Evaluation of Data Mining applications in Insurance Sector

Mrs Sunita D.Patil
Yashaswi Education Society, Pune

Dr. Smita Chavan, PhD
Associate Director NBN Sinhgad School of Computer Studies, Pune, Maharashtra

Abstract - **In the global era, Insurance systems rapidly a lot of tremendous development in our society. Due to the increased stress in day-to-day life, the growth of demand of insurance increased. Data mining helps insurance sectors to finding useful patterns from the customer database. Its aim is to retain valuable customers to exploit the profit of a company. Thus, purpose of the paper aims to present how data mining is useful in the Insurance Sector, how its techniques produce good results for extract information in Insurance Sector and how data mining enhance in decision making using insurance data, The conceptual paper is written based on secondary study, observation from various journals, magazines, White paper and reports.**

Keywords: **Applications, Benefits, Data mining, Insurance Sector.**

I. INTRODUCTION

Data mining is a process which finds useful pattern from large amount of data and to help the insurance firms for acquiring new customers, retaining existing customers and taking critical business decisions. The Insurance Sector is mainly dependent on customer's base. Every insurance sector is necessary to maintain effective management of customer data with the help of data mining techniques like customer demographic, transaction history etc. Data mining is help to manager and marketing professional extracting information from the database for to take decision and to predict information. In that situation, Data mining very helpful to the Insurance Sector for access the data easily.

This paper is organized as follows, first summarized role of Data Mining in Insurance Sector. Secondly, describes Data Mining Process and its trends and Challenges in Insurance Sector. Third, how data mining techniques used for Insurance Sector are summarized, finally the conclusion of the study is described.

II. LITERATURE REVIEW

This section author explained the background of the data mining techniques in the insurance domain. Lijia Guo et al., [3] focused on property/casual insurance and illustrated the data mining tools and techniques which related for the process of property insurance. T. L. Oshini Goonetilleke et al., [14] reviewed the analysis of customer analysis when mining a life insurance data, determined on the customer retention and executed the tools and techniques for customer attrition also introduce suggested problem of attrition analysis of life insurance domain was fruitfully removed with the execution of data mining techniques. Katharina Morik et al., [4] talked about

the customer churn management using insurance dataset. He identified churn customer how to minimize the churn in the Insurance Sector. A. B. Devale et al.,[6] listed about the strength and weekness, application of data mining techniques in life insurance, he explains how the data mining methodologies useful for the insurance firms for extract information from the customer database. Nanthawadee Sucharittham et al., [16] summarized the review of application of data mining method of life insurance domain. Rekha Bhowmik et al., [17] detected the auto insurance fraud using the various fraud anomaly detection techniques and focused to make out the behaviour of customer and identified profitable customers for the insurance company. S. Balaji et al., [7] extracted the classification techniques used for the prediction of data over life insurance customer data. He introduced the various algorithms .H. Lookman Sithic et al., [11] analyzed the various review for the fraud detection. He planned the number of research paper correlated to insurance fraud activities. Jayanthi Ranjan [20].

III. OBJECTIVES OF THE STUDY

1. To study the role of Data Mining in Insurance Sector.
2. To review of Data mining process in Insurance Sector.
3. To identify the prospective Data Mining techniques that can be used in Insurance Sector.
4. To explore the future trend and Challenges of Data Mining in Insurance Sector.

IV. ROLE OF DATA MINIING IN INSURANCE SECTOR

Data mining can be used for extracting hidden patterns from the large database and data warehouse; it is defined as the process of exploring and analysis for large amount of data with a specific target. In the Insurance Sector, data mining can help insurance firms to gain business advantage. For example, by applying data mining techniques acquiring new customers, retaining existing customers, predict churn customers, fraud detection etc. In Insurance sector data mining applied for identified risk factor, fraud detection, customer retention .It helps to manager and marketing professional to support to take critical decision with the help of exploit data about customers' buying patterns and churn customer behaviour and insurers increase customer satisfaction, increase product sales.

