questions like 'if a customer opts for a combination of insurance products then he is also likely to opt for another combination of insurance products from the set of products being offered', the number of such question will be in millions with just 15 products. There might be strong associations existing, but the insurer may miss them if he has to keep asking such millions of questions.

Data mining techniques free the users from asking such questions. The algorithms used for mining association rules are such that they explore all combinations of products and wherever strong enough associations are found they are brought to the notice. The user of data mining technique has to specify what is strong enough for him. May be that he specifies minimum support value and minimum support probability and based on that the mining algorithm takes out all such associations where a combination of products purchased with minimum support frequency relates to purchase of another combination of products with a minimum support probability. Thus, these techniques are capable of finding hidden valuable patterns. Since the computation load is tremendously high in such application, the feasibility of these techniques depend largely on computational efficiency of the algorithms and the computation power available with the present technology. Fortunately, both have improved so much in recent past that various business organizations can now easily use them in their functioning.

Data mining techniques mainly comprise of various mathematical and statistical techniques to do some intelligent work with the data like intelligent prediction, classification, mining association, clustering, neural network etc. With the social digital media and virtual world becoming as powerful as real world, the sentiments and emotions expressed in these media too provide valuable knowledge and cannot be ignored. Data mining techniques are used even in unstructured data mining like text mining, sentiment analysis etc. and are capable of providing valuable trends from the digital world that can save or break any organization.

1.3 Challenges in Adopting Data Mining Techniques in Insurance Service Operations

If an industry has to be picked that requires data mining most then it is perhaps the insurance industry. Insurance industry has a long and rich history. It has been using data for its operations and various analytic requirements for quite long. Insurance professionals are quite used to data and analytics. This can be a boon or can be a big challenge in adoption of techniques like data mining techniques.

It may not be feasible for most of the insurance professionals to get to the algorithm level of data mining techniques. It is not expected as well. However, data mining techniques, if implemented in strategically right way, can change the way one looks into the data and can dramatically change the expectations of insurance professionals about usage of such techniques in their field.

Data mining techniques can be implemented with some software only. Since, it may require data from various functions of the organization and also from large number of external sources, it is implemented on data warehouse. The software enabling the data mining techniques need to be integral part of the core software used by the organization in the core operations. This is required to meet the requirements of discipline and control in the sensitive operation like insurance. Most of the enterprise level software (ERPs) have data mining or business intelligence module in-built in their software. Implementing these modules are the best way of adopting data mining techniques. However, implementation of such modules are extensive exercise involving almost all the functions of the organization. In most of the cases, these implementations are done by technology implementation partners. They gather the requirements by interacting with selected persons from various functions of the organization, document the requirements, get the sign-off and then map these requirements into the features available in the software. Some requirements may be met by customization in the software if the consultants are not able to identify the standard features of the software suitable to meet them.

Users mostly spell their requirements in terms of the existing practices and focus more on the difficulties faced in day-to-day operations. Since they are comfortable with the data analysis techniques and various reports they use, their requirements mostly move around them. Data mining or business intelligence module of the software are anyway capable of handling these data analytics requirements. Thus even after implementing data mining, most of the organizations continue using the same kind of data analytics and report that they were using earlier.

Hence, the real challenge in adopting data mining techniques in insurance service operations is to understand the capabilities of data mining techniques, brainstorm on various exciting usage of them and formulate the analytic requirements based on them. These requirements will be futuristic in nature hence merely gathering the analytic requirements will not help.

Also, once the software is implemented, the execution of these techniques get reduced to working with the menu of the software and selecting few commands from them. Output of some of the data mining techniques may look similar but they are technically quite different. For example, clustering of customers in 3 clusters will give 3 lists containing details of customers in each of the clusters. Also, classification of customers in 3 classes will give 3 lists containing details of customers in each of the classes. Classification and clustering are two different techniques and can have different impacts. It is important for the users to get at least one layer deep and understand at least the intuitive concepts involved in these techniques. This will help in formulating the requirements using the power of data mining and thus the real benefits of these techniques can be achieved by the organization.

2. Data Mining in Insurance Service Operations

Data mining techniques bring additional strength in such functions that can use data in innovative way. Those techniques of data mining that are doing this and are significantly different than the data analysis techniques of the previous decade are being discussed here with their possible usage in various functions of insurance service operations.

2.1 Clustering

These techniques cluster the entities in specified number of clusters based on some criteria technically called distance. So, the entities in a cluster are closer to each other compared to entities in different clusters. Clustering algorithms quantify the distance based on given criteria and makes clusters of them such that overall distance within clusters get minimized. The clustering algorithm may be complex. It is not desired that a user of clustering understands the algorithm. But if he knows the kind of inputs he has to provide and what they mean then he will be able to use these techniques quite advantageously. The user is required to provide the number of clusters in which the entities are to be clustered and criteria for measuring the distance or closeness among entities.

