Role of Data mining in Insurance Industry

K. Umamaheswari[1], Dr. S. Janakiraman[2]
[1]Research scholar, Dept of Computer Science Bharathiar university, Coimbatore, Tamilnadu, India
[2]Assistant Professor, Dept of Banking technology, Pondicherry university, Pondicherry, India

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 firms to discovery useful patterns from the customer database. The purpose of the paper aims to present how data mining is useful in the insurance industry, how its techniques produce good results 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 and reports.

Keywords : Applications, benefits, CRM, Data mining, Insurance industry,

I. INTRODUCTION

Data mining is an interdisciplinary field of astronomy, business, and computer science, economic and other to discover new patterns from large datasets. Data mining technology can help the insurance firms for taking crucial business decisions. The insurance sector is primarily dependent on customer's base. The most scenario of any insurance firm, effective management of customer data is essential one. With the help of data mining techniques, the customer data handled effectively. Data mining is to help the market specialists for decision making process. Companies in the insurance industry collected huge amounts of data about their customers. Due to their protective regulations, extracting information from the database caused a lot of time. In that situation, Data mining very helpful to the firm for access the data easily.

A. Insurance sector

The insurance broadly divided into life, health and non-life insurance. The insurance companies have vital role in insurance provides which meet the requirements of the customers at the same time are affordable. "Insurance is the equitable transfer of the risk of a loss, from one entity to another in exchange for payment. It against the risk of a contingent, uncertain loss". Growing interest towards insurance among people, innovative products and distribution channels are sustaining the growth of the insurance sector

The Insurance Regulatory and Development Authority (IRDA) were constituted as an autonomous body to regulate and develop the insurance industry.

The key objectives of the IRDA include promotion of competition so as to enhance customer satisfaction and ensure the financial security of the insurance market. Various types of insurance such as Life insurance, Property insurance, health insurance, vehicle insurance and other insurance (Travel, Liability, Credit insurance) used the data mining technique depends upon their requirements. Today, insurance sector may gain valuable information that helps to other profitable policies to the customer community.

B. Data Mining

Data mining refers to extracting or "mining" knowledge from large amount of data. Data mining as a synonym for another popularly used term, knowledge discovery from data or "KDD" The goal of this technique is to find pattern that was previously unknown data [1].

The steps of knowledge discovery process as discussed as follows,

Selection: Selecting data relevant to the analysis task from the database.

Preprocessing: Removing noise and inconsistent data, combining multiple data sources.

Transformation: Transforming data into appropriate forms to perform data mining

Data mining : Choosing a data mining algorithm which is appropriate to pattern in the data, extracting data patterns.

Interpretation/ Evaluation: Interpreting the patterns into knowledge by removing redundant or irrelevant patterns. Translating the useful patterns into terms that human understandable.

This paper is organized as follows, first summarized about the background of the study. Secondly, describes about the data mining and its tasks. Third, how data mining techniques used for insurance industry are summarized, finally the conclusion of the study is described.

II. RELATED WORK

Literature survey is an essential for any study. This section described the background work of the data mining techniques in the insurance domain. Lijia Guo et al., [3] focused on property/casual insurance and described the data mining techniques which used for the process of property insurance. T. L. Oshini Goonetilleke et al., [14] summarized the analysis of customer analysis when mining a life insurance data. He focused on the customer retention and implemented the techniques for customer attrition. That recommended that problem of attrition analysis of life insurance domain was successfully removed with the implementation of data mining techniques. Katharina Morik et al., [4] discussed the customer churn management using insurance dataset. He focused on customer churn how to minimize the churn in the insurance industry. A. B. Devale et al.,[6] listed about the application of data mining techniques in life insurance, he explains how the data mining methodologies useful for the insurance firms.

Nanthawadee Sucharittham et al., [16] summarized the survey of application of data mining techniques of life insurance domain. Rekha Bhowmik et al., [17] detected the auto insurance fraud using the various fraud anomaly detection technique. And focused to identify the behavior of customer and analyzed the profitable customers for the insurance company. S. Balaji et al., [7] summarized the classification techniques used for the prediction of data over life insurance customer data. He analyzed the various classifier algorithms (Naïve Bayes, Bayesian Network classifier etc). H. Lookman Sithic et al., [11] analyzed the various review for the fraud detection. He listed the number of research paper related to insurance fraud activities. Jayanthi Ranjan [20] summarized the applications of customer relationship management in insurance company. She done the case study with insurance data

III. ROLE OF DATA MINIING IN INSURANCE INDUSTRY

Data mining is a powerful new technology with great potential to help insurance firms focus on the most important information in the data they have collected about the behavior of their customers and potential customers. Data mining assist insurance sector in predicting fraudulent claims & medical coverage and predicting the customer's pattern which customer will buy new policies. Data mining is applied in insurance industry, but tremendous competitive advantages the companies who have implemented it successfully.

