International Journal of Community Well-Being (2026) 9:33
https://doi.org/10.1007/s42413-026-00305-3

ORIGINAL RESEARCH ARTICLE

Machine Learning for Community Well-Being: Identifying Factors Affecting Resilience in Disaster-Prone Regions

Nasrin Sultana[1] · Md Sarwar Kamal[2] · Ho Leung IP[3] · Michael Bewong[2] · Md Zahidul Islam[2] · Azizur Rahman[2]

Received: 18 March 2024 / Accepted: 29 June 2026

Abstract

This paper presents a machine learning approach to identify factors affecting community well-being in New South Wales (NSW), Australia, based on regional well-being survey (RWS) data. The study focuses on NSW, which has recently experienced multiple types of disasters, and utilizes machine learning techniques to identify indicators that can help prevent the decline of community well-being using community users' responses to survey data. However, obtaining labelled data for this type of study is challenging, so the paper introduces a recommender-based approach that utilizes randomly generated instances, eliminating the need for labels to identify factors. The primary objective is to investigate the factors that affect the current level of community well-being. The study identifies these factors using a recommender-based model and highlights the adaptability of this model across regional well-being survey datasets, indicating its potential to make significant contributions to the field of disaster resilience and enhance community well-being.

Keywords Community well-being · Machine learning · Correlation coefficient · Resilience · Exploratory descriptive analysis

Introduction

Disasters, for instance, droughts, bush fires, storms, and floods, often affect not only individuals but the whole community, which could be a region, a state or territory, or even a nation. In recent decades, there has been a concerning increase in the frequency and severity of climate-related disasters (Leaning & Guha-Sapir, 2013). In order to prepare for the adverse impacts brought by disasters, it is imperative to strengthen the well-being of a community to improve the overall resilience (Graugaard, 2012; Jenkins, 2011; Lei et al., 2014). Within Australia, the impact of disasters tends to be

Extended author information available on the last page of the article

Published online: 18 August 2026

uniquely localised. The readiness after disasters, coping mechanisms, and recovery prospects of each local community depend on the presence of diverse assets within their community. These assets encompass local community groups, governmental bodies, service providers, businesses, meeting spaces, health services, transportation, infrastructure, and services. The ability of a community to build resilience against disasters and restore community well-being depends on the capabilities and capacities of these local assets, as well as the support they can access from external organizations before, during, and after disastrous events.

According to Khalili et al. (2015), a community is defined as a group of people who share common interests. Well-being is a comprehensive concept encompassing the living conditions and quality of life experienced by individuals in the present (referred to as current well-being) and the resources that contribute to sustaining well-being over time, which include natural, economic, human, and social capitals. Community well-being involves the ability of a community to support a high quality of life for its residents, encompassing the provision of services, governance, amenity, social connections, among many other things (Schirmer et al., 2021; UNDRR, 2020). Schirmer et al. (2021) defined three highly inter-related concepts: community recovery, community resilience and community well-being. Having higher resilience by definition contributes to more rapid and effective recovery from disaster, which in turn supports community well-being. However, these associations are not one-way or linear. A recovery process may act to successfully restore well-being and build resilience to future disasters; or may, if not successful, result in both lower well-being and resilience due to recovery requiring many community members to draw down on their resources (UNDRR, 2020). At the community level, well-being plays a key role in building resilience, where communities come together to support each other through difficult times (Mental Health Commission of NSW, 2023).

To promote community well-being, factors that affect community members' quality of life need to be identified. Naturally, these are often factors which individuals are likely to express dissatisfaction, which may arise from different domains such as health services, transportation, communication, telecommunication and livability. Through identifying and understanding factors leading to low well-being, policy makers can take informed actions to improve the status, which may include provision of essential services and amenities, effective governance, and promotion of robust social connections among community members (Schirmer et al., 2021; UNDRR, 2020). Community well-being is especially important for resilience against disasters, which pose significant challenges to societies worldwide. The comprehensive framework for disaster resilience highlights the importance of embracing a holistic approach that includes evaluating risks, developing capacities for adaptation, implementing response strategies, engaging communities, and incorporating technological advancements (Homeland Security, 2023). This approach guides policymakers, practitioners, and communities to enhance their resilience efforts, reduce vulnerabilities, and promote sustainable development in the face of disasters.

Community well-beings are often measured through surveys. For example, in Australia, the Regional Wellbeing Survey (RWS) has been run for ten years, generating an enormous amount of valuable insights. The survey respondents were asked to rate various aspects, including health services, transportation, communication, telecommunication, livability, governance, and more. However, with over several hundred questions and lots of missing values, the analysis of these kinds of surveys is challenging. In addition, a clear label representing the level of well-being is often not available, meaning that supervised machine learning techniques may not be suitable. Therefore, in order to analyze community well-being through these kinds of survey data, unsupervised machine learning approaches that can handle missing values are required (Dara et al., 2002).

In fact, in recent years, machine learning techniques have found extensive application in personal well-being and resilience domain, including flood risk management, bush-fire response, and healthcare systems, on a global scale (Chang et al., 2022; Kumar et al., 2022; Pabreja et al., 2022). These advanced techniques have also been leveraged in the development of mental health monitoring systems that rely on ubiquitous sensing technologies and machine learning algorithms (Flesia et al., 2020) to track and manage stress. Similarly, machine learning has been employed to detect stress among university students, contributing to a deeper understanding of mental health (Ahuja & Banga, 2019), and identifying suitable metrics for evaluating disaster resilience and risk reduction (Klainin-Yobas et al., 2021; Wikipedia Contributors, 2023).

The primary objective of this study is to identify factors affecting community well-being in the New South Wales (NSW) region of Australia by analyzing RWS data using machine learning approaches. We have employed a recommendation-based machine learning approach that incorporates principles from collaborative filtering recommender systems, which can automatically impute missing values based on similarity between users. Based on the community members' responses to the survey, the model can identify key factors that lead to loss in community well-being. As the model stems from user-item recommender systems, we will use the term "user" and "community member" interchangeably. Like-wise, "item" may be used to refer to a variable, and "user-item interaction" may be used to refer to the rating given by a community member to a specific question. We demonstrate that incorporating collaborative filtering principles effectively addresses scenarios involving incomplete user-related interaction information in recommendation systems. This integration improves recommendation diversity.

The rest of this paper is structured as follows. Section "Current Trend in Well-being Analysis" provides some background and describes earlier approaches that have been employed. Section "Methods: Social Filtering (SF) Model" describes the proposed models. Sections "Simulation Study" and "Analysis of RWS Data" describe the experiments used to evaluate the performance of our models. Finally, Sect. "Conclusion and Future Recommendation" presents our conclusions.

Current Trend in Well-being Analysis

Community well-being is a multidimensional concept encompassing various aspects of individuals' lives within a community. Understanding and enhancing community well-being have gained significant attention in recent years, with machine learning playing an increasingly pivotal role in this field. In this brief literature review, we examine the current trends in research related to community well-being using machine learning techniques, and we identify gaps in the existing literature that need further exploration.

Current trend in the personal well-being analysis is most of the time predictive models (Jaidka et al., 2020; Kutty et al., 2022; Makridis et al., 2021). Mukherjee et al. (2021) demonstrated a data-driven approach to study the complex associations between mental health of adults within a metropolitan community and the physical and socio-economic aspects of the built environment. They proposed to figure out how well-designed neighbourhoods could make adults in the community feel better and created a framework called a "multi-level scenario-based predictive analytic framework" to study these connections. Makridis et al. (2021) used machine learning to create models that predict the physical health and overall well-being of veterans by considering three types of factors: demographic, socio-economic, and geographic. Kido and Swan (2016) discussed how machine learning and personal genetics could help happiness research. Specifically, they used deep learning and other machine learning methods to predict social intelligence levels from genetic profiles.

