Data-Driven Approach to Understanding Senior's Needs: An Automated Unsupervised Learning Solution for Feedback Analysis in Singapore

Hock Lin Sng, Agency for Integrated Care, Singapore
Juan Zhen Koh, Synapxe, Singapore
Joycelyn Yun Ting Woo, Synapxe, Singapore
Yu Heng Tan, Agency for Integrated Care, Singapore
Winston Zhao Yang Ma, Agency for Integrated Care, Singapore
Andy Wee An Ta, Synapxe, Singapore

The Asian Conference on Aging & Gerontology 2025
Official Conference Proceedings

Abstract

The Silver Generation Office (SGO), under the Agency for Integrated Care (AIC), supports seniors in Singapore through home visits to understand their situation and connect them with services to address their needs, if any. Through these visits, valuable qualitative feedback on policies affecting them are collected. This study aims to develop a self-help tool using unsupervised natural language processing to analyse uncategorised free-text feedback to reduce manual effort required in summarising the feedback. A total of 41,891 anonymised and uncategorised free-text feedback collected from April 2022 to March 2023 were analysed using topic modelling algorithm, Non-Negative Matrix Factorisation (NMF), developed on Anaconda JupyterLab. The feedback was analysed by creating a Document-Term Matrix which represents the frequency of terms in each feedback, followed by applying NMF to extract topics with representative keywords. Human evaluation with inter-rater reliability (IRR) assessment was conducted with ten evaluators to assess its accuracy. Results showed that the model achieved over 75% accuracy, with high IRR coefficient above 0.876 after two rounds of evaluation. The model uncovered valuable insights that were previously challenging to obtain through manual efforts. The extracted topics help SGO to better make sense of the data, facilitating sharing of insights with stakeholders to highlight seniors' needs and preferences which will improve existing policies, programs, and services for seniors. The self-help tool is developed and currently in-pilot, allowing users to automate data preprocessing, conduct textual analysis, and generate visualisation charts. It may potentially enhance SGO's operational efficiency and reduce man-hours spent on data analysis.

Keywords: aging, seniors, topic modelling, text analysis, natural language processing, unsupervised machine learning, preventive health visit

Introduction

Background

Singapore will be a super-aged society by 2030 with one in four citizens aged 65 and above (NPTD Singapore, n.d.). Within the community, Agency of Integrated Care (AIC) adopts a Touch-Hold-Care model to build an integrated ecosystem of care that supports seniors and clients. Silver Generation Office (SGO), the outreach arm of AIC, aims to care for seniors by supporting their aspirations and addressing their ageing needs. SGO's volunteers, the Silver Generation Ambassadors (SGAs), conduct Preventive Health Visits, identifying seniors' health, financial and social needs through the Senior Engagement Form. This helps SGO to refer seniors to relevant support through SGO's network of community partners and government agencies. SGAs also record seniors' feedback on senior-centered government policies and schemes through the form, which encompass feedback regarding Healthcare, Finance, Digital Literacy and other general feedback. The feedback is captured through either pre-defined sub-categories (e.g. High daily living expenses, Employment challenges, Request for senior-friendly equipment/infrastructure etc) or a free-text option in the form.

With an average of 300,000 home visits conducted by SGAs and staff yearly across 17 Satellite Offices, SGO receives a significant amount of qualitative feedback data from seniors. Limited by manual effort alone to scan and sieve through the feedback, SGO is unable to gain a comprehensive understanding of seniors' sentiments and identify emerging trends. These insights can be shared with relevant stakeholders to improve the provision of care and support for seniors. With Singapore's growing ageing population, the needs of seniors are becoming even more diversified. Furthermore, with recent national initiatives such as Healthier SG (preventive health) and Age Well SG (ageing in place), the insights gathered from the feedback may be valuable in the shift towards active ageing. The search for a more effective solution to address this analytical need thus becomes even more imperative.

Importance of Text Analysis

Currently, feedback captured under pre-defined sub-categories (Healthcare, Finance, Digital Literacy, General Feedback) is analysed by sentiment (i.e. Happy or Concern). This allows for an understanding of areas of concern or satisfaction among seniors nationwide. However, a significant portion of feedback, particularly the free-text responses that fall under the "Others" subcategory, remains largely unexamined. Hence, it leaves a significant gap in understanding the sentiments of seniors. This unexamined feedback is a rich source of information, offering insights that are not covered in the pre-selected categories. The manual process of reading through and interpreting this free-text feedback is not only labour-intensive and time-consuming but also prone to inefficiencies and inaccuracies. It is challenging to identify and categorise common themes across numerous responses manually, especially when dealing with large volumes of data.

Advanced text analytics and unsupervised machine learning techniques can be applied to systematically mine this unstructured data, allowing data analysis of seniors' needs, concerns, and experiences. Insights gained can contribute to national initiatives to tailor services that better meet the needs of seniors, and ultimately enhance the quality of care and support provided. This novel approach is also significant to uncover emerging trends from the older adults to allow us to better appreciate and address new expectations from the cohort effects of the fast-ageing population in Singapore.

Goal of this Study

The goal of this study is to develop, validate and implement a comprehensive framework for analysing free-text feedback provided by seniors during SGA visits, particularly focusing on the unstructured responses gathered under the “Others” subcategory.

By leveraging unsupervised machine learning technique, specifically Non-Negative Matrix Factorisation (NMF), this study aims to systematically identify and extract underlying topics and themes from seniors’ feedback. Other than generating insights, a key objective is to establish a robust, self-help text analytics tool that can be utilised by SGO to perform data analysis of seniors' feedback and to uncover insights and emerging trends for reporting to stakeholders.