Since cluster is a kind of group based on some similarity, it can be understood that the entities in a cluster are likely to behave in somewhat similar way. Insurance organizations can apply these techniques to increase their penetration by designing insurance products to meet the insurance needs of various clusters of individuals or corporates. Even cluster focused policies and strategies can improve the reach of insurance significantly.

Formulation of marketing strategy in insurance operation can take advantage of clustering, cluster based product design can do risk management better, identifying new market place can be very effective based on clustering, distribution network in insurance operation can take important lead based on clustering, employee policies based on clustering can help in improving productivity, and so on. In fact, with the kind of knowledge and experience insurance professional have related to their field of operation, they can bring drastic improvement in every aspect of their operation with proper understanding and usage of clustering techniques.

Clustering algorithms are such that, the clusters of existing entities already clustered may change if new entities are brought in the system and clustered. Selecting number of clusters and the measuring criteria of distance play crucial role in the kind of experience one gets after implementing this technique. These are to be selected by the managers, hence effect of these on the overall outcome should be carefully evaluated by either experimenting with clustering before implementing in business or by using experience and intuition or by combination of both.

2.2 Classification

Classification techniques classify entities in different classes based on a set of predefined classification rules. An insurance company can cluster their customers in certain number of clusters based on their closeness with each other on some criterion. They can also classify the customers in certain number of classes based on some classification rules. Outcome of both the exercises will contain list of customers in different clusters or classes. Clusters are based on relative closeness of entities based on some criterion. It is a relative system. Whereas, in classification, a customer meeting the criterion to be classified in a particular class gets classified in that class irrespective of the characteristics of other customers. While forming classification rules, this point needs to be taken care. Computation part is done by the software, but forming a set of classification rules that is complete and consistent is a real task for managers.

This is a common tendency to pick large number of factors for classification thinking that larger the number of factors, more thorough the classification be. Similarly, in every factor, one tends to make large number of possible values. This makes the number of possible combinations too high. Take a case in which customers are to be classified based on their age, annual income, premium paid, claim history, number of policies taken, loyalty (number of years he has been a customer of the same insurer) and total sum assured. Say, age has five age groups, annual income has 6 income groups, premium paid has 8 premium ranges, claim history has 4 types, number of policies taken too have 4 policy ranges, loyalty has 3 possible value ranges and total sum assured has 8 value ranges. The number of combination of possible values of these factors will be equal to product of the value ranges in various factors. This is equal to 5*6*8*4*4*3*8 = 92160. So, this set of factors and possible values for classification has 92160 different combinations. Suppose the customers are being classified in 3 classes of Excellent, High valued and important customers – excellent being the best in that order. This too is a common tendency that to get a customer classified in the best class he needs to have the best values in the factors identified for classification. If, the classification rules are formed in similar way and even if all the possible value ranges of all the factors are included in some class, the classification rule remains badly incomplete. Including the 5+6+8+4+4+3+8 = 38 possible value ranges doesn't help in making the classification rule complete because there are 92160 value combinations and all should get classified in one class or other. Making use of multiplication rules is the key in forming a complete classification rule. But, this may lead to inconsistency if the same combination gets classified in more than one class. If this inconsistency is undesirable, proper considerations are required to handle that as well.

Having large number of factors and possible value ranges in them do not make the classification better in any way. It only adds to the complexity and makes it difficult to form a complete and consistent classification rule. Hence, it is a better idea to keep the number of factors small by picking only the important factors. And, also keep the number of value ranges less. Make range different only when it really makes a difference in the class of the entity. The number of combinations is a product and every combination should be classified in one and only one class. In case of customer classification, millions of customers are represented by these combinations. If any combination is missed in classification, thousands of customers represented by that combination will not get classified in any class.

Classification techniques can be used in insurance service operations for classifying various entities like customers, employees, agents, offices etc. in various classes based on some pre-formulated classification rules. These classes can be used in forming policies related to relationship management, various promotional activities, task assignment, grievance handling, human resource development etc.

2.3 Association

Predicting customer's likely behavior based on present behavior and generating some business lead has been a well-practiced concept for quite long. But recent technology has made it possible to automate this. A person listens to few songs and the system suggests him what he may like to listen next. Someone reading various news articles gets automated suggestion from the system about news of his interest that he may like to read. A customer purchasing few products gets suggestion of some other products in which he may be interested. A customer opting for few insurance products is suggested about what more he may be interested in based on his present options. These are some examples of association. These associations are found using some algorithms and it involves quite large volume of computation to mine the hidden valuable associations from the data.

The success of mining association rules depends largely on whether the predictions or suggestions turn true or not. If a person who has listened to two different songs is suggested that he may like to listen a particular song and he really likes that song then it is a success. But, if it is not liked then such association rules may turn out to be destructive to the business. Though sufficient number of incidences are required to reach to such conclusions, mined association rules should be used with caution in the beginning.