Some of the areas data mining can be applied to Insurance industry are

- Identify risk factors that predict profits, claims and losses.
- Customer level Analysis
- Marketing and sales analysis
- Developing new product lines
- Reinsurance
- Financial analysis
- Estimating outstanding claims provision
- Detecting fraud

With rapid changes taking place in the field of insurance industry, decision support system plays an important role. Data mining used to support the controlling of policies, the administrative and management tasks, efficient management of organization and financial data.

A. Data Mining Techniques

Data mining techniques have applied to various insurance domains to improve decision making. Data mining use predictive modeling, market segmentation, market basket analysis to answer business questions with greater accuracy. Various data mining techniques used for the insurance industry development are Classification, Clustering, Regression and Association rules, summarization used for knowledge discovery from database. [25]

Classification

It is one of the most commonly used techniques, to develop models that can population records at large. The classifier training algorithm uses this technique used for business development. Various classifiers used for the classification algorithms such as Decision tree, Bayesian classifier, neural network, Support vector machine etc. Customer database can be segmented into homogeneous groups, classification maps data into predefined group into segments. The application of data mining classification algorithms are applied on insurance benchmark dataset. Types of classification are

- Supervised classification
- Unsupervised classification

Clustering

It used for identification of similar classes of objects. It's used for grouping based on the customer's behavior. It is applicable for customer segmentation and targeted marketing. Types of Clustering are

- Partitioning methods

- Hierarchical agglomerative methods
- Density based methods
- Grid based methods
- Model based methods

Regression

It can be used for prediction. Regression analysis used to model the relationship between one or more independent and dependent variables. In insurance firm, more complex techniques needed to predict future values. Types of regression includes

- Linear regression
- Non-linear regression
- Multi-variate linear regression
- Multi-variate non linear regression

Association

Insurance companies faced a lot of problems on customer retention. Association used for this task, because it finds all the association where customers bought a frequent item set. Association helps business firms to make certain decisions. Market basket analysis and cross selling programs are typical examples for which association modeling is usually adopted.

When the customers want to insure some policy then this technique helps us finding the associations between different items. Types of Association rules are

- Multi-variate association rule
- Multi-dimensional association rule

Table 1 shows how a how data mining techniques will add value 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
Classification & Prediction	Predicting consumer behavior Predicting the likelihood of success of policies Classifying the historical customer records Prediction of what type of policy most likely to be retained, most likely to be left
	Predicting insurance product behavior 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
Association	discovery of such association that promotes business technique
Summarization	provides summary information Various multidimensional summary reports Statistical summary information

Summarization:

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 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.

B. Data mining tasks

Data mining is becoming common in both the insurance sectors like private and public. Data of the customer are one of the most valuable assets of any firm. The traditional methods, which were used for handling huge amounts of data generated by insurance transactions, are too complex. For transferring huge amount of data for decision making, data mining makes the methodology. Insurance firms use the data mining methodologies to enhance research and increase sales among the customers.The data mining used for various tasks in the insurance sector as follows.

Acquiring new customers:

Acquisition of new customer is most important scenario of any firm. Traditionally, the insurance companies used the services of brokers to acquire the customers, but today a lot of ways helps to acquire the new customers [8]. Insurance firm focused of both acquiring new customer & retaining existing ones.

Cluster Analysis used in the private sector to identify target group of customers. It involves targeting the population who are most likely to become customers or most profitable to the company.

Customer level analysis

Analysis of customer purchase patterns and behavior. Using associated discovery technique, most insurance firms accurately select which policies and services to offer which customers [8]. According [21], it used data mining technology for insurance settlement and analyzed the customer records and also developed function structure model for customer analysis using data mining method.