Methodology

Overview

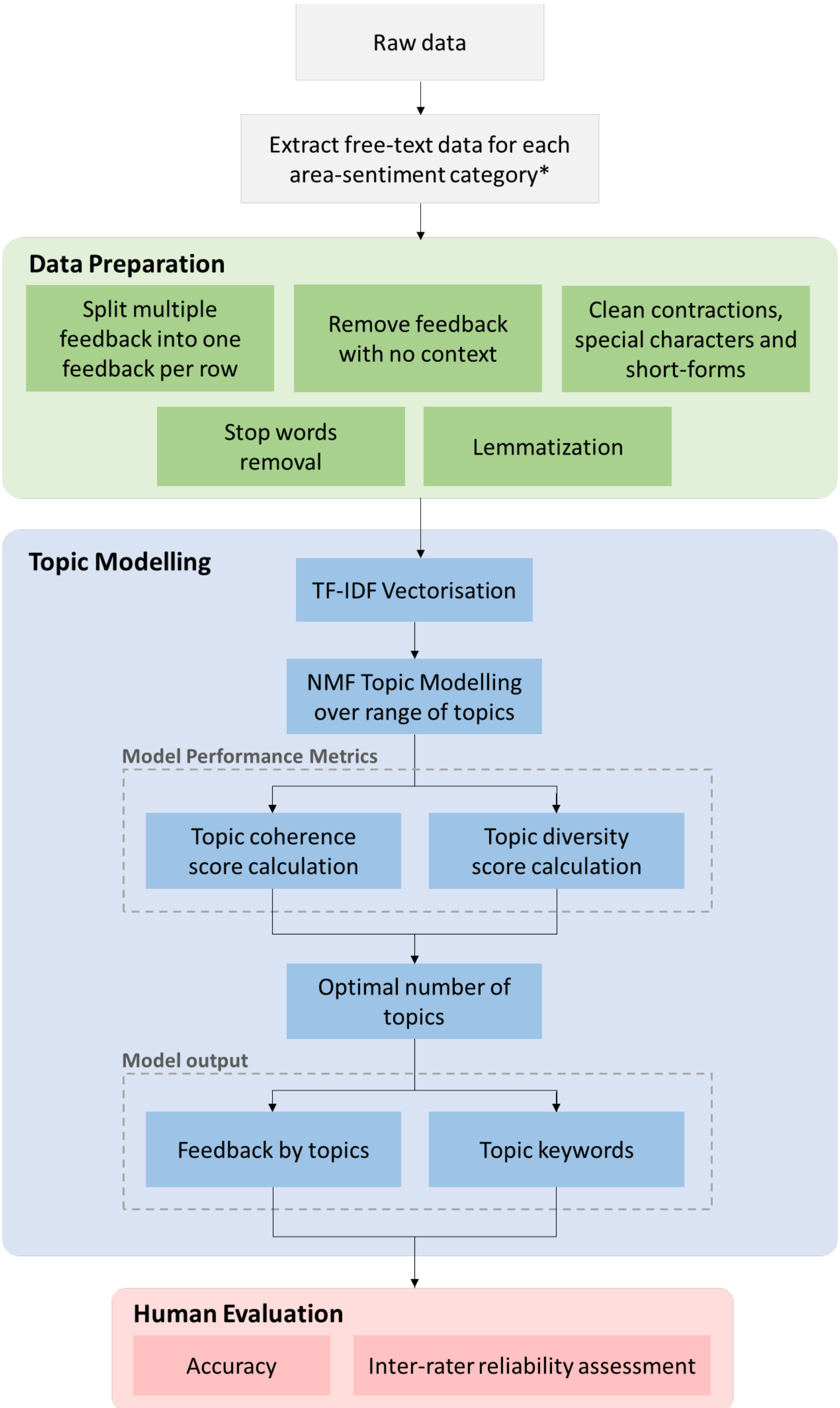
The approach of this study involves the implementation of a self-help tool that automates feedback analytics and reporting. Data preparation involves preprocessing of the raw data so that it is suitable for further processing and analysis. Key steps include cleaning, stop words removal, and lemmatisation. The data was categorised into eight area-sentiment categories. The processed free-text data are then analysed via topic modelling to extract latent topics underlying each area-sentiment category. Human-in-the-loop evaluation was conducted to evaluate the accuracy of the topics predicted, with inter-rater reliability (IRR) assessment to measure degree of agreement between evaluators. Finally, the workflow was implemented for users to self-help for management reporting.

Figure 1 outlines the research methodology for this study which involves three main stages: data preparation, topic modelling, and human-in-the-loop evaluation.

Data preparation involves preprocessing of the raw data so that it is suitable for further processing and analysis. Key steps include cleaning, stop words removal, and lemmatisation. The data was categorised into eight area-sentiment categories. The processed free-text data are then analysed via topic modelling to extract latent topics underlying each area-sentiment category. Human-in-the-loop evaluation was conducted to evaluate the accuracy of the topics predicted, with inter-rater reliability (IRR) assessment to measure degree of agreement between evaluators. Finally, the workflow was implemented for users to self-help for management reporting.

Figure 1

The Research Methodology for the Analysis of Seniors' Feedback



*Refer to Table 1 for list of area-sentiment categories

Data Preparation

Feedback data from SGA visits conducted between April 2022 and March 2023 was extracted for this study, with a total of 162,269 raw feedback collected regarding Healthcare, Finance, Digital Literacy and other general feedback.