The data mining applications in insurance industry are listed below:

- Acquiring new customers.

- Retaining existing customer
- sales analysis
- Financial analysis
  - Detecting fraud

In Insurance sector customer retention, liability prediction applications and loyalty enhancement program are equally important, Data mining is applied in Insurance Sector lately but brought wonderful competitive advantages to the sector who have implemented it successfully. The data mining applications in Insurance Sector are listed below:

- Data mining is applied in claims analysis such as identifying which medical procedures are claimed together.
- Data mining enables to forecasts which customers will potentially purchase new policies.
- Data mining allows insurance companies to detect risky customers' behaviour patterns.
- Data mining helps detect fraudulent behaviour.[24]

Challenges

In insurance sector, handing large quantities of data during data mining has some of the challenges.[27] Data mining systems faces a lot of problems and when manage customer's data such as

- Noisy data
- Missing data.
- Hybrid one or more techniques
- Corrupted values
- Missing attribute values

One of the biggest challenges that insurance faces is improve the customer retention and higher revenue.

- What Service and benefits would current customer likely desire
- Which cluster or segment of a customer likely to be?.
- Which product would best suit as per customer requirement?.
- What is the risk rating of customer related to the amount of assurance?
- Purchasing behaviour of customers.

Future Trends

Data mining creates new challenges and opportunities in insurance sector so that can be helpful for manger and marketing professional to take a decision for improving business growth.[25]

- Standardization of data mining language
- Predict customer responses
- Customer profitability.
- Semantic and Image mining

V. DATA MINING PROCESS

The growth of the insurance industry entirely depends on the ability of converting data in to the knowledge by analyzing large amounts of data for prediction necessary to accomplish different phases of data mining process , It depends on the conclusion of each phase or which particular task of a phase, has to be performed next. Let's discuss the cross industry standard process for data mining.

5.1 Business understanding

1. This phase is required to understand what the business need and try to understand business objectives clearly. have to assess the current situation by finding about the resources, assumptions, constraints
2. Create data mining goals to achieve the business objectives within the current situation. Finally, a good data mining plan to establish to achieve business and data mining goals.
- Does the tool have a proven record of solving a wide range of business problems, including the problems that we face?
- How does the tool provide a bridge between business understanding and the technical aspect of data mining?

5.2 Data understanding

This phase depends upon initial data collection from available data sources according to goals. Its performed data load and data integration to make data collection successfully. Data quality examined by remove noise data, normalize data and, missing values.

- How does the tool preserve my existing investment in IT infrastructure?
- How does the tool enable interactive exploration and visualization of the data?

5.3 Data preparation

In this phase introduced complete all activities for construct the final data set. This task includes:

Selecting relevant data: attribute selection (filtering and wrapper methods), removing anomalies, or eliminating duplicate records.

Reducing data: sampling or instance selection.

Recover incomplete data: filling the values missed, or reducing ambiguity.

Purify data: correcting errors, or removing outliers (unusual or exceptional values).

Resolve data conflicts: using domain knowledge or expert decision to settle discrepancy.

- How does the tool prepare data?
- Can the tool automatically extract data for preparation?

5.4 Modelling:

In this phase modelling refer to a group of processes in which multiple sets of data are combined and analyzed goal of data modelling is to use past data to inform future efforts to uncover relationships or patterns. Data modelling involves various types such as statistical Regression, genetic programming, structural equation modelling.

The goal of Statistical Regression is to be describing the relationship between a dependent variable and independent variable.   Genetic   Programming

Modelling, Structural Equation Modelling generally used to confirm models already hypothesized by other means .It allows for the study of complex relationships among variables. Unique in that it allows for inclusion of latent variables – not directly observable or measurable is often limited by the imagination and time constraints of the modeller – what variables to include, what combinations to create?

