A Text Mining Approach to Explain What Being Actively Engaged in Life Means

Richi Nayak

School of Information Systems, Queensland University of Technology, Brisbane, Australia

Introduction

- It is widely accepted that many first world countries are experiencing extended life expectancies and reduced fertility rates, which is resulting in an ageing population.
  - A definition of active ageing is “the process of optimizing opportunities for physical, social, and mental well-being throughout life in order to extend healthy life expectancy.
- The Australian Active Ageing (Triple A) study at the Queensland University of Technology (QUT) has conducted a national-wide postal survey to collect the responses of older people on a wide range of questions related to ‘work’, ‘learning’, ‘social’, ‘spiritual’, ‘emotional’, ‘health and home’, ‘life events’ and ‘demographics’.
- The survey also included an open-ended question that the applicants responded to was “Briefly describe what being actively engaged in life means to you?”.
- The purpose of this paper is to present the findings and analysis of the open-ended question in the survey using the text mining techniques.
- This paper attempts to gain a better understanding of the impact of aging upon older Australians’ engagement in society and their sense of well being.

Method: Text Mining

- Text mining extracts the important concepts and emerging themes from the collection of text sources using computation linguistics, machine learning and information science methods.
- We utilize the Leximancer text mining application [3] to extract relationships from the text data.
  - It is based on the bayesian theory using the words which make up a sentence to predict the concepts being discussed.
  - Each *term* appearing in the text data is analyzed to form a *concept* and then many concepts are put together to form a *theme* based on the co-occurrence of terms and concepts in the text.

Data

- The Triple A study data includes responses of 2645 members (45% response rate) of a large Australia wide seniors organization on a wide range of questions related to their life.
- The survey includes 178 questions grouped into 8 sections covering the life aspects of social, health and home, emotional etc and an open-ended question.
- The sample included a reasonable balance of gender (57% female) and state representation (NSW: 32%, Vic: 24% and Qld: 20%).
- The respondents were quite comfortable financially, well educated and majority were living with a companion, in their own home in a metropolitan area.

Process

- Preprocessing steps were performed.
  - (1) removal of surveys (instances) that did not have sufficient information of age, gender etc,
  - (2) application of spelling and grammar checking function in Excel in order to locate and correct the misspell words, incorrect abbreviations and slang,
  - (3) stop word removal and stemming were performed.
- The data file was then arranged and split into categories of both Age and Gender that allowed us to thoroughly compare and contrast the different combinations of categories.

Combinations of categories for analysis

Grouping	Category	Qty
All comments		1964
Grouped by age	55-64	1252
	65-74	474
	75 & over	238
Grouped by gender	Male	870
	Female	1094
Males grouped by age	55-64	523
	65-74	230
	75 & over	117
Females grouped by age	55-64	729
	65-74	244
	75 & over	121

Comparison by Age Groups

Age group 55 – 64

- The most significant concept and theme is ‘Family’.
- ‘Life’ and ‘Active’ are next in the level of frequency.
- There is an integral relationship between ‘family’ and ‘time’ as well as having close ties with the ‘enjoy’ entity.
- ‘Work’ is considered a significant concept ranking highly for both males and females.
- ‘Work’ and ‘Active’ aspects comes to the fore in Males whereas ‘Health’ and ‘Social’ aspects are more frequent in the Females.
- This may indicate that Males place a greater importance in physically active aspects while Females signal social themes as figuring in their perception of active ageing.

Comparison by Age Groups

Age group 65 – 74

- This age group shows that they consider an active life to being engaged in it, living a full life, having a life with purpose and friendship in terms such as visit, involvement and participation.
- In this age group, males considered ‘work’ a more important concept than females.
- Both males and females emphasize on Family. However the terms that form the ‘Family’ concept in both groups differ.
  - The females were concerned with interaction with emphasis on meeting, touching, visiting, whereas the males, although the terms indicate a need to be interactive, are less personal.
  - Their terms include social, involved, community, and activities.
  - Although it is a subtle difference, it could possibly indicate that for males, family is also a source of activity or social events.
  - Whereas with females, it could be perceived as a more emotional connection as opposed to as a source of activities or social connectivity.

Comparison by Age Groups

Age group 74+

- Ageing is perceived as having a close association with family and friends and maintaining an engagement in life through activities with these people.
- The concept of ‘health’ emerges as a significant concept, more highly ranked by the males than the females.
  - This differs from the prior age groups indicating that this group may be considering health issues more frequently in relation to active ageing.
- A very interesting difference that appears between male and female of this age group is the ‘enjoy’ concept.
  - Female associate enjoy with life concept.
  - However, male segregate the enjoy concept from either of family or active concept.
  - Male population subtly emphasizes more on ‘active’ concept in comparison to female.
  - Male population being actively means being involved with people.
  - The female population seems to relate being actively with being involved with their loved ones, friends and social groups such as church.

Comparison by Gender

- All Females:
  - Ageing is perceived as having a close association with family and friends and enjoying the activities with these people.
  - Ageing is also perceived as being active in life through the time based work activities.
- All Male:
  - Ageing is perceived as being active in life and spending time with family and friends.
- The most significant difference between the results of the males and females, are that where the females population were concentrated around ‘Life’, ‘Family’ and ‘Friends’, the male population seems to include a wider range of concepts including ‘Active’ and ‘Work’ as heavily weighted.

Summary

- A strong common themes across all groups, male, female and age groups were that ‘Family’, ‘Friends’ and ‘Activities’.
- The younger age group considers ‘Work’ an important variable in active ageing, and this drops off to be important for only males in the next group, and is not ranked in the 75 and older group.
- Within the eldest group however, ‘Health’ emerges as an important concept, and given that this group is in a minority in the population as a whole, could explain that lack of representation overall of health concerns.
- These trends seem to indicate that the respondents are considering factors that are relevant to them in the immediate timeframe when responding to the question without consideration toward their future.
- There are various areas where these results may have impact and warrant some consideration such as future planning for government funding of community activities.

References

1. Hearst, M. A. (1999): *Untangling Text Data Mining*. The 37th Annual Meeting of the Association for Computational Linguistics, Maryland, June 20-26, (invited paper).
2. Kalache, A. 1999. *Active Ageing Makes the difference.* World Health Organization. Bulletin of the World Health Organization. 77 (4): 299.
3. Leximancer (Accessed September 28th, 2006). http://www.leximancer.com/cms//index.php?option=com_content&task=view&id=59&Itemid=114
4. SAS Text Miner. (Accessed September 8th, 2006) http://www.sas.com/technologies/analytics/datamining/textmine