Several data cleaning and corpus preparation steps were involved during data preparation. As each recorded feedback from a senior could contain multiple feedback, splitting of each record into individual feedback was done to ensure that every feedback was captured separately. Rows of data that did not provide contextual information were then removed from the dataset. The resulting corpus was then cleaned by removing contractions, special characters, and stop words. Lemmatisation was also performed to reduce words to their base form. There were 88,676 feedback after data pre-processing, out of which 47.2% (41,891 feedback) were tagged to subcategory "Others". The data were categorised into their respective area-sentiment category according to how the feedback question was structured in the engagement survey which helps to guide SGAs when filling in the feedback of seniors (Table 1).

Table 1
List of Area-Sentiment Categories

No.	Area-sentiment category
1	Digital Literacy, Concern
2	Digital Literacy, Happy
3	Finances, Concern
4	Finances, Happy
5	General Feedback, Concern
6	General Feedback, Happy
7	Healthcare Services and Schemes, Concern
8	Healthcare Services and Schemes, Happy

Topic Modelling

Topic Modelling Algorithm

Latent Dirichlet Allocation (LDA) and NMF are two commonly used and popular topic modelling techniques (Mifrah & Benlahmar, 2020). In this study, NMF topic modelling was used to extract meaningful topics underlying the corpus of feedback. NMF has two main advantages over LDA. First, NMF provides consistent and reproducible results due to its deterministic algorithms, as opposed to LDA which is a probabilistic technique (Svensson & Blad, 2020). Second, NMF allows for an easier tuning and manipulation of its parameters (Purpura, 2018).

NMF was chosen over LDA due to its deterministic nature, making it easier to interpret and more consistent in producing similar results across different runs. NMF has also proven its effectiveness in mining short texts (Athukorala & Mohotti, 2022), which is suitable for this use case where the word count of feedback recorded is of median of nine words. In addition, NMF is found to be more in line with human judgment than LDA (Egger & Yu, 2022).

Applying NMF on Text Corpus

NMF is a linear algebraic method that factors the document-term matrix (V) into two smaller non-negative matrices, document-topic matrix (W) and topic-term matrix (H), such that $V \approx W \times H$. Matrix W shows the distribution of the topics across the corpus of feedback, while matrix H captures the significance of terms within each topic.

Feedback from each area-sentiment category is processed separately, creating distinct corpora of text for each category. Topic modeling using NMF is then applied to these individual corpora. For each lemmatised text corpus, TfidfVectoriser transfers the textual data into the term frequency-inverse document frequency (TF-IDF) matrix which represents the importance of terms within each feedback, relative to the entire corpus. The vectoriser was configured to remove common stop words specific to each area-sentiment category, and it also considers both single words and bigrams to capture more context.

NMF topic modelling was then iterated over a range of number of topics from two to 20. Two metrics were used to assess the performance of each iteration – topic coherence and topic diversity. Topic coherence score was generated to measure the degree of semantic similarity among the top words of each topic, which helped in evaluating the interpretability of topics generated by the model. Topic diversity score was calculated using the pairwise word embedding distances to measure how distinct the top words in each topic were from other topics, which helped in evaluating how well-separated the topics were. Ideally, the top words within each topic are expected to be highly semantically related, while those across topics are to be as distinct as possible. The optimal number of topics is determined based on the highest coherence and diversity scores, ensuring that the topics generated are both interpretable and diverse.

The dominant topic for each feedback was determined based on the weights of topics in the document. After review and discussions amongst the researchers, consensus was reached to recommend a minimum weight threshold of 0.04, which was appropriate to determine the outlier feedback from the topics identified. The topic with highest weightage will be selected as the dominant topic for each feedback.

Evaluation

A 10% stratified sample was extracted for each area-sentiment dataset and evaluated by a panel of ten evaluators to measure the accuracy of the topic model in predicting the topic for each feedback. The panel of evaluators (refer to Appendix, Table 4) recruited were from the Headquarters and Satellite Offices of SGO with expertise and domain knowledge in seniors feedback reporting.