- How does the tool boost analyst productivity?
- Does the tool offer a wide range of techniques?
- Does the tool enable the combination of techniques?
- Does the tool preserve my existing investment in technology such as algorithms and other tools?

5.5 *Evaluation*:

In this phase you have built models that appear to have high quality from a data analysis perspective. Review of each phase before proceeding evaluation of model. A key objective is to determine if an important business issue has not been sufficiently considered. The checklist items for the evaluation phase relate to how well input from business users have been incorporated into the model

- Does the tool achieve consistently high results?
- Does the tool provide results that are easy to understand? Strange
- Can the full range of visualization be applied to validate the results of a model?

VI. *DATA MINING TECHNIQUES TO INSURANCE DATA*

Data mining	Techniques Patterns
Clustering	Customer having similar characteristics Analysis of customer attrition in Insurance Sector Policy most likely to be used, most unlikely to be used Segments related to policy [8]
Classification & Prediction	Predicting consumer behaviour Predicting the likelihood of success of policies Classifying the historical customer records [17] Prediction of what type of policy most likely to be retained, most likely to be left Predicting insurance product behaviour and attitude Predicting the performance progress of segments throughout the performance period Prediction to find what factors will attract new avenues in Insurance Sector Classify trends of movements through the organization for successful/unsuccessful customer historical records[7]
Association	Discovery of such association that promotes business technique[18]
Summarization	Provides summary information Various multidimensional summary

reports Statistical summary information

This technique which used for report generation provides better decision making for large volume of customer database with the help of visualization tools. It will provide more functionality in business decision making for manager and marketing professionals. For solving the business problems and making decision, this data mining techniques can be help to the organization but selecting the appropriate techniques can important for the organization.[6]

Application and advantages of data mining in Insurance Sector:

1. Customer perspective
2. Company perspective
3. Administrative perspective.

VII. *CONCLUSION*

In this study Data mining help to make use of data mining process in a Insurance Sector .We hope that the information generated after the implementation of Data mining concepts helpful for manager and marketing professional in Insurance Sector for increased revenue and efficient processes and to faced future challenges. This study may be useful for managing customer data and improve insurance advisor's performance; reduce failing ratio by taking appropriate steps at right time to improve the quantity of customers. The primary aim of this paper is to provide a big picture for contributors to help them determine the potential research points and areas. This paper also can be of use for business managers who are planning to implement a data mining intervention in Insurance Sector.

REFERENCES

1. Han and Kamber, "Data Mining: Concepts and Techniques",Second Morgan Kaufman Publisher, 2006,
2. Micheal J. A .Berry and Gordon S.linoff. EBook "Data Mining Technique for marketing Sales and CRM", Wiley Publishing, Inc., Indianapolis. Indiana, 2004.
3. Lijia Guo. "Applying Data mining Techniques in Property/Casualty Insurance", Casualty Actuarial Society Forum Casualty Actuarial Society - Arlington, Virginia Winter 2003.
4. Katharina Morik and Hanna Kopcke. "Analyzing Customer Churn in Insurance Data". Lecture Notes in Computer Science, vol. 3202, 2004, pp. 325-336 .
5. Keyvan Vahidy Rodpysh,, Amir Aghai and Meysam Majdi ," Applying data mining techniques for Customer Relationship management", International Journal of Information Technology, Control and Automation Vol.2, No.3, July 2012
6. A. B. Devale and Dr. R. V. Kulkarni, "Application of data mining techniques in life insurance", International journal of Data mining and Knowledge management Process", Vol.2, No.12,.pp. 31-40. July 2012
7. S. Balaji and Dr. K. Srinivasta, "Naïve Bayes Classification Approach for Mining Life Insurance Databases for Effective Prediction of Customer Preferences over Life Insurance Products" , International journal of Computer Applications, vol.51, No.3, pp.22-26. August 2012
8. Kanwal garg, Dharminder kumar and m.c.garg, ",Data mining techniques for identifying the customer behavior of investment in life Insurance Sector in india", International journal of information technology and knowledge management, Vol.1, pp.51-56.
9. Cheng C.H. and Chen Y.S. (2009), "Classifying the segmentation of customer value via RFM model and RS theory", Expert Systems with Applications, Vol.36, pp. 4176-4184.
10. Gulshan Vohra and Dr. Bhart Bhushan " Data mining techniques for analyzing the investment behavior of customers in life Insurance Sector",

International Journal of Research in Engineering & Applied Sciences, Vo 2, Issue 1 January 2012.

