

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 6, Special Issue 12, July 2017

Data Mining on Social Security and Welfares

Shiddalingappa Kadakol¹, Pooja Patil², Shriprada Laxmeshwar³

KLE's BCA Karnatak University, India

ABSTRACT: Data mining is the process of analysing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analysing data. It allows users to analyse data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. Data mining has attracted a great deal of attention in the information industry and in society as a whole in recent years, due to the wide availability of huge amounts of data and the immediate need for turning such data into useful information and knowledge. The government faces new and intense pressure to collect and use personal data. The need for a deep understanding of public and public government interactions through advanced data analytics has been increasingly recognized by the community at large. The need for a deep understanding of customers and customer government interactions through advanced data analytics has been increasingly recognized by the community at large. So far, however, no substantial work on the mining of social security and social welfare data has been reported.

KEYWORDS: Data Mining, Types of Data Mining, Social Security and welfares.

I. INTRODUCTION

Data mining it is the process of analysing data from various perspectives and summarizing it into information that is helpful. The issues in data mining are noisy data, interestingness, static data, sparse data, and dynamic data, and algorithm efficiency, relevance, missing values, heterogeneity, size and complexity of the data. Such problems often occur in large amounts of data. Data mining is having a large number of applications. Advancements in the Machine Learning, Statistics, Artificial Intelligence, Computation capabilities and Pattern recognition have given the present days data mining functionality a new range. The important application areas of the data mining are fight against terrorism, web and semantic web, business trends, fraud detection and management, scientific applications Bio-informatics and cure for the diseases, etc. There are many application areas. But in this seminar consider the importance of the data mining on social security and social welfare data. Because everyone as a member of society has the right to social security. Now a days the importance of social security and welfare business has been increasingly recognized in more and more countries. In our country itself, most of our personal data should be kept secured. Australia is one of the most developed social welfare countries in the world in terms of government policies, infrastructure, the population of benefit recipients, and the advancement of social security techniques and tools. Since 2004, we have been engaged in conducting data mining for the Australian Commonwealth Government through a series of projects. We have developed models, algorithms, and systems to identify key drivers, factors, patterns, and exceptions, indicating high risk of customers, customer circumstance changes, declarations, and interactions between customers and government officers.

II. TARGET AUDIENCES

This paper is targeted at researchers and practitioners who have interest in data mining, especially in the area of social security and welfare.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 6, Special Issue 12, July 2017

III. SEMINAR OUTLINE

The seminar is designed into 4 kinds. First one is the review on social security or welfare research and then next going from Social security services and data to Framework of social security data mining. And last the varying types used for data mining.

IV. REVIEWS ON SOCIAL SECURITY/WELFARE

Research

In the review on social security and social welfare data mining includes the analysis of data collection. The data mining is an important task in the knowledge of discovery process [3]. To performing the data mining of data will be collected from different perspectives. In which some are more important than others. The technical perspective data collection has an important role in this era. After the performance of the data mining some observations are made. There are various perspectives for data collection. By collecting data in these perspectives it is easy for data selection. The different perspectives are Political Perspective, Economic Perspective, Sociological Perspective, Regional Perspective and Technical Perspective. There are various perspectives for data collection. By collecting data in these perspectives it is easy for data selection. The different perspectives are Political Perspective, Economic Perspective, Sociological Perspective, Regional Perspective and Technical Perspective. Studies on social security issues started in the middle of the 20th century. Till now most of the researches are done under the political, economic, sociological and regional perspectives. Only few researches are done in the technical perspective. In this seminar the technical perspective is taken into account. In the technical perspectives the main issues are problem analysis, process and policy modelling, business oriented analysis, correlation analysis, infrastructure support, and data driven analysis.

V. SOCIAL SECURITY SERVICE AND DATA

There is a large number of social security services are provided by each government. And each services carried out through different steps. So huge amount of data is created with each service. For the data mining process the data collected from Center link. Center link is a program in the Department of Human Services that delivers a range of payments and services for retirees, the unemployed, families, careers, parents, people with disabilities, Indigenous Australians, and people from diverse cultural and linguistic backgrounds. In this paper the discussion about a simple business workflow is performed. As a result of implementing day-to-day social security services, a huge amount of data has been accumulated which increases dramatically every day. The social security and social welfare area contains a huge amount of data. So it is very difficult to perform on the data. If the data collection is based on some classification, then the analysing of data is little easy. The data have mainly eleven classifications. The social security data have many characteristics. Each of them are important in different views. The knowledge about the characteristics of data, make the data mining process more efficient. There is eleven characteristics for social security data.

VI. FRAMEWORK OF SOCIAL SECURITY AND DATA MINING

The data mining applications in social security are driven by business objectives and underlying data. The creation of the basic frame work of Social security Data mining after studying the data classifications and characteristics makes more efficient. The basic frame work consists of three layers. Analysing the social security data mining goals and challenges the tasks for the frame work created.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 6, Special Issue 12, July 2017

In this paper a basic framework of Social Security Data Mining is provided that consists of three layers: the data layer, the business objective layer, and the data mining goal layer. The data layer summarizes the main data resources. It consists of customer data (customer demographic and circumstance information), service data (service usage and procedural information) etc. The business objective layer includes the main aims and expectations for the implementation of social security services. For instance, a customer service enhancement (to instantly provide high-quality services to those with particular needs) etc. The data mining goal layer lists the main goals in mining social security data to enhance business objectives; for instance, customer-centric analysis, payment-centric analysis, debt-centric analysis, income-centric analysis etc. The main data mining goals in the social security area are divided into the following five classes, according to the understanding and practices of key entities, problems, and challenges in social welfare business by data mining: 1) over-payment-centric analysis; 2) customer-centric analysis; 3) policy-centric analysis; 4) process-centric analysis; and 5) fraud-centric analysis. Each of these classes contains different goals. The study of the nature of the data reveals that the social Security Data Mining faces a lot of challenges. The key challenges consider in terms of the main procedures of social security data processing, pattern analysis, and knowledge delivery.