Figure 2
Evaluation Workflow

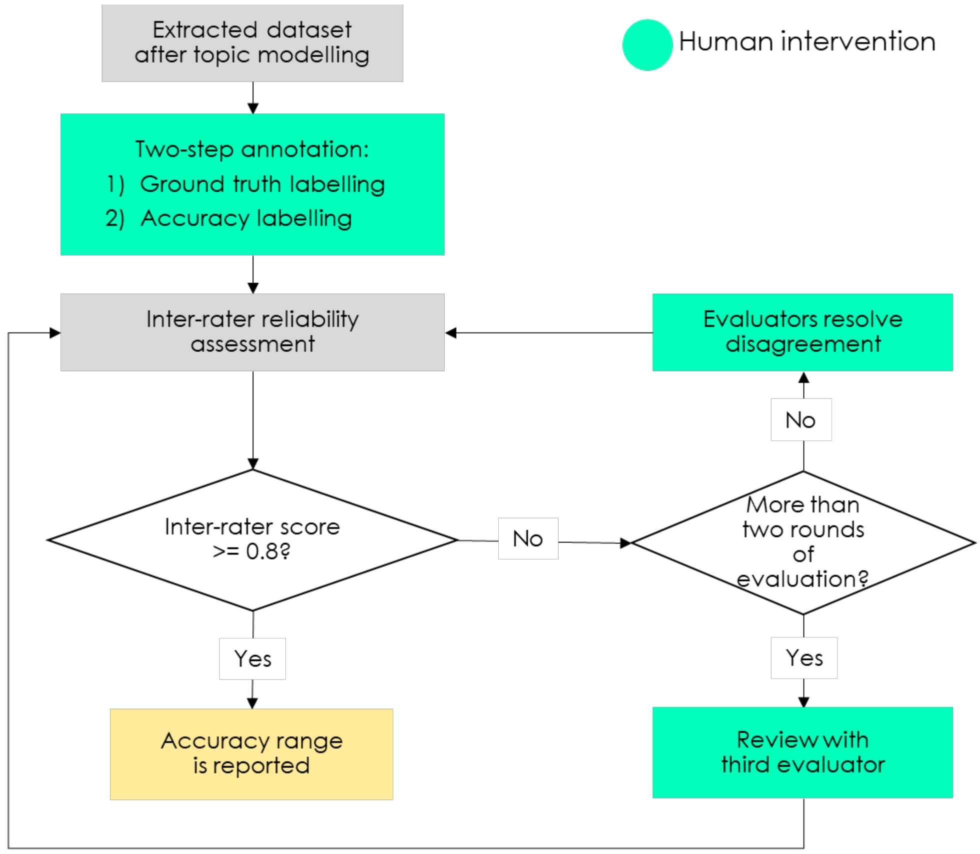


Figure 2 illustrates the workflow for the evaluation exercise. Each feedback was assessed by two evaluators to ensure unbiased evaluation. Each evaluator independently conducted a two-step annotation by firstly labelling the ground truth for each feedback, followed by reviewing the predicted topic against the ground truth label and annotating the accuracy of topic prediction.

IRR assessment was integrated in the workflow to quantify and validate the level of agreement between evaluators. IRR was calculated using the *irrcac* Python package, which closely follows the general framework of IRR assessment presented by Gwet (Gwet, 2014). The Landis and Koch scale was used to interpret the level of concordance between evaluators, providing a systematic approach to evaluating the consistency of the evaluators' assessments (Landis & Koch, 1977).

A threshold of 0.8 was set as the acceptable level of agreement according to the Landis and Koch scale (Landis & Koch, 1977). If the IRR coefficient was equal to or exceeds 0.8, it indicated a high level of agreement, and the accuracy of the evaluations was reported based on the evaluators' responses. Conversely, if it fell below 0.8, indicating a lower than acceptable level of agreement, the evaluators were required to resolve their disagreement in responses. Following the resolution discussion, the re-evaluated responses were subjected to a second round of IRR assessment to ensure that a minimal level of agreement is achieved. In the event if the first round of resolution failed to achieve an acceptable level of agreement, the involvement of a third reviewer is recommended to provide an additional layer of evaluation. This approach ensures that the evaluation process is robust, thereby enhancing the validity of the assessment outcomes.

Results

Topic Modelling

The topic modelling analysis aims to uncover underlying topics present in the feedback provided by seniors under the "Others" subcategory. By applying NMF to the textual data, topics were identified for each area-sentiment category.

Table 2
Optimal Number of Clusters by Area-Sentiment Category

Area-sentiment category	Number of optimal clusters
Digital Literacy, Concern	5
Digital Literacy, Happy	19
Finances, Concern	20
Finances, Happy	7
General Feedback, Concern	4
General Feedback, Happy	19
Healthcare Services and Schemes, Concern	9
Healthcare Services and Schemes, Happy	5

Table 2 summarises the optimal number of clusters identified for each area-sentiment category, representing the number of distinct topics that emerged from the feedback data. The determination of these topics involved evaluating metrics such as topic coherence and topic diversity to ensure the relevance of each topic and distinction across topics. The variation in optimal number of clusters highlights the diverse nature and the complexity of feedback across different area-sentiment categories. For example, "Digital Literacy, Concern" was found to have five distinct topics, while "Finances, Concern" had a higher complexity of feedback nature, yielding 20 optimal clusters.

The topics extracted for each area of feedback are included in the Appendix (Table 5). For each topic identified, the top five keywords by weightage are listed, providing a snapshot of the underlying theme.

Digital Literacy

Key concerns in digital literacy revolved around the use of government apps and online payment services, fast-paced digitalisation, and fear of scams. Topics like "scam, afraid, government, worry, afraid scam" and "fast, digitalisation, catch, digitalisation fast, pace" may reveal the seniors' struggles with adapting quickly to digital changes and are concerned with online safety. Topic "English, understand, Chinese, language, learn" highlighted the challenge of understanding English or learning Chinese, possibly pointing to issues of accessibility in digital services for non-English speakers.

The happy sentiment focused more on appreciation for learning opportunities, with topics like "good, educate, government, government good", "learn, slow, computer, good, thankful", "lesson, glad, glad government, conduct lesson" and "skill future, fund" indicating satisfaction with the digital literacy programs and skills upgrading support provided by the government. There were also recurring mentions of online transactions, such as payment and appointment booking, showing an acknowledgment of the convenience brought by these digital solutions, despite some ongoing concerns with fast pace of digitisation.

Finances

Concerns around finances were focused on the rising cost of living, with topics like "cost living, high", "food expensive, market, costly", "expense, medical, living expense, expense high" indicating worry over inflation and affordability of essentials, especially food, medical and housing expenses. Seniors also expressed hope for more government support, as seen in the topics "government, hope, government help" and "support, financial support, government" which suggested expectation or desire for further financial assistance or subsidies.

Positive finance feedback reflected satisfaction with the care provided by government (e.g., "care, government, care citizen") with the various government payouts and support received. There were also mentions of specific support programs such as Silver Support Scheme (SSS), Community Development Council (CDC) vouchers, Assurance Package scheme and Goods and Services Tax vouchers (GSTV), signalling a sense of relief from receiving financial support to curb rising cost of living.