11. H.Lookman Sithic, T.Balasubramanian, "Survey of Insurance Fraud Detection Using Data Mining Techniques", International Journal of Innovative Technology and Exploring Engineering, Vol-2, Issue-3, February 2013.

12. Kuo-Chung Lin , Ching-Long Yeh," Use of Data Mining Techniques to Detect Medical Fraud in Health Insurance", International Journal of Engineering and Technology Innovation, vol. 2, no. 2, pp. 126-137, 2012

13. Derrig R. A., "Insurance fraud," Journal of Risk and Insurance, vol. 69.3, pp. 271-287, 2002.

14. T. L. Oshini Goonetilleke and H. A. Caldera ," Mining Life Insurance Data for Customer Attrition Analysis ",Journal of Industrial and Intelligent Information Vol. 1, No. 1, March 2013

15. Gabor Melli , Osmar R. Zaïane ,Brendan Kitts, "Introduction to the Special Issue on Successful Real-World Data Mining Applications ", SIGKDD Explorations Vol 8, Issue 1.

16. Nanthawadee Sucharittham, ,Thanaruk heeramunkong Choochart Haruechaiyasak, Bao Tu Ho, Dam Hieu Chi, "Data Mining for Life Insurance Knowledge Extraction: A Survey".

17. Rekha Bhowmik, "Detecting Auto Insurance Fraud by Data Mining Techniques", Journal of Emerging Trends in Computing and Information Sciences , Volume 2 No.4, Aprill 2011

18. Abdhesh Gupta Anwiti,"Life Insurance Recommender System Based on Association Rule Mining and Dual Clustering Method for Solving Cold start problem ",International Journal of Advanced Research in Computer Science and Software Engineering , Vol 3, Issue 10, October 2013.

19. Dilbag singh,Pradeep Kumar, "Conceptual Mapping of Insurance Risk Management to Data Mining", International Journal of Computer Applications, Vol. 39, No.2, pp.13-18 2012.

20. Jayanthi Ranjan, Raghuvir Singh, "Analytical customer relationship management in Insurance Sector using data mining: a case study of Indian insurance company", International Journal of Networking and Virtual Organisations,Vol. 9, No. 4, pp. 331-366 .2011.

21. Xu Zhikun1, Wang Yanwen2 and Liu Zhaohui, "Optional Insurance Compensation Rate Selection and Evaluation in Financial Institutions ", International Journal of u- and e- Service, Science and Technology,Vol.7, No.1 , pp.233-242 ,2014

22. Miguéis V.L., Camanho A.S., João Falcão e Cunha (2012), "Customer data mining for lifestyle segmentation", Expert Systems with Applications,Vol.39, pp. 9359-9366.

23. M. Venkatesh "A Study Of Trend Analysis In Insurance Sector In India ",International Journal Of Engineering And Science (IJES) ,Volume 2 ,Issue 6,pp 01-05, June 2013.

24. Mrs.Bharati M. Ramageri, "Data mining techniques and Applications", Indian journal of computer science & engineering, Vol1, No.4, pp.301-305.

25. Jayanthi Ranjan " Data mining in pharma sector : benefits", IJHCQA, Vol. 22 No. 1, 2009, pp. 82-92.

26. Ashok Goudar, "Predictive Analytics in insurance Servies", MPHASIS, Hp company.

27. "Insurance Sector, challenges reforms and allignement", Confederation of Indian Industry