Another trend is the sentiment analysis to assess well-being. Social media platforms have emerged as valuable sources of real-time information during disasters. Artificial intelligence and machine learning techniques could be used to identify emergency messages on social media during a natural disaster, and categorize tweets relevant to first responders (Powers et al., 2023). Sentiment analysis techniques can help gauge public sentiment, identify urgent needs, and assess community resilience based on user-generated contents (Sarvari et al., 2018; Cao et al., 2022; Kankanamge et al., 2020; Zhou & Zhang, 2021). This information can inform emergency responders and facilitate targeted interventions.

Machine learning techniques have been applied to assess community resilience by analysing diverse data sources (Ainuddin and Routray, 2012; Cutter et al., 2014; Sherrieb et al., 2010; Zou et al., 2018). These sources may include socioeconomic data, infrastructure information, environmental indicators, and historical disaster records. Many methods have been developed recently for predicting natural disasters such as drought, hurricanes, floods, and earthquakes (Linardos et al., 2022; Ogie et al., 2018; Wu & Cui, 2018). These systems leverage historical data, sensor networks, and satellite imagery to forecast the occurrence and severity of disasters.

Recently, another branch of unsupervised approaches, named recommendation-based machine learning approaches, emerged. Despite the extensive literature on recommender systems spanning various domains, there is a notable lack of emphasis on disaster resilience and community well-being recommendations. A recent study by Alcaraz-Herrera et al. (2022) introduced an innovative recommendation framework that employs evolutionary algorithms as the primary recommendation engine. This approach views the generation of personalized well-being recommendations as a multi-objective optimization problem, capturing the interconnectedness of different well-being aspects. The framework constructs configurable recommendations in the form of bundled items with dynamic properties, taking into account both user preferences and a predefined well-being goal.

Generally speaking, two common techniques exist within the domain of personalized recommender systems: content-based filtering and collaborative filtering. Content-based recommender systems require manual feature engineering, thus demanding a significant amount of domain knowledge. In the opposite, collabotrative filtering does not require subject matter knowledge as the embeddings can be learned automatically. In other words, contextual features are not required. Collaborative filtering has the advantage of making recommendations solely based on ratings. However, it also has limitations. For instance, items without any ratings cannot be recommended, and accurate recommendations are challenging for users with only a few rated items. Our proposed model attempts to fill this gap by calculating the similarity between users to impute missing values. As a final output, our proposed model can recommend personalized variables to different groups of people, offering a more adaptive and user-specific recommendation system.

We believe that this new line of research in this field can be developed to recommend factors to indicate factors that may lead to low well-being in the community. As a result, in this paper we propose a recommender system to indicate priority factors for the organizations and stakeholders.

Methods: Social Filtering (SF) Model

This section describes our proposed models. The SF model uses the filtering method (Wang et al., 2018), which is usually based on collecting and analyzing information on users' behaviours, activities, or preferences, and predicting missing ratings based on the similarity with all users in the community. The model is composed of several steps:

Step 1 The similarities between users are calculated using cosine similarity.
Step 2 The k nearest neighbours are determined based on the similarity measure. Here, k can be specified by the researcher.
Step 3 Missing values are imputed based on ratings provided by the k nearest neighbours obtained from the previous steps.
Step 4 Variables with the lowest average ratings are extracted.

Figure 1 shows a schematic flow of the model. Below we elaborate the steps in greater details.

In the first step, the SF model calculate the similarity between users. For an user A, let $r_{A,i}$ be the response from user A on variable i, and $I(A)$ be the set of variables which A has provided responses (i.e., non-missing). By means of centred cosine similarity (Chowdhury, 2010; Widiyaningtyas et al., 2021), which is closely related to the Pearson's correlation coefficient, the similarity between two users A and B can be calculated as

$$sim(A,B) = \frac{\sum_{i\in I(A)\cap I(B)}(r_{A,i}-\bar{r}_A)(r_{B,i}-\bar{r}_B)}{\sqrt{\sum_{i\in I(A)\cap I(B)}(r_{A,i}-\bar{r}_A)^2}\sqrt{\sum_{i\in I(A)\cap I(B)}(r_{B,i}-\bar{r}_B)^2}}, \quad (1)$$

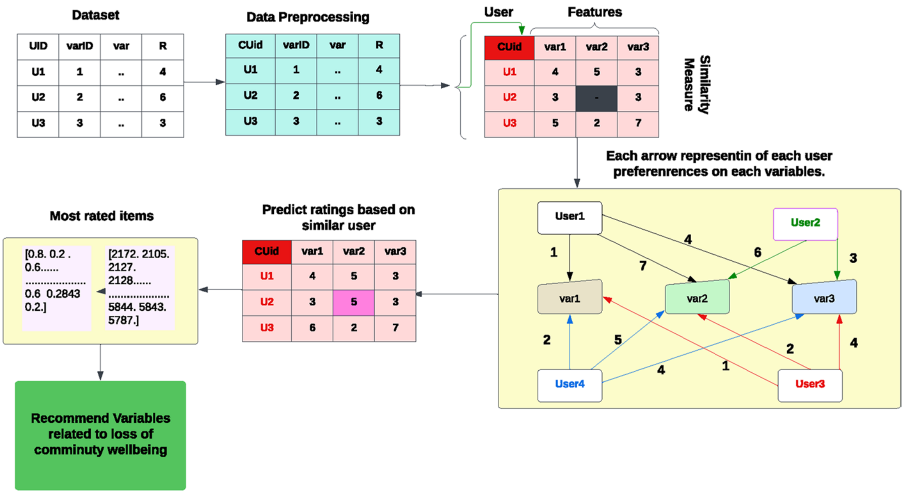

Fig. 1 Social Filtering (SF) Model

where $\bar{r}$ denotes the average of all ratings within the set $I(A) \cap I(B)$. Since the similarity is calculated over the set of variables both users have rated, the similarity may not be reliable when the number of co-rated items is small. In this case, significance weighting can be used such that

$$sim(A, B) = \frac{min(|I(A) \cap I(B)|, \gamma)}{\gamma} sim(A, B), \tag{2}$$

for some predefined threshold parameter γ. The same is repeated for each pair of users. In the second step, for each user, the k other users with the highest degree of similarity are defined as the nearest neighbours. The missing values are imputed in Step 3 based on the ratings available among the neighbours. Specifically, suppose a missing value was provided by user x for item i, the predicted value of $r_{x,i}$ is calculated as:

$$P(r_{x,i}) = \bar{r}_x + \frac{\sum_{u \in S(x)} sim(x, u)(r_{u,i} - \bar{r}_u)}{\sum_{u \in S(x)} sim(x, u)}, \tag{3}$$

where $S(x)$ be the set of users who are neighbours of x. In the final step, all ratings are aggregated across users and variables. In the procedure, the model investigates how many users interacted with most of the variables, which helps us understand which variables are responded to by the users. The model goes through each of the users' ratings for each variable to generate factors with either the highest or lowest ratings, depending on context. For example, in community well-being studies, the SF model can recommend factors with the lowest ratings which are associated with low community well-being.

Simulation Study

This section is intended to demonstrate the performance of the SF model under controlled conditions using an artificial dataset, mimicking a representative sample of data.

Simulated Dataset

For the randomly generated dataset, a data generator was employed, which simulated user ratings for the different variables. These ratings were generated on a scale ranging from 1 to 7, indicating the users' interactions with each variable. The generator ensured that the ratings were randomly distributed throughout the dataset, mirroring real-world variations in user preferences. In total, the dataset consisted of twenty-five variables, each assigned a unique identifier. These 25 variables were assumed to come from three domains: infrastructure and services, natural hazard resources, and livability. Table 1 shows the variables. We have categorized the variables into three distinct domains, with each domain comprising two indicators. In Table 1, first column represents variable names, second column represents rating range, third column represents domain names, fourth column represents the indicator names and fifth column the true rankings generated. To enhance visual clarity and differentiation, we have assigned unique colours to each domain and indicator in the Table 1. Variables with a lower ranking received lower rates. To compare with our model we generated the true ranking in simulated dataset. For example, "Mental Health Service" was scored predominately 1 or 2 by the users while "Long waits for tradies" was scored predominately 6 or 7 by the users.