General Feedback

The main concerns in general feedback echoed the financial worries faced by the seniors, particularly on the cost of living, with topics such as "cost live, high cost". The mention of "complaint, town council, neighbour" highlighted localised issues the seniors faced regarding the community they lived in.

In contrast, positive sentiments revolved around overall satisfaction with government schemes and policies. Repeated compliments for the government such as "good government", "satisfied government" and "appreciate government" showed the seniors' appreciation towards government's policies and schemes. Other areas of satisfaction include community relationships and environment, with topics like "good neighbour" and "environment, facility good" indicating positive social and physical environments, which could contribute to overall well-being of the seniors. Some seniors also expressed gratitude towards SGA visits ("SGA visit, appreciate visit") and safety of Singapore ("Singapore, country, safe, live").

Healthcare Services and Schemes

Concerns in healthcare were focused on rising medical costs, with topics such as "increase medical, cost" and "expensive, medication, fee expensive" being dominant. Other concerns include service standards at polyclinics/hospitals with topics like "polyclinic, long, wait time, appointment".

Positive feedback was centred on the government's healthcare subsidies and care programs, with topics like “subsidy, medical, healthcare, health, care” indicating appreciation for government’s efforts in keeping healthcare services affordable.

Evaluation Outcomes

The results of the evaluation process, including IRR and accuracy assessments are summarised in Table 3. This table presents the IRR coefficients from two evaluation rounds and the accuracy of topic predictions for each area-sentiment category.

Table 3

IRR Assessment and Accuracy Results

Area-sentiment category	IRR coefficient (Evaluation round 1)	IRR coefficient (Evaluation round 2)*	Accuracy
Digital Literacy, Concern	0.735	0.952	78.1-81.4%
Digital Literacy, Happy	0.876	-	86.7%
Finances, Concern	0.935	-	87.9-91.4%
Finances, Happy	0.622	1.00	86.0%
General Feedback, Concern	0.529	0.974	75.9-76.8%
General Feedback, Happy	0.908	1.00	95.0%
Healthcare Services and Schemes, Concern	0.962	-	96.6-97.7%
Healthcare Services and Schemes, Happy	0.979	-	97.7-100%

*Round 2 is conducted when round 1 IRR falls below 0.80

Responses from the panel of evaluators were in concordant within two rounds of evaluation. High levels of agreement were observed in categories related to the Healthcare, with coefficients above 0.962, indicating a strong consensus among evaluators. Conversely, “General Feedback, Concern” showed the lowest IRR coefficient of 0.529 in the first round, which improved to 0.974 in the second round after resolving disagreements.

In terms of accuracy, the topic model achieved over 75% accuracy across all categories, with six categories exceeding 85% accuracy. In general, for machine learning models, industry standards are between 70% and 90% (Hendricks, n.d.). Healthcare exhibited the highest accuracy above 96.6%, whereas “General Feedback, Concern” and “Digital Literacy, Concern” displayed relatively lower accuracy, possibly due to the complexity of these feedback categories and highlights the opportunity for further refinement in the topic modelling technique.

Discussion

Principal Findings

Overall, the extracted topics via topic modelling has provided a granular view of the various concerns and satisfactions expressed by seniors, offering valuable insights into the areas where support is needed or appreciated.

In Digital Literacy, seniors express anxiety around fast-moving digitalisation and scams. However, government-supported educational initiatives seem to mitigate some of these concerns, providing a sense of empowerment through learning. Inflation is a major theme across both Finances and General Feedback categories. Financial pressures are significant, but at the same time, government interventions are acknowledged and appreciated. This indicates that while there are financial concerns, seniors feel that they are not left unsupported. Positive feedback in General Feedback tends to focus on satisfaction with governance, particularly in areas of healthcare support and financial handouts. This could suggest that policies resonate well with seniors when they directly benefit from these programs. For Healthcare, medical expenses are a significant concern. At the same time, there are positive feedback on the government's healthcare subsidies and support that assisted the seniors in alleviating healthcare cost.

There were several overlapping themes across the different categories of feedback, which highlights interconnected concerns and sentiments that seniors express across various aspects of their lives.

1. Cost of Living and Financial Concerns

The recurring theme of financial concerns across Finances, General Feedback, and Healthcare Services and Schemes emphasises that economic pressures are central to seniors' lives. The rising cost of living and healthcare is a constant concern, while positive feedback about government support suggests relief from these concerns. This indicates that seniors view financial stability as integral to their overall well-being.

2. Government Support and Services

Seniors consistently express satisfaction with government support across all categories. Output from the positive sentiments categories frequently revealed appreciation for government's schemes and policies, emphasising positive perceptions of the government's support in various areas, from digital literacy initiatives to healthcare and financial support. Across the categories, seniors also frequently expressed their hope for further government support. The recurring theme of government support across domains underscores how the importance of public assistance programs to seniors' well-being. Regardless it being financial aid, healthcare subsidies, or digital education programs, seniors are dependent on government intervention to improve their quality of life.

3. Healthcare Access and Affordability

The rising cost of healthcare is a common concern highlighted across in Finances, General Feedback and Healthcare categories. Worries about affordability of healthcare amplifies overall financial concerns. Though there are worries on healthcare cost, several topics reflect appreciation for financial support related to healthcare.

The overlapping themes across categories illustrate the interconnected nature of the issues seniors face. These recurring concerns and satisfactions underscore the importance of integrated solutions that address multiple aspects of seniors' life.

