Market Opportunity and Policy Support for Chinese Old Aging Industry: An Application of Text Mining

Chenguang Li[1, 2,*]

[1]Business Management Department, North China University of Technology, 100144 Shijingshan Distinct, Beijing, China
[2]Enterprise Management Intelligent Decision Research Centre, North China University of Technology, 100144 Shijingshan Distinct, Beijing, China

Abstract. Text mining has become a vital method of natural and social science research. A policy documents text mining framework with three phases has been introduced. Using 52 policies of aging industry development, this paper explores five business opportunity and three areas policy support for Chinese old aging industry. These findings contribute both foreign and local the elder service enterprise.
Key words: text mining, market, opportunity, policy, age

1 Introduction

A growing consensus recognizes that the population aging has substantially impacted on economic growth. Classical works in population aging, focusing on the government consumption budget [1], the labour market [2], work model [3], etc. Albeit most scholars believe that social problems (such as financial strains comes from social security, Medicare and Medicaid and so on) is attributed to people are getting older faster, ageing population has been creating business opportunities for old aging industry.

China has the largest population in the world and represented about 33% of the Asian population. For the past few years, the population size was expending dramatically due to improvement in medicine and sanitation that was benefiting from rapid economic growth. And the population as a whole was aging rapidly. The investigations from National Bureau of Statistics of the People's Republic of China shows population aged over 60 has reached 231 million in 2016, and it will increase to 480 million by 2050.

When 1980s Chinese government pushed a "later, longer, fewer" approach that encouraged people to marry later, wait longer between children, and have fewer babies [4]. In 30 years, most families in cities have only one child. A few years later, the children from single-child parents will face what is known as the "4-2-1" phenomenon (when the child grows up to working age, he or she must have to care for two parents and four grandparents in retirement [5]. These middle-aged men have been bearing a tremendous burden.

---

[*] Corresponding author: dr.lcg@qq.com

So China's government encourages the development of aging industry. And most polices were set for getting more local and foreign companies involved in that industry development process. Scholars advise finding implications for public policy and identifying a profitable business opportunity is important to a company, especially in developing countries [6]. Chinese old ageing industry brings a great business opportunity. For that reason, it is necessary to analyse policies for mining market opportunity and government support.

Text mining applications have been developed for policy analysis [7], thus this study will use this method for exploring the market opportunity and government support. By reading over the official aging industry documents, the study explores the presence of meaningful clusters of terms that appear across documents in the collected corpus [8]. Then it uses semantics to make sense of the opportunity and support [9].

This article presents a framework for text mining. The remainder of this paper is organized as follows: Section 2 presents a framework of policy documents mining. Section 3 describes the text mining process. Section 4 provides the analysing results. Finally, we briefly give the suggestions in Section 5.

2 A Framework of Text Mining

Text mining, also known as text data mining or knowledge discovery, has become a vital research tool with the rising popularity of data mining [10]. This method, referring multidisciplinary theory, is the complex task with unstructured and fuzzy data processing. And the process includes information index, text analysis, feature extraction, clustering, categorization, networking, semantic analysis, machine learning, and data mining. In recent years, the commercial potential of text mining has been emphasized on account of most market information is contained in text documents.