Table 1 Dataset description

Variables	Range	Domains	Indicators	True ranking
Mental health service	1 to 7	Infrastructure and services	Access to health	1
Community disaster preparedness	1 to 7	Natural hazard resources	Preparedness	2
Self-rated community safety	1 to 7	Livability	Crime and safety	3
Access public transport	1 to 7	Infrastructure and services	Infrastructure	4
General health service	1 to 7	Infrastructure and services	Access to health	5
Special health service	1 to 7	Infrastructure and services	Access to health	6
Self-rated disaster preparedness	1 to 7	Natural hazard resources	Preparedness	7
Hazard risk perception	1 to 7	Natural hazard resources	Hazard risk perception	8
Self-rated coping skills	1 to 7	Natural hazard resources	Preparedness	9
Self-rated ability to cope with disaster	1 to 7	Natural hazard resources	Preparedness	10
Self rated safety	1 to 7	Livability	Crime and safety	11
Hard to get rent	1 to 7	Livability	Housing	12
Hard to get accommodation	1 to 7	Livability	Housing	13
Self rated general health	1 to 7	Infrastructure and services	Access to health	14
Job Opportunity	1 to 7	Livability	Crime and safety	15
House price increases	1 to 7	Livability	Housing	16
Trust in leaders	1 to 7	Natural hazard resources	Preparedness	17
Community alcohol abuse	1 to 7	Livability	Crime and safety	18
Community drug abuse	1 to 7	Livability	Crime and safety	19
Community domestic violence	1 to 7	Livability	Crime and safety	20
Psychological distress level Community	1 to 7	Natural hazard resources	Hazard risk perception	21
Local groups get things done	1 to 7	Natural hazard resources	Preparedness	22
Involved in local decision making	1 to 7	Infrastructure and services	Infrastructure	23
Actively contribute	1 to 7	Infrastructure and services	Infrastructure	24
Long waits for tradies	1 to 7	Infrastructure and services	Infrastructure	25

The generator produced random ratings for these variables, representing varying levels of user satisfaction. To add a layer of realism to the data, we incorporate missing ratings for a subset of users. This is achieved by introducing a chance that a rating is set to None. This None value signifies a missing or unspecified rating. A Python code is designed to create the synthetic ratings across 1000 individual users. It introduces missing ratings for few users, assigns unique variable names and IDs, and generates a range of ratings based on specific criteria. The resulting dataset is then saved in a CSV file.

Experimental Setup

For our SF model, first we do the training and test split as 80% and 20% respectively. In these experiments, the two hyper-parameters k and γ were set as 5 and 20, respectively. These values were found to perform well based on our preliminary studies. In other words, the similarity measure was adjusted when the number of co-rated items between two users was less than 20. Based on the similarity measure, 5 neighbours were defined for each user. We have demonstrated the Hyper-parameter Sensitivity Analysis of the SF model. (details are reported in the supplementary.)

The dataset and corresponding code are available on the GitHub page located at the following address: (https://github.com/nasrin21/resilience).

Simulation Results

In this section, we show our experimental results. Our model recommends three different levels of variables, such as , low medium and high. For clarity, in the figures below, red colour boxes represent low well-being, the blue colour boxes represent average well-being, and the green colour boxes represents high well-being.

Factors Associated with Low Well-Being

Table 2 shows the top ten factors extracted by the SF model, as well as their original ranking [1]. It can be observed that the model accurately identified the same set of variables. Although the ranks provided by the SF model do not always match with the original ranks, the first two factors "Mental health service" and "Community disaster preparedness" were identified in the correct order. Even without any labels in the dataset, this result clearly demonstrates the superior performance of the SF model in recommending key factors that influence community resilience. Variables in each of the domains can also be assessed in greater details, as illustrated below.

Infrastructure and Services Domain In this specific domain, there are two indicators that are used: 'access to health' and 'telecommunication and infrastructure'. These indicators are further connected to specific variables. The visual representation in Fig. 2 displays the well-being level for each variable. The red colour boxes indicate factors such as general health, special health and mental health that were related to

[1] Described in Section "Simulated Dataset"

Table 2 Top Factors related to low well-being

Factors	True Ranking	SF Ranking
Mental health service	1	Mental health service [1]
Community disaster preparedness	2	Community disaster preparedness[2]
Self-rated community safety	3	Access public transport[4]
Access public transport	4	Special health service[6]
General health service	5	General health service[5]
Special health service	6	Self-rated ability to cope with disaster[10]
Self-rated disaster preparedness	7	Self-rated coping skills[9]
Hazard risk perception	8	Hazard risk perception[8]
Self-rated coping skills	9	Self-rated disaster preparedness[7]
Self-rated ability to cope with disaster	10	Self-rated community safety[3]

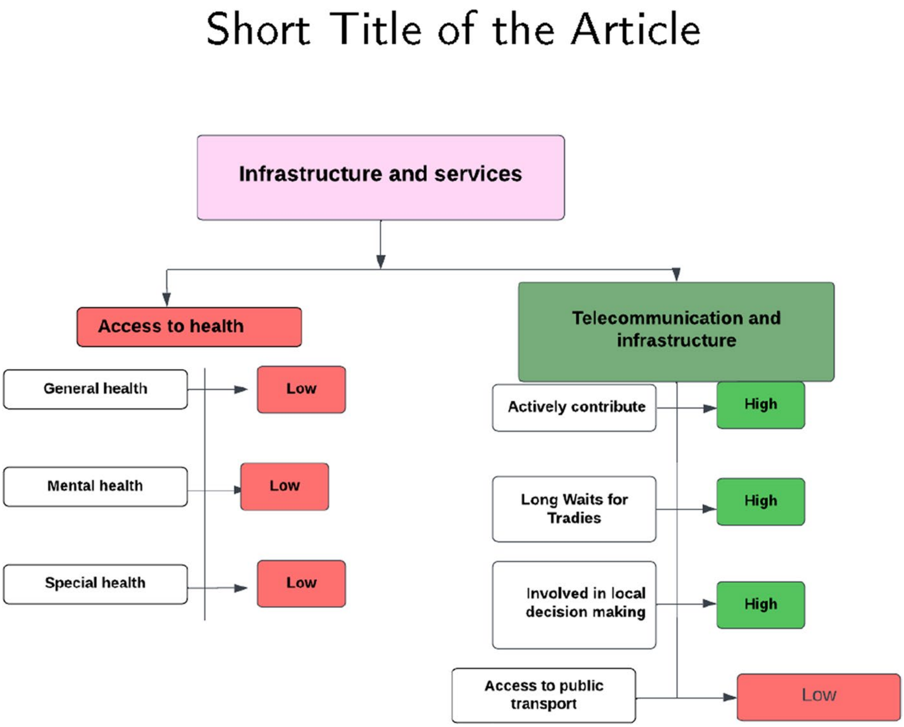

Fig. 2 Well-being levels based on simulated data: Infrastructure and services domain

low well-being. On the other hand, in telecommunication and infrastructure, few variables in green colours are identified as factor related to high community well-being based on random dataset.

Livability Figure 3 illustrates the impact of variables within the livability domain on well-being. In this domain, there were two indicators: 'crime and safety' and 'housing'. The housing sector indicated an average level of well-being. On the other hand, self-rated community safety was one of the factor related to low well-being based on random data.

Natural Hazard Resources In the domain of natural hazard resources, there were two key indicators considered: 'preparedness' and 'hazard risk perceptions'. Figure 4 demonstrates that all variables within preparedness and hazard risk perceptions were related to low well-being.