The evaluation outcomes underscore the effectiveness and limitations of the topic modeling framework. High IRR in categories such as Healthcare suggests that these categories are relatively straightforward for evaluators to interpret and classify, while lower reliability in "General Feedback, Concern" reflects the broader and diverse nature of this category. The significant improvement in IRR reliability from the first to the second round demonstrates the efficacy of iterative resolution in enhancing evaluator agreement.

The accuracy results show that the topic modeling framework overall performs well, with most categories achieving high accuracy of above 85%. Categories such as “General Feedback, Concern” and “Digital Literacy, Concern” achieved slightly lower accuracy (>75%). The diverse nature of feedback within “General Feedback, Concern” may have contributed to challenges in effectively capturing the context of all feedback through the predicted topics. On the other hand, “Digital Literacy, Concern” had the smallest dataset size, hence some feedback which had context related to “cyber wellness” and “learn”, were not fully captured by the predicted topics. This suggests room for further refinement in the topic modeling approach to better handle complex feedback categories and improve the model's predictive accuracy regardless of dataset size.

Overall, the topic modelling analysis has deepened our understanding of seniors' needs and preferences. The following outcomes were achieved from the analysis:

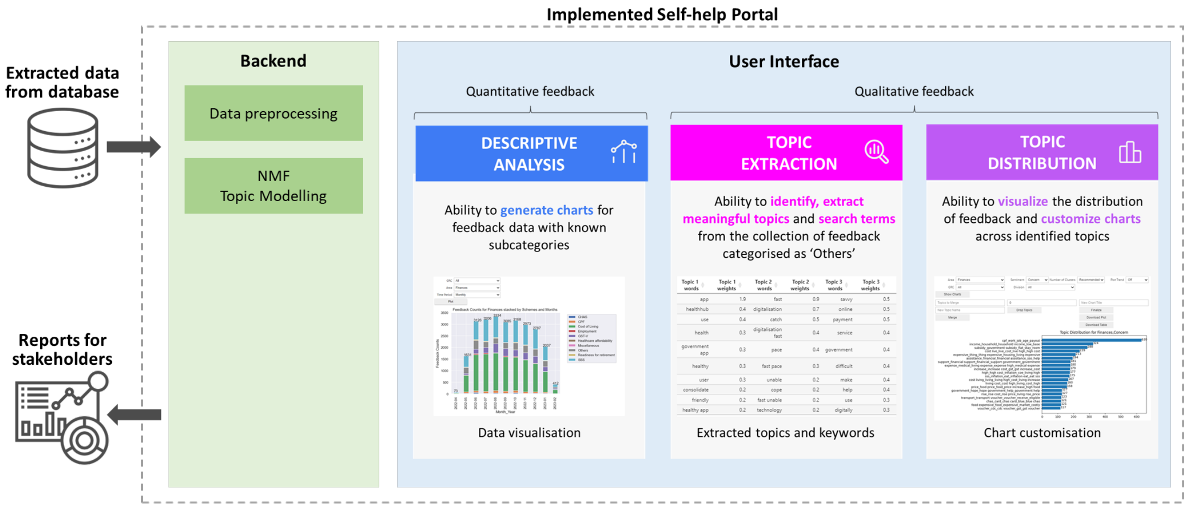
- Pinpointed the areas where seniors feel vulnerable, particularly in relation to modern challenges like digitalisation, rising living costs, and healthcare affordability.
- Identified areas where seniors are satisfied, allowing for reinforcement and scaling of programs or services that work well, such as healthcare subsidies and the various financial assistance provided.
- Provided a comprehensive understanding of how seniors perceive different aspects of their lives — in areas of digital literacy, finances, healthcare, and general well-being — and how their concerns are interconnected.

With a systematic analysis of seniors’ free-text feedback which was previously unexplored, stakeholders can now leverage these insights from the topics to suggest targeted improvements to existing policies, programs, and services for seniors.

Model Implementation

Figure 3

Backend and User Interface of Implemented Self-Help Tool



The study demonstrated the use of NMF to develop an automated solution for analysing seniors’ feedback. The results of the textual analysis through human evaluation showed that the topic model can accurately identify topics from the feedback. This study also saw the implementation of a self-help tool (Figure 3), offering an interactive environment for users to run automated data preprocessing, textual analysis, visualise outputs and review the output for management reporting.

Key Features of the Self-Help Tool

The implemented self-help tool provides three key features designed to enhance the analysis of feedback.

1. Descriptive Analysis

This feature is for the quantitative analysis of feedback tagged with preexisting sub-categories, allowing users to generate data visualisations to gain insights into the distribution of feedback over time.

2. Topic Extraction

This feature is for the qualitative analysis of free-text feedback, allowing users to run NMF topic modelling to extract meaningful topics, providing insights from the unstructured feedback which were previously unexplored. During this process, users have the flexibility to specify the exact number of topics they wish to extract from the feedback data. Alternatively, they can opt to use the number of topics recommended by the model, which is determined based on topic coherence and topic diversity scores. Additionally, users can adjust the minimum weight threshold for topic inclusion, which is set to 0.04 by default.

3. Topic Distribution

Once the topics have been extracted, this feature enables users to visualise how these topics are distributed. Users can also customise the charts to facilitate reporting.

Users and Stakeholders Reporting

The users of the tool is representatives from SGO Headquarters (HQ) and SGO cluster offices. HQ disseminates feedback to public agencies, while cluster offices analyses the feedback more granularly for each Satellite Office within their cluster for sharing with Grassroots Advisers.