Customer Segmentation

Segment based products for targeting the customers. Data mining can be used for customer segmentation, for promoting the cross-selling of services, and in increasing customer retention. Customers are assigned to lifestyle segment based on their purchase history. Market segmentation is the key issue for the development of loyal relationships among the customers [22]

Policy designing and policy selection

The insurance firm made the investigation whether people tend to purchase policy for the reason and the policy designed. In that case, to compete successful in the market , insurance companies used data mining technologies.

Prediction

Data mining used for variety of applications such as predicting and classifying customer's and clustering customer characteristics for achievement of profitability. From the customer point of view, predictive analytics provides some benefits such as simplified claim handling process, reduced policy premium for low risk customers, faster and automated claim settlement.[26] . Data mining tools predict behaviors and future trends, allowing businesses to make proactive, knowledge-driven decisions. It performs inference on the current data (insurance dataset) order to make predictions [7]

Claims management

It is one of the most function is the insurance; data mining handled the claims management function such as claim analysis and fraud analysis.

Developing new product lines

To develop the new product/plan depends on the customer needs. Insurance firms utilized all of their available information to better develop new product and marketing campaigns [8].

Underwriting and Policy management

Data mining can be used in this application to optimize the function of the insurance value chain (Premium Analysis and Loss analysis)

Risk management

One of the main stages in the process of risk management is risk financing is, of course, insurance. Insurance industry is keen in identifying the risks pertaining to their business.
Risk management contains the six phases are risk identification, risk analysis, risk prioritization and risk monitoring [19].

Reinsurance:

Reinsurance comes under in the fields of risk management. The reinsurer may be either a specialist reinsurance company, which only undertakes reinsurance business, or another insurance company. Data mining tools can develop predictive models to arrive at the reinsurance level for the book of business based on the historical claims data. These predictive models can be identified suitable policies for reinsurance based on the loss experience of similar policies in past.

Fraud detection

Detecting fraud claims is important in the insurance firm. Data mining isolates the factors that lead to fraud waste and abuse. To identify which transactions are most likely to be fraudulent. This is called as Fraud anomaly detection. In medical insurance, various medical insurance agencies suffered due to fraud claim in the health insurance; here he developed a model with three steps for the health insurance fraud detection. And he discussed the characteristics of fraud detection are high claims payment data is incorrect, suspicious data analysis, problem of hospital or physician [12]. The types of fraud and how much of fraud activities in the insurance firm discussed and he developed a claim sorting algorithm for the claim processing systems [13].

Trend analysis

Trend analysis often refers to the science of studying changes in social patterns, including fashion, technology, and consumer behavior. In insurance, to reveal difference between the typical customer this month and last [23]. Data mining used in the different service industry especially in insurance firms the most frequently used applications for customer segmentation, customer retention, risk assessment and fraud detection and Policy approval process.
Insurance firms have a lot of improved changes in Information technology. Insurance industry has

historically been a growing industry. It plays an important role in insuring the economic well being one country. Many insurers use data mining techniques to identify the new customers and other are applying this technique to reduce portfolio risk, and to identify policies that were based on fraudulent information

C. Challenges
In insurance organizations, processing large quantities of data during data mining has some of the challenges. Data mining system faces a lot of problems and pitfalls when handling customer's data such as

- Noisy data
- High volume and high complexity for different kinds of 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.

D. Future Trends
Due to technology increased, future application areas also creates new challenges and opportunities for data mining. Advanced data mining techniques can be developed and used by the insurance firm that helps business development to ensure the growth and development of insurance companies.

Future data mining technologies involves [25]

- Standardization of data mining language
- Predictive analysis
- Advanced Text mining
- Semantic and Image mining

IV. CONCLUSION
Data mining takes a vital role in the entire applications domain especially in insurance sector.It highlighted the importance and role of data mining techniques are useful in insurance sector for managing the customer data and gain business advantage. It can enhance an insurance business process. For business decision making, data mining techniques are implemented in this domain. This paper discusses how insurance firms can benefit by data mining methodologies and thereby reduce costs, increase profits, acquire new customers, retain current customers and develop new products. We can conclude that Insurance sector is having increasing growth rate and adopted with various data mining techniques. Insurance firms use the proper data mining techniques prove to be a boon for the

organization. The future of the insurance lies in the increasing the product plans and improving service levels using advanced data mining techniques.

V. 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 industry 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 Industry, challenges reforms and allignement", Confederation of Indian Industry.