Correlation Analysis Based on Simulated Data

Based on the results of SF modelling, it is evident that both the natural hazards domain and infrastructure domain exhibit notably lower levels of well-being. Furthermore, using correlation techniques, we have discovered significant relationships among certain variables, namely general health, special health, mental health, self-rated ability to cope with disaster, community disaster preparedness, and self-rated disaster preparedness (Fig. 5). In essence, this correlation analysis assesses how users interact with these variables, and it becomes apparent that all four of these factors share low ratings compared to other factors. These findings serve as clear warning signs for community well-being. In conclusion, our extensive analysis recommends

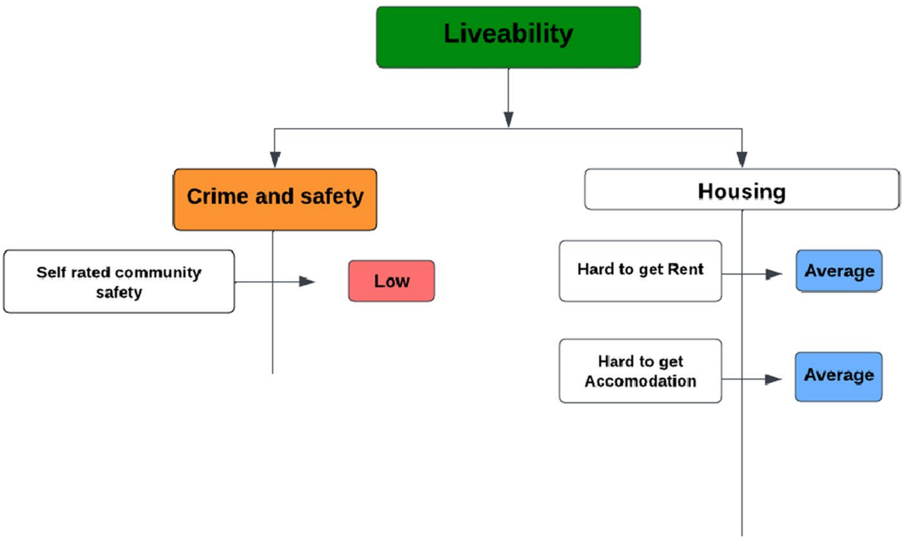

Fig. 3 Well-being levels based on simulated data: Livability domain

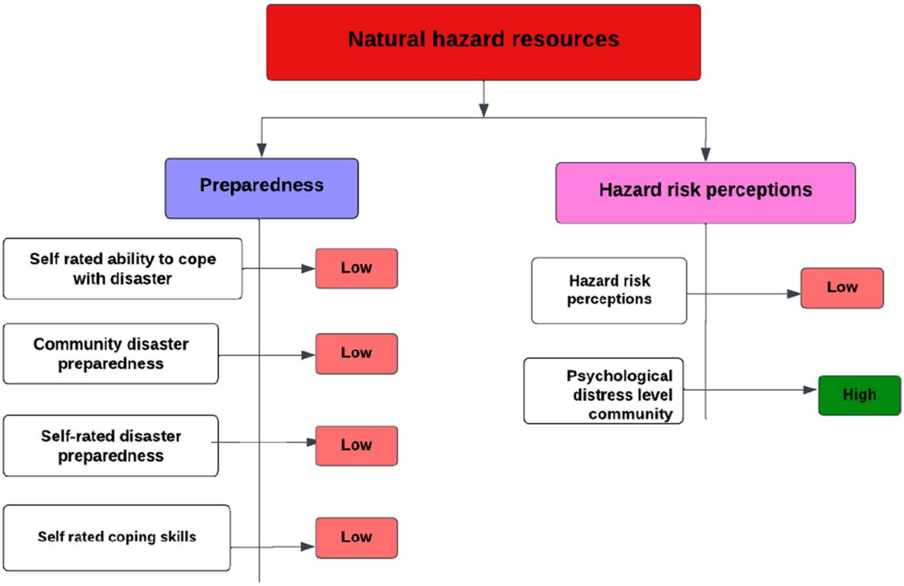

Fig. 4 Well-being levels based on simulated data: Natural hazard resources domain

the substantial influence of these two domains, represented by the indicators, on the outcomes of well-being.

Analysis of RWS Data

In this section, we present the findings based on the analysis of the data collected from Regional Well-being Survey in 2021. The next subsections describe the dataset, experimental setup and the results of the proposed models.

RWS Dataset (2021)

In this section a brief overview of the key attributes of the RWS are presented. The RWS is a survey focused on residents living in rural and regional areas of Australia. Rural and regional areas are defined as those areas located outside the cities of Sydney, Melbourne, Adelaide, Perth, Canberra, and Brisbane. Its scope thus includes people living on rural properties, in small rural and regional towns, and in larger regional cities. We focused on NSW region based on the tenth wave of the survey (2021). The objectives of the RWS are to better understand the well-being of individuals living in, and communities located in, rural and regional areas; and identify how this well-being is influenced by different types of change common to rural and regional areas. Thus, the survey has a strong focus on remote, rural and regional areas, with less focus on urban areas. When examining challenging and positive events, the focus is on both individual, household, and community-scale events experienced, with a particular focus on understanding resilience to climatic variability and change in

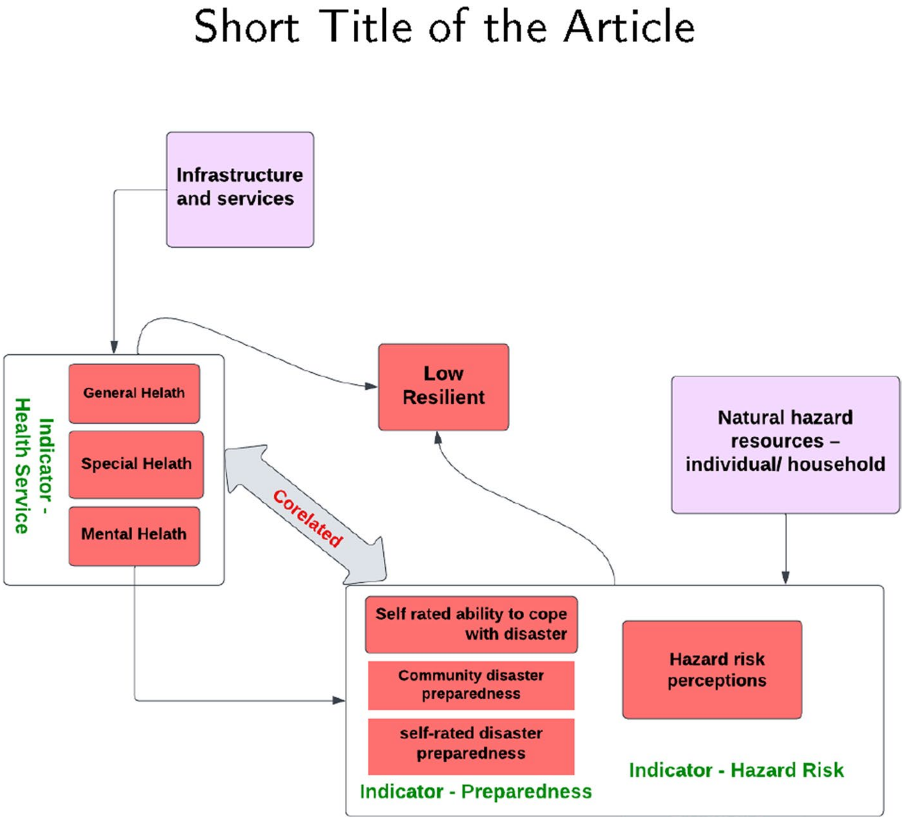

Fig. 5 Co-related community well-being affected factors

natural resource dependent industries, particularly agriculture. For the RWS dataset we have number of domains that classifies as various indicators. For each indicator there were many variables. Since some of the variable names in the dataset were not immediately clear to readers, we have provided further descriptions of the variables in Supplementary material.

Factors Affecting Community Well-Being in the NSW Region

With the objective to identify the key factors that related to low well-being of the community, we focused on five specific domains that we believe have a substantial impact on the well-being: infrastructure and services, livability, skill and capacity, natural hazard resources and institutional resources. The SF modelling was performed using the same setup as described in Sect. "Experimental Setup".