At HQ, the outputs will be grouped by topics which are relevant to respective public agencies. For example, if there are topics generated on speeding bicycles on pedestrian footpaths and long waiting times for bus services, these topics would be grouped together and shared with Singapore's transport regulatory government agency. The aim is for such public agencies to act on the feedback to improve policies and services for the elderly population.

At cluster offices, the outputs are grouped by the areas and sentiments in Table 1 for each precinct which are represented by a group of Grassroot Advisers (GRA). The outputs are represented in the form of charts to give an overview and included in a quarterly update to GRAs. The aim is to leverage on their influence to move community resources towards addressing the feedback.

Benefits and Impact

The use of the self-help tool has benefited SGO in several ways.

1. Greatly increased SGO's operational efficiency, reducing approximately 210 man-hours per year spent to manually scan the feedback to provide insights for the quarterly GRA reports. The time saved freed up more productive hours to perform analysis of the charts and to provide insights from feedback data.
2. Uncover insights which might not have been previously detected through manual effort. Through the categorisation of feedback in topics, SGO was able to identify topics with significant numbers and uncover emerging trends from the seniors' sentiments.
3. Increase awareness of ground issues by community leaders. These feedback topics provide insights to the seniors' sentiments within their constituencies and help Grassroots Advisers focus on the relevant support and services required by seniors.
4. Improve policies and enhance services by public agencies through actionable insights. Public agencies are made known the growing concerns and issues that the elderly face which would help them to formulate better guidelines and programmes that address ground concerns.
5. Improve training curriculum for SGAs through the review of feedback themes generated, to educate them on the common feedback received and ways to better record the feedback for effective data analysis.

Limitations and Future Work

While the topic modelling framework has demonstrated effectiveness in analysing seniors' feedback, several limitations have been identified that suggest areas for improvement for future research and implementation.

Several predicted topics exhibited overlapping contexts, which could affect the interpretability of results. This overlap suggests a need for semantic understanding of the textual data. Transitioning to embedding-based algorithms allow semantic understanding and enhance the model's ability to distinguish between subtly different topics. This will reduce the manual effort required to combine duplicated topics together.

The developed solution was helpful to provide a high-level categorisation of feedback into topics but has room for enhancement to delve into the deeper details of each topic. A potential enhancement could complement topic modelling with language models to summarise the feedback keywords and themes into a short phrase to represent the category, and to query actionable insights on specific concerns or praises shared by seniors on a particular area (e.g. Frequency of bus 183 could be improved). This will reduce manual effort required to summarise the feedback themes and allow sharing of specific, deeper insights with stakeholders.

Furthermore, as some agencies have requested for aggregated data or anecdotal information, it would be very helpful if the raw feedback could be further categorised by the deeper and sharper topics. This would allow less resource-intensive methods to isolate and analyse raw feedback to generate further insights for stakeholders.

Conclusion

The study successfully developed and validated a framework for analysing free-text feedback with NMF topic modelling. The framework demonstrated its effectiveness in identifying and extracting underlying topics from unstructured feedback, and providing valuable insights into seniors' experiences and needs. Notably, the topic model, which was validated through human evaluation and IRR assessment, achieved high accuracy over 75% across all area-sentiment categories, with six categories exceeding 85% accuracy.

While the study demonstrated the developed framework's capabilities, several areas for future enhancement were identified. Opportunities for enhancement include the use of embedding-based algorithms for semantic understanding of feedback, and complementing with the capability of language models to derive deeper insights.

Overall, the study's outcomes contribute to better understanding of seniors' concerns and satisfaction, which might not have been previously detected through manual effort. A self-help tool has been implemented which enhances SGO's operational efficiency and reduce man-hours spent on data analysis. This novel approach is also significant to uncover emerging trends from the older adults to allow us to better appreciate and address new expectations from the cohort effects of the fast-ageing population in Singapore.

References

Athukorala, S., & Mohotti, W. (2022). An effective short-text topic modelling with neighbourhood assistance-driven NMF in Twitter. *Social Network Analysis and Mining, 12*(1), 89. https://doi.org/10.1007/s13278-022-00898-5

Egger, R., & Yu, J. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. *Frontiers in Sociology, 7*. https://doi.org/10.3389/fsoc.2022.886498

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (4th ed.). Advanced Analytics, LLC.

Hendricks, R. (n.d.). *What is a good accuracy score in Machine Learning?* Deepchecks. https://www.deepchecks.com/question/what-is-a-good-accuracy-score-in-machine-learning/

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159.

Mifrah, S., & Benlahmar, E. H. (2020). Topic Modeling Coherence: A Comparative Study between LDA and NMF Models using COVID'19 Corpus. *International Journal of Advanced Trends in Computer Science and Engineering*. https://doi.org/10.30534/ijatcse/2020/231942020

NPTD Singapore. (n.d.). *Longevity*. National Population and Talent Division, Singapore. https://www.population.gov.sg/our-population/population-trends/longevity/

Purpura, A. (2018). Non-negative Matrix Factorization for Topic Modeling. *CEUR Workshop Proceedings.* Bertinoro, Italy.

Svensson, K., & Blad, J. (2020). *Exploring NMF and LDA Topic Models of Swedish News Articles.* Uppsala, Sweden: Uppsala University.