This article proposes three phases of a policy mining process (as Fig. 1 illustrating). First, pre-treatment that transforms policy documents into a text mining set could normalize the policy text format. The policy intermediate form is semi-structured, which contains word segmentation, feature extraction, clustering, etc. by segmenting long text. There is no connected concept at this stage. Next, text mining, as a significant task, discover the relationship across features and concepts based on frequency and network analysis. In this phase, some appropriate filter and model can be useful in keywords' extraction. Then, analyzing the association of high frequency and centrality words in postprocessing of the concept intermediate form, the path semantics and topic could be explored. At last, the results are summarized from path meaning of path correlation.
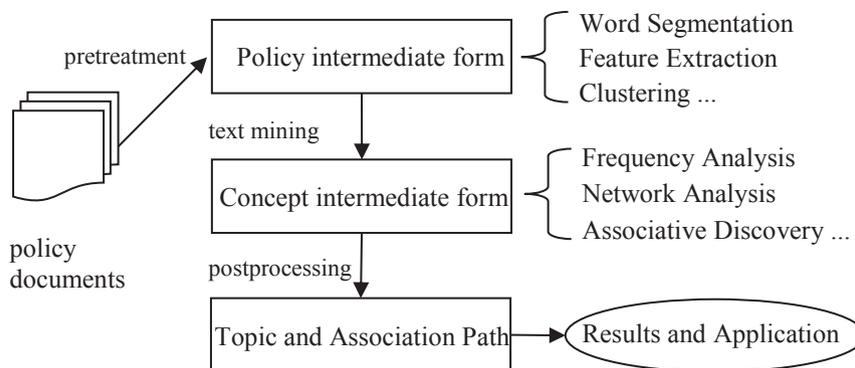

Fig. 1. A policy text mining framework

Policy documents will be put into a set with unified format. Policy intermediate form should split statement, checking and removing interference and nonsense words. The concept form is prepared for frequency and association analysis based on word segmentation. Topic and association path mining depends on network analysis with the use of semantics.

3 Text Mining Process

3.1 Sampling

The study used 52 policies of aging industry development have been issued by Chinese Ministry of Civil Affairs, Commerce and 32 provinces' Department of Civil Affairs from 2013 to 2017 (as Table 1 shown). The sample is involving in planning, opinion, inform and law, which are the important policy forms in China.

Table 1. Aging Industry Development Policy sample

Type	Number
planning	32
Opinion	6
Inform	10
Law	4

3.2 Text Mining

Firstly, the policy documents are unified Unicode form and stored in text (.txt) document files, which are made up of the text library $V(d)$. And $w_n(d)$ represents semantic words in each policy p_n, so we have

$$V(d) = (p_1, w_1(d); \ldots; p_i, w_i(d); \ldots; p_n, w_n(d)). \tag{1}$$

This study extracts 5251 segments after browsing through the policy text. By filtering nonsense words, there are 1741 words constitute policy intermediate form. Then, 302 words have been screened in feature extraction.

In the next place, word frequency statistics have been done for searching key words. However, to avoid topic highly correlated non-significant influence, it must remove some frequently used words such as "old", "aging", "policy", etc. Redundant words, should be added to filter list, need researchers to distinguish from synonyms.

By applying maximus matching method, the word segmentation weighting coefficient α_i can calculated with

$$a_i = \frac{TF_i * \log(\frac{N}{DF_i})}{\sqrt{\sum_k \left[TF_k * \log(\frac{N}{DF_k}) \right]^2}} \tag{2}$$

where TF_i represents the frequency of words, DF_i describes the text's, i is occurrence, and N=52.

Thus, 40 highest frequency words have been found, which illustrates in Table 2.

Table 2. Highest frequency words

Words	Freq.	Words	Freq.	Words	Freq.	Words	Freq.
care service	3643	culture	333	talent	202	quality	137
community	1013	service industry	290	project	200	security	134
medical treatment	831	standard	280	technology innovation	196	assistance	133
facility	660	rehabilitation	242	retirement homes	188	hospital	128
safeguard	602	resource	238	disease	186	fund	127
insurance	559	demand	232	training	181	civil rights	125
health	430	rescue	222	fitness	145	psychology	125
education	414	cooperation	221	subsidization	143	travel	120
nursing	412	sanitation	213	study	141	investment	119
family	368	market	207	charity	138	Internet & informatization	118