This analysis delves into the examination of factors influencing community well-being in the NSW region, categorizing them into three levels: low, average, and high. We illustrate the analysis in two ways. First, the primary aim is to provide a comprehensive understanding of the current factors that communities are coping with, highlighting those that have a low, average, or high impact on well-being based on the 2021 RWS data in Figs. 6, 7, 8, 9 and 10, Sect. "Comprehensive Insights: Assessing the Impact of Current Factors on Varied Community Coping Levels". Second, aim is to extract two top most priority variables in terms of low well-being from each indicator based on the 2021 RWS data in Tables 3, 4, 5, 6 and 7. By conducting this investigation, we enable stakeholders to gain insights into the challenges that communities are currently facing, as well as the factors that contribute to community satisfaction. This knowledge serves as a valuable resource for informed decision-making and targeted efforts to enhance community well-being in the NSW region.

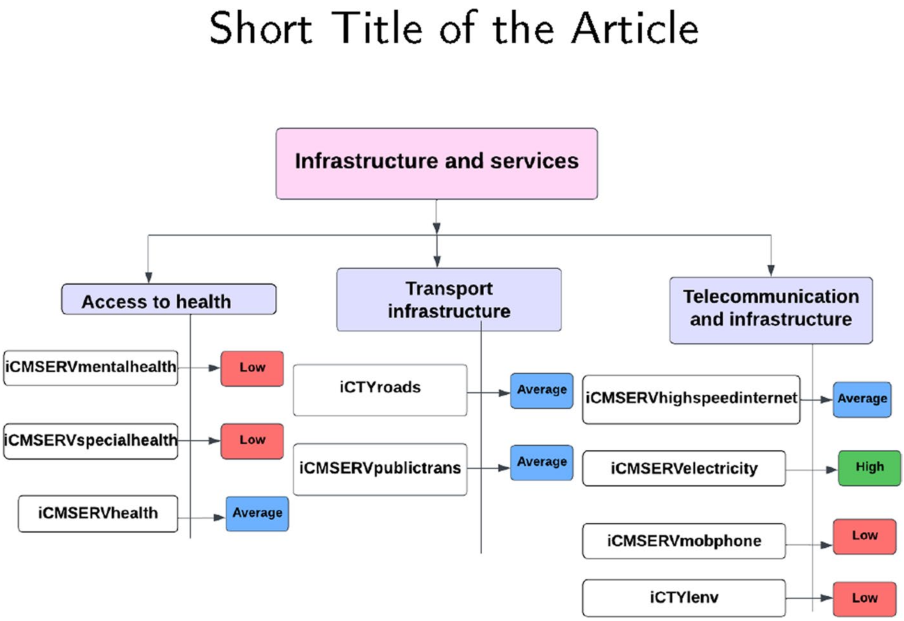

Fig. 6 Well-being levels based on 2021 RWS dataset: Infrastructure and services domain

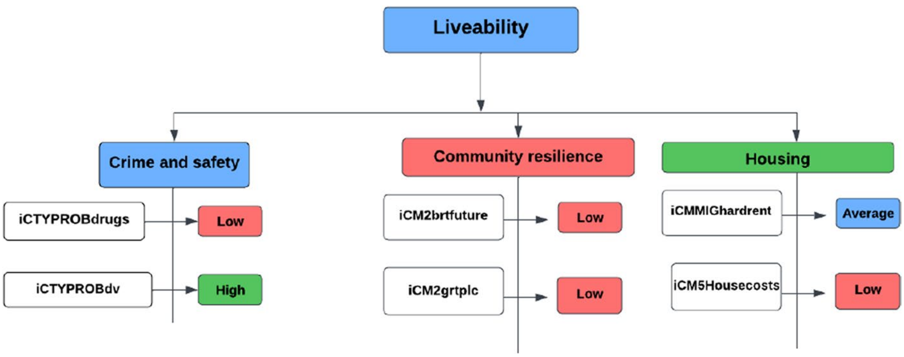

Fig. 7 Well-being levels based on 2021 RWS dataset: Livability domain

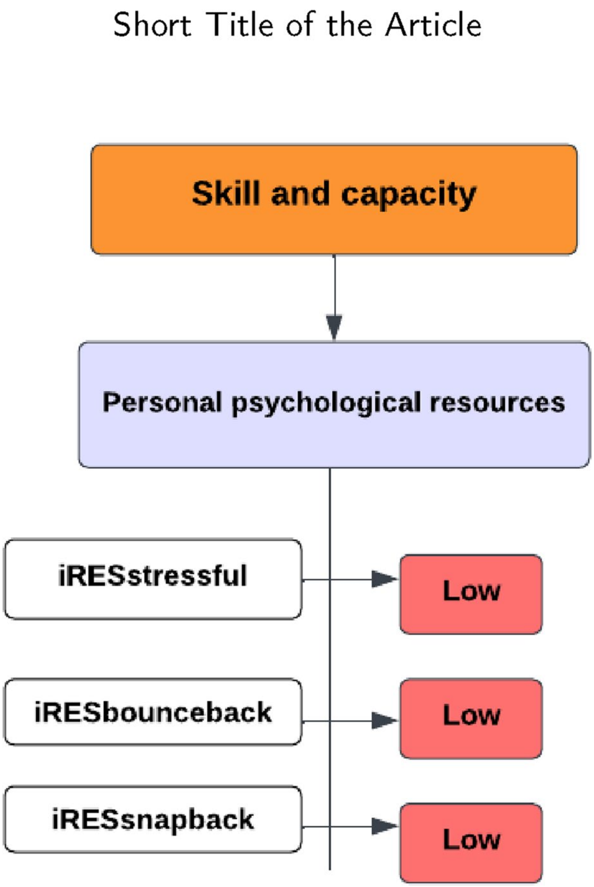

Fig. 8 Well-being levels based on 2021 RWS dataset: Skill and capacity domain

Comprehensive Insights: Assessing the Impact of Current Factors on Varied Community Coping Levels

This section, we show our model recommends three different levels of factors from each domain, such as, low, average and high.

Infrastructure and Services Domain: Priority Factors In the infrastructure and services domain, we have examined three key indicators: 'access to health,' 'transport infrastructure,' and 'telecommunication and infrastructure.' These indicators are further linked to specific variables. Figure 6 shows a visual representation of the well-being levels associated with each of these variables.

The elements highlighted in red boxes, such as mental health service (iCMSERVmentalhealth) and special health service (iCMSERVspecialhealth), are factors that have a noticeable impact on community well-being. However, in the case of transport infrastructure, factors like roads (iCTYroads) and public transport (iCMSERVpublictrans) have a more moderate influence (blue box) on well-being.