Appendix

Table 4
Panel of Evaluators

SN	Description of Role	Title	Years of Service
Evaluator 1	User rep for IT systems of SGO	Senior Assistant Director	More than 3 years
Evaluator 2	User rep for IT systems of SGO	Assistant Manager	1 – 3 years
Evaluator 3	Media, and Communications rep for SGO to manage communication materials	Senior Executive	1 – 3 years
Evaluator 4	Strategy rep for SGO, liaise with external stakeholders	Assistant Manager	More than 3 years
Evaluator 5	Deputy Head for Satellite Office managing ground outreach operations of SGO	Manager	More than 3 years
Evaluator 6	Deputy Head for Satellite Office managing ground outreach operations of SGO	Manager	More than 3 years
Evaluator 7	Deputy Head for Satellite Office managing ground outreach operations of SGO	Manager	More than 3 years
Evaluator 8	Deputy Head for Satellite Office managing ground outreach operations of SGO	Manager	More than 3 years
Evaluator 9	Deputy Head for Satellite Office managing ground outreach operations of SGO	Senior Manager	More than 3 years
Evaluator 10	Deputy Head for Satellite Office managing ground outreach operations of SGO	Senior Manager	More than 3 years

Table 5

Extracted Topics for Each Area-Sentiment Category With NMF Topic Modelling

Area of Feedback	Topics Extracted for Concern Sentiment	Topics Extracted for Happy Sentiment
Digital Literacy	1) app_healthhub_use_health_ government app 2) fast_digitalisation_catch_ digitalisation fast_pace 3) savvy_online_payment_service_ government 4) scam_afraid_government_worry_ afraid scam 5) english_understand_chinese_ language_learn	1) good_educate_government_ government good_good government 2) learn_slow_computer_good_ thankful 3) app_health_life_like_book 4) savvy_savvy concern_concern scam _generation_difficult 5) lesson_glad_glad government_ conduct lesson_conduct 6) course_organise_government good_ good organise_organise course 7) use_payment_need_use mobile_cash 8) government_handle_public_fast_ process 9) skill_future_skill future_fund_future fund 10) help_voucher_help learn_ government help_appointment 11) singapore_accessible_internet_ world_fast 12) phone_mobile phone_mobile_ coaching_smart phone 13) keen_keen learn_know_competent_ usage 14) scam_wary_concern scam_savvy concern_government 15) digitalisation_come_book_ appointment_book appointment 16) class_good_attend_conduct_want 17) care_content_government_ government good_smart 18) online_singpass_provide_ implement_make payment 19) function_want_want learn_initiative _handphone
Finances	1) cost living_living_living high_cost_ living increase 2) food expensive_food_expensive_ market_costly 3) cost live_live_cost_live high_high cost 4) government_hope_hope government_help_government help 5) price_food price_food_price increase_high food 6) living cost_cost high_living_ cost_high	1) care_government care_ government_care citizen_citizen 2) receive_sss_receive sss_cash bonus _bonus 3) voucher_cdc_cdc voucher_ receive cdc_help 4) support_financial_government_ financial support_provide 5) payout_government payout_ government_appreciate_appreciate government 6) package_assurance package_ assurance_satisfied_satisfied

	7) sss_inflation_eat_inflation eat_eat sss 8) increase_increase cost_gst_gst increase_cost 9) support_financial support_ financial_support government_ government 10) high_high cost_inflation_coe_living high 11) transport_transport voucher_ voucher_receive_eligible 12) rise_rise cost_rise price_living rise_price 13) voucher_cdc_cdc voucher_gst_gst voucher 14) subsidy_government subsidy_flat_ stay_room 15) expense_medical_living expense_ expense high_medical expense 16) cpf_work_job_age_payout 17) chas_card_chas card_blue_blue chas 18) assistance_financial_financial assistance_sso_help 19) expensive_thing_thing expensive_ housing_living expensive 20) income_household_household income_low_base	assurance 7) gstv_gstv receive_recently_receive recently_receive
General Feedback	1) cost live_live_cost_high_high cost 2) living_cost_high_cost living_living cost 3) government_hope_hope government_help_support 4) complaint_town_council_town council_neighbour	1) government_good government_ government scheme_thank_overall 2) government good_good_good job_ job_good government 3) government care_care_care elderly_ elderly_care elder 4) neighbour friendly_friendly_ neighbour_friendly environment_ facility good 5) policy_government policy_policy good_policy care_content 6) care citizen_citizen_government citizen_government_support citizen 7) appreciate_appreciate government_ financial handout_financial_handout 8) support_government support_support scheme_support government_scheme 9) satisfied_satisfied government_ overall satisfied_overall_ government government 10) help_government help_help elder_ elder_help citizen 11) good care_care_elderly_care

		elderly_benefit 12) payout_government payout_payout yearly_yearly_receive 13) environment_facility_facility good_ environment facility_friendly environment 14) look_government look_look people_ people_good look 15) visit_sga_sga visit_appreciate visit_ appreciate 16) subsidy_medical_medical subsidy_government subsidy_sss 17) singapore_country_safe_singap ore safe_live 18) neighbour good_neighbour_good_ good neighbour_good government 19) recognise_job_good job_indonesia_ proud singaporean
Healthcare Services and Schemes	1) increase medical_medical cost_increase_medical_cost 2) expensive_medication expensive_ medication_cost expensive_fee expensive 3) subsidy_hope_government_ho pe government_hope subsidy 4) high_cost_cost high_high cost _medical cost 5) expense_medical expense_ medical _expense high_fee 6) polyclinic_long_wait_time_ appointment 7) card subsidy_low_subsidy help _help low_help 8) chas_card_chas card_blue_blue chas 9) medisave_use_use Medisave_ pay_cash	1) discount supermarket_ supermarket _discount_ supermarket pay_pay 2) good_government_care_ government care_service 3) satisfied_privilege_satisfied privilege_privilege government_ privilege subsidy 4) support_government support_ government_support subsidy_ subsidy 5) subsidy_medical_health care_ health_care