To support the data mining goals many tasks need to be performed in SSDM. We categorize the SSDM tasks according to the main entities in social security/welfare business and to address the SSDM goals data processing, activity analysis, customer risk analysis, earnings analysis, change detection, debt analysis and fraud detection. Taking the Australian social welfare business as an example, as a one-stop-shop, the Australian Commonwealth Department of Human Services (Centrelink) delivers a range of Commonwealth services to the Australian community and is responsible for the distribution of around \$86.8 billion, i.e., 30% of the Commonwealth's outlay, to about one-third of Australians. In day-to-day interactions between the government and customers, Centrelink accumulates a large amount of interaction data. For instance, Centrelink provides 361 000 face-to-face services each working day and processes 6.6 billion transactions against customer records each year. It has been shown that the number of such interactions increases every year. The government has progressively recognized the importance of analysing these interactions to obtain a deep understanding of customers and organization–customer relationships, to actively manage customers, to improve government service quality and objectives, and to inform policy design. An issue of particular interest is the identification of the drivers that cause noncompliance in organization–customer interactions. Noncompliance drivers may result from many aspects or staff errors. The 2007–2008 audit report by the Australian National Audit Office (ANAO) drew attention to the importance of deeply understanding customers, and of addressing the behavior and behavioural changes in rising debt from the perspective of the customer, government administration, client group, and community.

VII. VARIOUS TYPES OF DATA MINING

A series of researches and commercial projects are conducted with Centrelink in data mining. The debt prevention and debt recovery are two of the most important issues in Centrelink. So based on these two concepts the data mining is performed. There are different techniques and methods for data mining. The data mining on social security and welfare data prefer five methods. That are Modelling impact of activity/activity sequence, Mining impact target patterns, Mining Positive and Negative sequential patterns [1], Mining

- 1) Customer demographic and circumstance data, recording information about a customer and his/her circumstances, circumstance changes, etc; for instance, home address and the history of address change;
- 2) Benefit/allowance data, regarding the information about specific benefit/allowance design and applicability in alignment with customer eligibility, and management processes;
- 3) Customer pathway data, reflecting the history and relevant details of a customer's use of government services, such as the number of services, when, and from which service centers the services have been applied for, and granted;

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 6, Special Issue 12, July 2017

- 4) Activity data, providing activity records information about who (maybe multiple operators) processes what types of activities (say change of address) from where (19 customer service centers) and for what reasons (say the action of receipt of source documents) at what time (date and time), as well as the resultant outcomes (say raising or recovering debt);
- 5) Facility usage data, regarding the resources used by or for customers, e.g., phone calls and online services;
- 6) Service policy data, information about policies, the applications of policies to customers with particular circumstances;
- 7) Service transactional data, day-to-day information recorded regarding the use of services, such as new registrations, new claims, debt review, etc.;
- 8) Service performance data, concerning service quality and performance, such as overpayments and their distribution.

VIII. CONCLUSION

Data mining play a critical role in social security for debt prevention, recovery and customer analysis. Mining social security and public sector data are still an open field for business applications in data mining and machine learning. There is few researches are carried in technical perspective. So there is a large scope for researchers. The researches in the social security and social welfare data helps the public and government to take important decisions about social security services. Such social science research could be priceless for better welfare program administration, program evaluation, and policy analysis. Mining social security data, including modelling the impact of activity or activity sequences, mining impact-targeted activity patterns, mining positive and negative sequential patterns, conducting impact-targeted sequence classification, and mining combined association rules. The reliability and security of social security systems could be investigated and enhanced using data mining. While many existing algorithms and models seem to be suitable for SSDM and public-sector data mining, it is necessary to carefully check that they fit perfectly into the public-sector data characteristics. Variations may be necessary. A typical situation is that many service patterns are identified, but most if not all of them simply reflect policy and process arrangements, which is not helpful for government service decision making.

REFERENCES

- [1] L. C. H. B. Y. O. Y. Zhao, H. Zhang and C. Zhang, "Data mining applications in social security," New York: Springer, pp. 81–96, 2009. D. K. Professor and G. J. U. o. S. . T. H. D. B. R. S.
- [2] D. o. C. S. . E. G. J. U. o. S. . T. H. Dean, Faculty of Engineering & Technology, "Rise of data mining: Current and future application."
- [3] Y. Z. L. Cao and C. Zhang, "Mining impact-targeted activity patterns in imbalanced data," IEEE Transactions, vol. 20, pp. 1053–1066, Aug. 2008.
- [4] R. J. Barro and C. Sahasakul, "Average marginal tax rates from social security and the individual income tax," J. Bus., vol. 59, no. 4, pp. 555– 566, 1986.
- [5] H. Berghele, "Identity theft, social security numbers, and the web," Commune. ACM, vol. 43, no. 2, pp. 17–21, 2000
- [6] L. Cao, "In-depth behavior understanding and use: The behaviour informatics approach," Inf. Sci., 180, no. 17, p.3067-3085, 2010.
- [7] R. W. Boadway and D. E. Wildasin, "A median voter model of social security," Int. Econom.Rev., vol. 30, no. 2, pp. 307–328, 1989.
- [8] M. J. Boskin and G. F. Break, The Crisis in Social Security: Problems and Prospects. Oakland, CA: Inst. Contemporary Stud., 1977.
- [9] G. Burtless and R. A. Moffitt, "Social security, earnings tests, and age at retirement," Public Finance Rev., vol. 14, no. 1, pp. 3–27, 1986