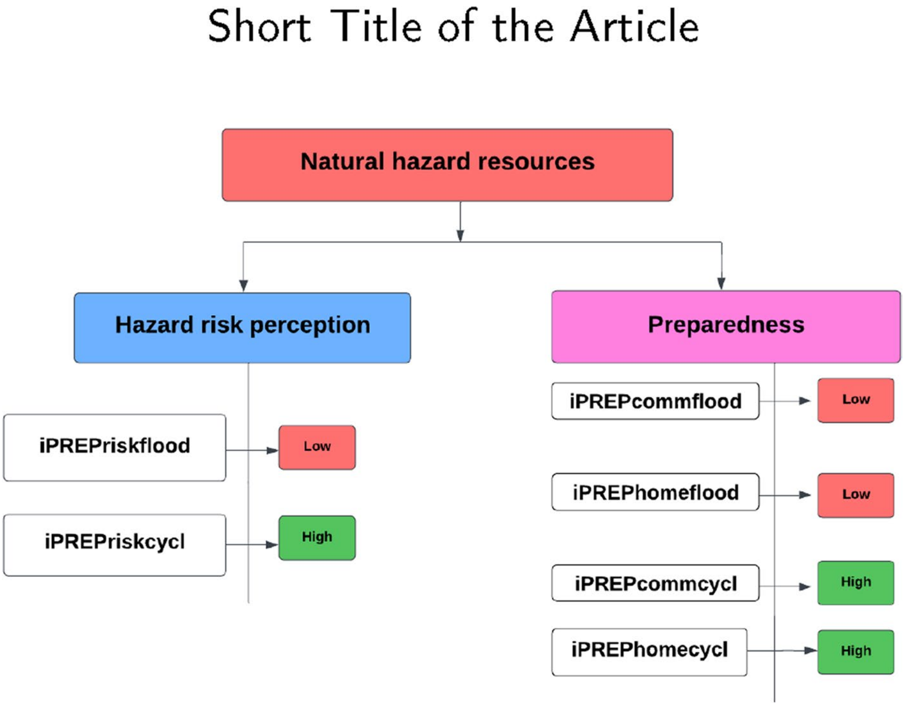

Fig. 9 Well-being levels based on 2021 RWS dataset: Natural hazard resources domain

In the telecommunication and infrastructure indicator, community opinions about variables are mixed. Factors related to the health of the local environment and mobile phone reception (iCMSERVmobphone) seem to be high-priority concerns for low well-being. The speed of internet access (iCMSERVhighspeedinternet) raises average concern. Conversely, the availability of electricity (iCMSERVelectricity) is one of the factors that community users express satisfaction with.

Livability In the livability domain, three indicators were looked at: 'crime and safety,' 'community resilience,' and 'housing.' As can be seen in Fig. 7, crime and safety, as well as the housing, show a combination of well-being outcomes.

However, when it comes to community resilience, it's a different story. In particular, community resilience is a factor related to lower well-being, which is shown by the red boxes in the Fig. 7. In this domain, community resilience is a top concern when it comes to lower community well-being, representing an essential factor that local government should pay attention to.

Skill and Capacity Domain There was only one indicator, personal psychological resources, within this domain. From Fig. 8, it can be visualized that all variables within this domain indicated a low level of well-being.

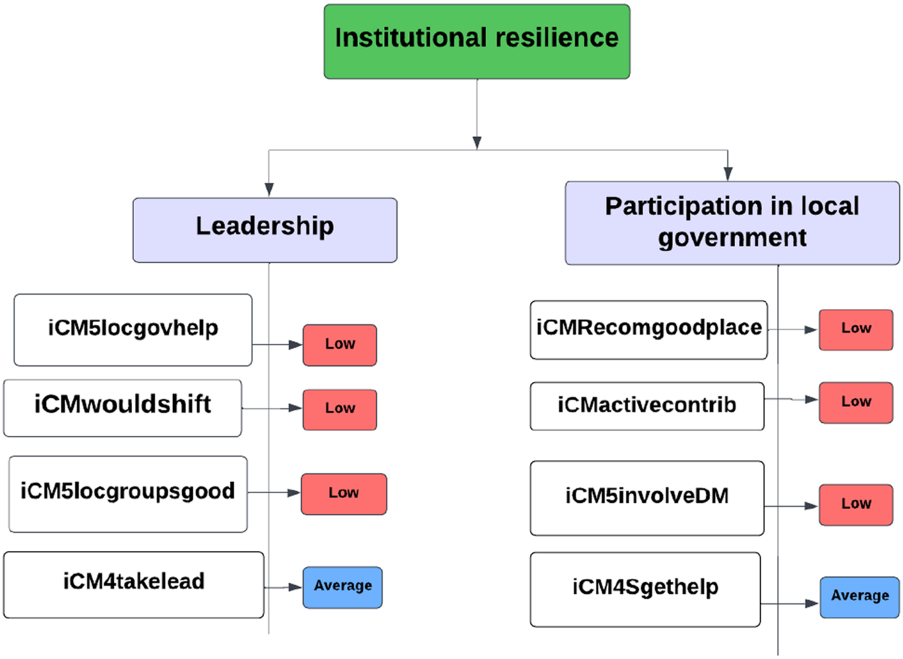

Fig. 10 Well-being levels based on 2021 RWS dataset: Institutional resilience domain

Table 3 Priority Factors: Infrastructure and services

Domain Name	Indicator Name	Priority Ranking	Variable
Infrastructure and services	**Health**	1	iCMSERVmentalhealth
Infrastructure and services	**Health**	2	iCMSERVspecialhealth
Infrastructure and services	**Transport infrastructure**	1	iCTYroads
Infrastructure and services	**Transport infrastructure**	2	iCMSERVpublictrans
Infrastructure and services	**Telecommunications infrastructure**	1	iCMSERVmobphone
Infrastructure and services	**Telecommunications infrastructure**	2	iCMSERVhighspeedinternet

Table 4 Priority Factors: Livability

Domain Name	Indicator Name	Priority Ranking	Variable
Livability	**Crime and safety**	1	iCTYPROBdrugs
Livability	**Crime and safety**	2	iCTYPROBalcohol
Livability	**Community resilience**	1	iCM2brtfuture
Livability	**Community resilience**	2	iCM2grtplc
Livability	**Housing**	1	iCM5Housecosts
Livability	**Housing**	2	iCM5Otherlivcosts

Table 5 Priority Factors: Skill and capacity

Domain Name	Indicator Name	Priority Ranking	Variable
Skill and capacity	**Personal psychological resources**	1	iRESstressful
Skill and capacity	**Personal psychological resources**	2	iRESbounceback
Skill and capacity	**Personal psychological resources**	3	iRESrecover
Skill and capacity	**Personal psychological resources**	4	iRESsnapback

Table 6 Priority Factors: Natural hazard resources

Domain Name	Indicator Name	Priority Ranking	Variable
Natural hazard resources	**Preparedness**	1	iPREPhome ood
Natural hazard resources	**Preparedness**	2	iPREPcomm ood
Natural hazard resources	**hazard risk perception**	1	iPREPrisk ood
Natural hazard resources	**hazard risk perception**	2	iPREPriskbf

Table 7 Priority Factors: act as early warning signals for domain: Institutional resilience

Domain Name	Indicator Name	Priority Ranking	Variable
Institutional resilience	**Leadership**	1	iCM5locgovhelp
Institutional resilience	**Leadership**	2	iCMwouldshift
Institutional resilience	**Participation in local government**	1	iCMRecomgoodplace
Institutional resilience	**Participation in local government**	2	iCMactivecontrib

This may suggest that, after any disaster strikes, the community finds it challenging to recover and get back on their feet. They are experiencing stress because they lack the necessary skills and capacity to cope with the aftermath of these disasters.

In essence, this domain is crucial to examine because it points to a significant decrease in community well-being. It highlights the fact that the community is struggling to bounce back and regain their stability after facing any kind of disaster.

Natural Hazard Resources Domain In the domain of natural hazard resources, as indicated by the 2021 RWS data, we focused on two important indicators: 'preparedness' and ‘hazard risk perceptions’. Figure 9 shows how these indicators relate to community well-being.

Our analysis reveals an interesting dynamic. The community is concerned about the risk of floods, which is associated with lower community well-being (iPREPrisk-flood). They are worried that if a flood were to occur, their community would struggle to cope with it (iPREPcommflood). Additionally, they feel that their own households are ill-prepared to handle a flood (iPREPhomeflood). This vulnerability to floods contributes to a decline in community well-being.

Conversely, when it comes to cyclones, the community does not perceive a significant risk (iPREPriskcycl). Since cyclones have not occurred recently, they do not feel the need for specific preparations in this regard. Therefore, their well-being is not currently affected by concerns related to cyclone preparedness.

Institutional Resilience Domain Within this domain, we have focused on two key indicators: ‘leadership’ and ‘participation in local government’. The results can be seen in Fig. 10. What becomes evident from this data is that several variables within these indicators are associated with lower well-being levels. This emphasises the need for heightened attention to the leadership sector.