There is always some probability that a customer opting few insurance products will also opt for a set of other insurance products. This can be generalized saying that there is always some probability that a set of behavior will lead to another set of behavior. But what probability value (called threshold probability value) will be considered as high enough to consider the association strong and use that in business is an important decision users of association rules have to take. Similarly, suppose there are 10 incidences of customers behaving in a particular way and on 9 out of these the customer also behaved in another way. Should one say that the second kind of behavior is strongly associated with the first kind of behavior assuming that a probability of 0.9 is strong enough? Here, the question is on the number 10. Is 10 good enough, or it should be 100 or 1000 or something else (called threshold support value)? Selecting this number too is an important selection for the users of association rules. By properly selecting the threshold probability and threshold support value, one can get association rules that are effective.

Association can be used in insurance service operations to increase the business by providing well targeted suggestions, it can be used in creating innovative insurance product mix, it can be used in predicting a likely fraud, and preventing them and so on.

The three data mining techniques discussed here bring clear difference in implementation approach compared to the data analysis techniques. From the user point of view, application of clustering requires deciding about the number of clusters required and the criterion to be used for measuring nearness among the entities. This decision is more like a managerial and strategic decision than a mathematical decision. Similarly, classification technique makes user formulate classification rules that are complete and consistent. This again requires managerial and strategic thinking more than mathematical thinking. Mining association rules frees the users from asking questions to identify some strong associations in the data. Instead, users are required to specify the threshold probability that is a comfortable probability for his managerial style and the area of application. He is also needed to provide the threshold support value to mainly keep the confidence level in comfortable zone. Thus data mining techniques encourages strategic thinking in even identifying the kind of inputs they need to provide to mine the data.

3. Implementing Data Mining

Since data mining uses high volume of data and does somewhat rigorous mathematical computation on them, its implementation in insurance service operations has to be through implementation of some software. It is strategically important that the software to be used for data mining is a part of the major software being used in the operation. It can be one or few modules of the main software. However, implementation of such module or software doesn't implement data mining.
The reason is that implementation of data mining should be changing the way organization generates leads in various activities. They become more intelligent. These needs to be spelled out clearly through the requirement determination exercise. It is not merely automation of data analysis through another software. Hence, the core concepts of data mining techniques and the way intelligence application is eased in the operations needs to be understood and brainstormed. Then only such requirements can be determined that can use the real power of data mining techniques.

Data mining should not engage the database for its data requirements because they make the database so busy that operational activities needing data from database gets practically stopped. Data warehouse is the right data infrastructure for implementation of data mining. Design of the data warehouse involves identifying various data elements to be used in data

ISBN: 978-1-943295-06-7

mining and the sources from where they need to be fetched in the data warehouse. This too requires inputs from insurance professionals because they know the data requirements and their source for various operational activities better.

Thus, apart from the technical aspects of implementing data mining techniques through implementation of some software, preparation by the insurance professionals in going deep into the relevant concepts and their usages, and providing valuable inputs in requirement determination involving data mining and source of data identification etc. play vital role in successful use of data mining in insurance service operations.

4. References

Following are the main sources referred during the work on this paper.

1. Stuart Rose; 'How Analytics is Changing Insurance – and Why the Industry Continues to Lag'; Insurance Innovation Reporter; http://iireporter.com/how-analytics-is-changing-insurance-and-why-the-industry-continues-to-lag/.
2. Aggarwal Charu C; 'Data Mining – The Textbook'; Springer 2015; p 93-152.
3. Inna Kolyshkina, Richard Brookes; 'Data Mining Approaches to Modelling Insurance Risk'; PricewaterhouseCoopers; https://pdfs.semanticscholar.org/47d7/d393bc5e9309204dae025efb8d8f1bb1b113.pdf; 22nd October 2002.
4. Christine Fettinger; 'Challenges and Opportunities in Data Mining in the Insurance Industry'; Cyber streams; https://www.cyberstreams.com/challenges-and-opportunities-in-data-mining-in-the-insurance-industry/.
5. Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman; 'Mining Of Massive Datasets'; Cambridge University Press; 2nd edition; Kindle book.
6. Umamaheswari K., Janakiraman S.; 'Role of Data Mining in Insurance Industry'; COMPUSOFT, An international journal of advanced computer technology; June 2014; (Volume III, Issue VI); p961-966.
7. Sumathi S, Sivanandam S N; 'Introduction to Data Mining and its Applications'; Springer, 2006; p 476-485.
8. Pathak Girijesh; 'Use Business Intelligence in Retail Insurance Marketing'; http://ezinearticles.com/?Use-Business-Intelligence-in-Retail-Insurance-Marketing&id=2797419.