To clear the meaning of the words, co-word matrix has to be compiled based on the successive appearance of one word to another. Next, cluster analysis of whole high-frequency words, as Fig. 2 shown, demonstrates the association among semantics. This networking process can help to observe the distribution of internal relationship, which is described by the core words and others associated with them.
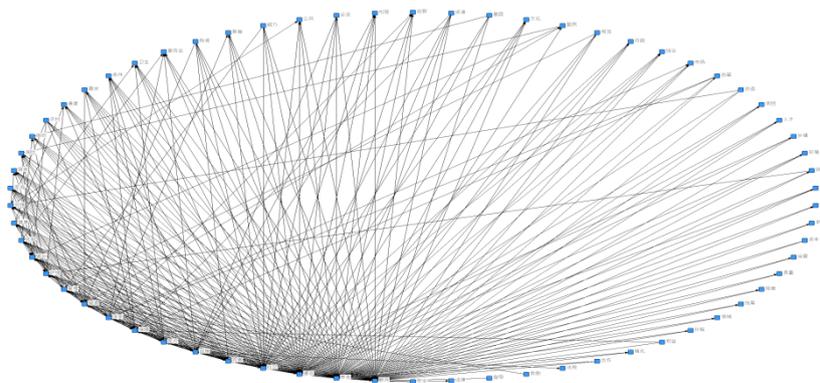

Fig. 2. Semantic network. The highest frequency words focus on care service, medical treatment, facility, safeguard, education, etc.

Then, the network analysis shows the centrality of the semantics, and measures the dispersion degree of the words. This article uses Ucinet 6.0 to calculate that parameter. As

Table 3 represents, the degree centrality of "care service" is largest while meaning it is the vital work the government encouraged.

Table 3. Centrality of semantic words

Words	Centrality	Words	Centrality	Words	Centrality	Words	Centrality
care service	47.724	culture	24.884	talent	14.544	quality	11.216
community	46.277	service industry	23.547	project	14.501	security	11.169
medical treatment	42.445	standard	23.445	technology innovation	14.385	assistance	10.801
facility	36.995	rehabilitation	22.696	retirement homes	13.767	hospital	10.762
safeguard	31.011	resource	20.439	disease	12.849	fund	9.718
insurance	29.589	demand	18.947	training	12.496	civil rights	9.143
health	29.150	rescue	18.635	fitness	12.114	psychology	9.096
education	27.409	cooperation	16.902	subsidization	11.375	travel	8.301
nursing	26.742	sanitation	15.354	study	11.375	investment	7.946
family	25.906	market	14.665	charity	11.253	Internet & informatization	6.119

To cluster around the concept network, it is necessary to clique analysis. With the centrality of semantic words, we can draw the graph for associative discovery (As Fig. 3 illustrated).
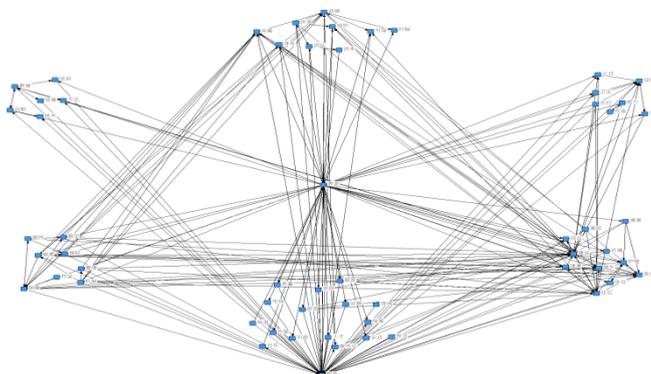

Fig. 3. Clustering semantic network. There are six cliques in the network.

Finally, we can categorize the related words from the semantic network cluster. Meanwhile, the interrelationships of the semantics component of the association path has provided foundations of decision analysis (as Table 4 shown).

Table 4. Association path.