Additionally, it is important to note that the community members were expressing dissatisfaction with the current level of local government participation. This suggests that they need to improve in the involvement and engagement of the local government in community affairs, especially following disasters.

Of course, these domains are interrelated, as shown in Fig. 11. Below we show the results of more in-depth analysis within each domain. The primary objective of this further analysis is to identify and extract the key factors that have a significant impact on community well-being. In order to achieve this, we conducted a comprehensive analysis of all variables within each domain associated with specific indicators, and

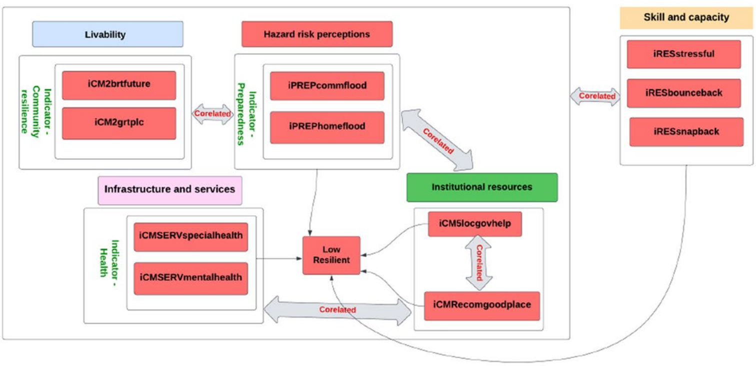

Fig. 11 Correlated community well-being affected factors in the NSW region

recommend the top two variables that stakeholders may wish to prioritize to improve. The results, presented in Tables 3 to 7, provide valuable insights into the most influential factors related to each indicator and domain. In these tables, the third column assigns a priority ranking to the variables, with ‘1’ denoting the highest priority factor impacting community well-being. Finally, the fourth column presents the variable names, where further clarification on the meaning of each variable can be found in Supplementary material.

Prioritizing Key Variables from Each Indicator that Impact Community Well-Being

In this section, we present the top two key variables recommended by our model, which impact community well-being across indicators, based on the 2021 RWS data.

Infrastructure and Services Domain Our findings, as presented in Table 3, highlight key areas where improvements are needed to enhance community well-being. Notably, there is a need for enhancements in mental health services and special health services to improve overall well-being. Additionally, in the context of transportation, community users have expressed dissatisfaction with road services and public transport, indicating a need for improvements in these areas. Furthermore, in the telecommunication infrastructure area, the analysis shows that the significance of improvements needed for mobile phone reception and high-speed internet services.

Livability Domain Table 4 provides valuable insights into the livability domain and pinpoints the top two factors from each indicator that exhibit the strongest correlation with early signs of well-being decline within the community.

Within the Crime and Safety indicator, it is evident that the stakeholders need to address the challenges related to drug and alcohol problems. Within the Community Resilience indicator, the variables iCM2brtfuture and iCM2grtplc offer important insights. They reflect the community’s concern regarding the future and current status of the community. In the context of the Housing indicator, the factors of housing cost and living cost stand out as the priority elements affecting well-being in the livability domain.

Skill and Capacity Domain Table 5, represents the top factors that has strongest relationship to early warning signs of well-being loss for the community from the personal psychological resources in this domain. In this domain our models recommend most of the factors related to the personal psychological resources indicator. This analysis reveals specific physiological resources that organizations and stakeholders should focus on to improve community well-being, based on the data from 2021.

Natural Hazard Resources Domain Table 6 presents the leading factor with the strongest relationship to early signs of well-being loss within the community, specifically within the domain of Natural Hazard Resources. This analysis demonstrates compare to other natural disasters flooding is a significant concern within the community, and the lack of preparedness after floods is contributing to well-being challenges within the community.

Institutional Resilience Domain Table 7 presents the leading factor with the strongest relationship to well-being loss within the community, specifically within the domain of institutional resilience. In this analysis we considered the participation in local government and leadership indicators.

In the leadership indicator category, community are facing challenge to trust in leaders. The top low resilience factors in this indicator is local government is not able to help community when face challenges. Another factors need to consider if they could, they would shift to live in another community.

In the participation in local government indicator category, community users indicate they will not recommend this community to other. Factors such as not actively contribute and no involvement in local decision making emerge as the critical variables showing a substantial correlation with well-being decline in the institutional resources domain.

Discussion

This study introduces the Social Filtering (SF) recommender-based method as a practically feasible approach to identify early-warning factors associated with declines in community well-being in regional settings. Applied to the 2021 NSW RWS data, SF consistently highlighted several policy-relevant issues: a) access to mental-health and specialist health services, b) community and household preparedness for flooding, and c) reduced institutional trust and active participation in local governance. These findings have three immediate implications for resilience policy.

First, the identification of mental-health access and specialist-health services among the highest-priority factors suggests targeted investment is needed in local mental health capacity (e.g., mobile mental-health clinics, telehealth expansions for rural areas), and that monitoring of mental-health access should be treated as a core resilience indicator. Second, the prominence of flood-preparedness variables indicates that resilience programming in NSW should prioritize household- and community-level flood readiness—this may include subsidized home-scale mitigation, community-based preparedness workshops, and enhanced early-warning communications targeted to the most vulnerable in New South Wales. Third, signals of weak institutional trust and low participation suggest policy levers beyond infrastructure are required: improved two-way engagement with communities, transparent post-disaster assistance processes, and mechanisms to co-design recovery and preparedness interventions will likely increase both wellbeing and resilience outcomes.

Beyond immediate operational actions, the SF model itself can support policy as a repeatable monitoring tool: by running SF on successive RWS waves or local surveys, stakeholders can track which variables rise into the top-priority set and allocate resources adaptively. However, we caution that SF outputs are dependent on the survey instrument and sampling frame; therefore, policy use should be complemented with qualitative community engagement and administrative data to validate and refine recommendations. Finally, we discuss limitations and future directions in the next Section.

Conclusion and Future Recommendation

A machine learning approach is used to identify factors affecting community well-being in New South Wales, Australia, based on regional well-being survey data. The study uses randomly generated instances and machine learning techniques to identify indicators that can help prevent the decline of community well-being, without the need for labels. The primary objective is to investigate the factors that affect the current level of community well-being. The findings indicate that the proposed model is adaptable across regional well-being survey datasets and has the potential to make significant contributions to the field of disaster resilience and enhance community well-being.

Future research should extend this study by incorporating longitudinal datasets to capture temporal variations in community well-being, as the current analysis is limited to cross-sectional data from 2021. The integration of multi-year data would enable the development of predictive models and early warning systems to detect declines in well-being. Given the absence of labeled data, future work could explore semi-supervised learning approaches to combine limited labeled instances with large volumes of unlabeled survey data, thereby improving model validation and interpretability. In addition, incorporating external data sources, such as environmental, economic, and real-time data, could provide a more holistic understanding of community well-being. Finally, future studies should validate the proposed framework in different geographic regions and develop stakeholder-oriented decision-support systems to translate analytical findings into practical interventions to improve community resilience and well-being.

Appendix A

Hyper-Parameter Sensitivity Analysis

Table 8 presents the sensitivity analysis conducted for the two hyper-parameters of the SF model, the size of the neighborhood (k) and the significance-weighting threshold γ (Sarwar et al., 2001). The performance of the model was assessed (Adomavicius & Zhang, 2012). using two complementary criteria: (a) mean absolute error (MAE) on canceled entries in the simulated data set, and (b) rank stability of the top 10 low-wellbeing variables in the RWS data.