Cliques	Semantic words
care service	community, family, nursing, standard, cooperation, market, project, retirement homes, quality, travel
medical treatment	health, rehabilitation, disease, psychology
safeguard	demand, insurance, rescue, civil rights, fitness, charity, security
facility	sanitation, hospital, Internet &informatization
education	culture, talent, technology innovation, training, study
policy support	resource, fund, subsidization, service industry, investment

4 Opportunity and Support

We can find that Chinese old aging industry, from the association discovery, focuses on care service, medical treatment, safeguard, facility and education. In these areas, elderly care service with a long path indicates huge market opportunity. For instance, community is a basic unit of social organism. And it is a large group of families. The common place of aged in China is community, where they have great fun, chatting, singing, dancing and exercising. Therefore, develop community care services is potential business prospects. Standard of service promotion is significant work to get profit, when our old aging market just emerging several years. The parents of adult who is the only child need home care for sharing care pressure. Cooperation with foreign companies, has been encouraged by the local government such as Jiangsu province, is also a wise choice of old aging industry development. Furthermore, both ORPEA from France and RIEI from Japan have built dozens of retirement homes. Each elder man charges 10,000 RMB (about 1,544 US dollars). And they can get two thousands to five thousands RMB ($309~$772) subsidies from the local government. Due to the foreign capital involved, the state enterprise has to improve the service quality for competition. Moreover, the senior tour project is another profitable business.

In medical treatment area, it is urgent for providing information to planners of health care for the aged. The importation of new medical equipment and the effective medicine can also help to develop old aging industry. In recent years, elder people's psychological health has been paid great attention, but the absence of the elderly psychiatrists is bad for their health.

The insurance for the aged is so important for developing the industry. We need more types of the insurance project for ensuring safety. The charity organization will be welcome for aiding the poor and offering social welfare projects to cope with the demand of security and fitness. Besides, elderly rights protection is an emerging market. And emergency project for rescuing is also a desired service.

With the rapid development of the Internet and mobile devices, the frequency of Chinese elder person using mobile phones is increasing gradually. The idea of Internet doctor and Mobile doctor will bring a new business in old aging industry.

Geriatric university gets a quickly develop in China. And study brings a lot of fans to the aged and voluntary activities make their lives more meaningful. In addition, training for elderly service and research new techniques for caring can create jobs for job seekers.

China's government encourages foreign cooperation and give subsidies, land resources and investment guarantee. In terms of tax, the investors will enjoy all the preferential policies for business invitation. And local governments are making a big push to streamline the permitting process.

5 Conclusion

This paper focuses on the market opportunity and policy support for Chinese old aging industry. We introduce a policy documents text mining framework with three phases, to analyze 52 reference policies. Applying association path discovery, the study finds out five business opportunity for foreign and local companies. These findings contribute to development of Chinese old aging industry. And the policy support from China's government is described briefly in this article. It is an useful reference for foreign enterprise.

Acknowledgements

References

1. R. Mao, J. Xu, Population aging, consumption budget allocation and sectoral growth. *China Economic Review,* **30**(44), (2014)

2. K. Lisenkova, M. Mérette, R. Wright, Population ageing and the labour market: Modelling size and age-specific effects. *Economic Modelling,* **35**, 981-989 (2013)

3. B. Beach*,* Ageing populations and changing worlds of work. *Maturitas*, **78**(4), 241-242 (2014)

4. M. Hvistendahl*,* Has China outgrown the one-child policy? *Science*, **329**(5998), 1458 (2010)

5. J. K. Miller, X. Fang, Marriage and family therapy in the people´s Republic of China: Current issues and challenges. *Journal of Family Psychotherapy*, **23**(3), 173-183 (2012)

6. J. C. W. Smith, R. L. Watts, The investment opportunity set corporate financing, dividend and compensation policies. *Journal of financial Economics*, **32**(3), 263-292 (1992)

7. E. Ngai, P. Lee, A review of the literature on applications of text mining in policy making. *Proceeding of the 20th Pacific Asia Conference on Information Systems*, 343-354 (2016)

8. C. Boussalis, T. G. Coan, Text-mining the signals of climate change doubt. *Global Environmental Change*, **36**, 89-100 (2016)

9. E. Tobback, H. Naudts, W. Daelemans, E. Junqué De Fortuny, D. Martens, Belgian economic policy uncertainty index: Improvement through text mining. *International Journal of Forecasting, In Press* (2016)

10. A. H. Tan, Text mining: The state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Disocovery from Advanced Databases*, (1999)