Across the grid of values tested, the results show that smaller neighborhood sizes (e.g., k = 3) produce relatively higher MAE and reduced rank stability, reflecting increased variance in local similarity estimates. Very large neighborhoods (e.g., k = 20) also perform poorly, with the highest MAE and weakest stability, suggesting an

Table 8 Sensitivity analysis for SF hyper-parameters (k and γ). Performance is evaluated on simulated data (MAE on held-out entries) and on RWS data (rank stability of top-10 low-wellbeing variables). Values shown are illustrative placeholders

k	γ	MAE (Simulated)	Rank Stability (%)	Top-Variables Overlap	Notes
3	10	0.62	78	7/10	Higher variance in neighborhoods
3	20	0.59	82	8/10	Moderate stability
5	10	0.55	84	8/10	Good accuracy, stable ranking
5	**20**	**0.53**	**90**	**9/10**	Selected default parameters
5	50	0.57	86	8/10	Over-penalizes sparse pairs
10	20	0.60	80	7/10	Oversmoothing across users
20	20	0.66	70	6/10	Neighborhood too large

over-smoothing effect that reduces observable heterogeneity. In contrast, k = 5 consistently produces lower error and stronger rank persistence across different γ values.

The interaction between k and γ also follows the expected pattern: With low γ, unreliable pairs have too much influence, while overly high γ thresholds (e.g. $\gamma = 50$) penalize sparse co-ratings too strongly, slightly degrading performance. The combination k = 5 and $\gamma = 20$ achieves the best overall results, achieving the lowest MAE (0.53), the highest stability of the rank (90%), and the greatest overlap in the variables of highest priority (9 out of 10). Consequently, this setting was selected as the default configuration. Overall, the table demonstrates that the SF model is robust to moderate variations in the hyper-parameters while clearly favouring a small neighborhoods size and a mid-range significance threshold.

Appendix B

Variable Description

In this section, we provide a comprehensive description of all the variables contained within the RWS dataset. Specifically, Tables 9, 10, 12, 11 and 13, we present a structured overview of these variables for the readers' reference.

The layout of these tables is designed to enhance understanding. The first column showcases the variable names employed in this paper, allowing for clear identification. In the second column, we provide the questions posed to the survey respondents, shedding light on the context and purpose of each variable. To facilitate a comprehensive understanding of the variables, the third column offers detailed descriptions and explanations pertaining to their significance and relevance. This systematic presentation serves to enhance the readers' comprehension of the variables within the RWS dataset.

Table 9 Variable description - Infrastructure and services

Variable name	Question seen by respondent	Description of variable
iCTYlenv	Health of the local environment	Local environment
iCMSERVmentalhealth	Access to mental health services e.g. psychologist, psychiatrist	Acc. mental health
iCMSERVspecialhealth	Access to specialist health services (other than mental health)	Acc. spec. health
iCMSERVhealth	Access to general health services e.g. GPs, drop-in centres	Acc. gen. health
iCMSERVprofserv	Professional services e.g. accountants, lawyers	Acc. prof. serv
iCTYroads	Quality of local roads	Quality of local roads
iCMSERVmobphone	Mobile phone reception	Acc. mob. phone
iREClongwaits	Long waits for tradespeople or other types of specialised workers have caused delays in rebuilding and recovering from this event	Long waits for tradies
iRECfinrec	I am financially recovered from the event/events	Recovered financially
iREClongime	It took several months after experiencing the disaster to realise I might need to ask for emotional or psychological support	Long time to ask for support
iCMSERVhighspeedinternet	Access to high speed, reliable internet	Acc. internet
iRECfinassnotlast	Available financial assistance didn't last long enough after the event/events	Financial assistance didn' last long
iCTYsafe	Safety of the local area	Safety of the local area
iCMSERVlocalschools	Quality of local schools	Quality of local schools

Table 10 Variable description - Livability

Variable name	Question seen by respondent	Description of variable
iCTYPROBdv	To what extent are the following problems challenges in your community at the moment?	Domestic violence
iCTYPROBdrugs	To what extent are the following problems challenges in your community at the moment?	Drug abuse
iCTYPROBalcohol	To what extent are the following problems challenges in your community at the moment?	Alcohol abuse
iCTYPROBcrime	To what extent are the following problems challenges in your community at the moment?	Crime
iCM4copewell	How are your local economy, government and community groups going at the moment?	Community copes with challenges
iCM2grtplc	What are your views about the community you live in?	Community great place
iCM2brtfuture	What are your views about the community you live in?	Community has a future
iCM2proudlive	What are your views about the community you live in?	Proud to live here
iCM2commspirit	What are your views about the community you live in?	Good community spirit

Table 11 Variable description - Natural hazard resources

Variable name	Question seen by respondent	Description of variable
iPREPriskflood	I live in a region with a high risk of floods.	Risk of flood
iPREPriskcycl	I live in a region with high risk of cyclones	Risk of cyclone
iPREPriskbf	I live in a region with a high risk of bushfires	Risk of bushfire
iPREPcommflood	My community more generally is well prepared for floods if they happen	Well prepared for flood
iPREPcommbf	My community more generally is well prepared for bushfires if they happen	Well prepared for bushfire
iPREPcommcycl	My community is well prepared for a cyclone if one happens	Well prepared for cyclone
iPREPcommstorm	My community is well prepared for storms or damaging winds if they happen	Well prepared for storm
iPREPcommbetter	The community I live in is better prepared to cope with future disasters now than it was five years ago	Overall community is better prepared for future disasters
iPREPhomeflood	My household is well prepared for floods if they happen	Household well prepared for flood
iPREPhomecycl	My household is well prepared for a cyclone if one happens	Household well prepared for cyclone

Table 12 Variable description - Skill and capacity

Variable name	Question seen by respondent	Description of variable
iRESsnapback	It is hard for me to snap back when something bad happens	Resilience - hard to snap back
iRESbounceback	I tend to bounce back quickly after hard times	Resilience - bounce back
iRESstressful	I have a hard time making it through stressful events	Resilience - hard time making it through stressful events
iRESrecover	It does not take me long to recover from a stressful event	Resilience - not long to recover
iRESlittletroub	I usually come through difficult times with little trouble	Resilience - little trouble getting through difficult times
iRESgetover	I tend to take a long time to get over set-backs in my life	Resilience - take a long time to get over set-backs

Table 13 Variable description - Institutional resources

Variable name	Question seen by respondent	Description of variable
iCM5locgovhelp	My local government is able to help our community face challenges?	Local government faces challenges
iCM5kdecisrep	The people who make decisions for my community represent the whole community, not just part of it.	Whole community represented
iCM5involveDM	I can get involved in local decision-making processes if I want to	Can get involved in local decision making
iCMactivecontrib	I actively contribute to discussion and decision making in my local region e.g. local government, school councils, or business groups	Actively contribute
iCM5locgroupsgood	Local groups and organizations here are good at getting things done.	Local groups get things done
iCMwouldshift	If I could, I would shift to live in another community.	Would shift
iCMfairgo	Most people get a fair go around here	Most people get fair go
iCM4copewell	This community copes pretty well when faced with challenges	Community copes with challenges
iCMRecomgoodplace	I would recommend my community to others as a good place to live	Recommend community to others

Acknowledgements We would like to express our sincere gratitude to the University of Canberra (UC) for their instrumental support, including supplying the Regional Well-being Survey (RWS) data during the Resilience Data Analysis Project (RM-No.:103803). The project, generously funded by the NSW Southern Drought Hub at Charles Sturt University within the Department of Agriculture, Water, Fisheries and Lands, has benefited greatly from the information provided by UC. The RWS data was used in this research with the Charles Sturt University Human Research Ethics Committee approval reference number: H23784. The UC colleagues' discussion has greatly enhanced our understanding of key concepts such as the definition of resilience and the different classifications of natural disasters.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability Data will be made available on request.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Adomavicius, G., & Zhang, J. (2012). Stability of collaborative filtering recommendation algorithms. *Citeseer, 10*(1.221), 7584.

Ahuja, R., & Banga, A. (2019). Mental stress detection in university students using machine learning algorithms. *Procedia Computer Science, 152*, 349–353.

Ainuddin, S., & Routray, J. K. (2012). Community resilience framework for an earthquake prone area in Baluchistan. *International Journal of Disaster Risk Reduction, 2*, 25–36.

Alcaraz-Herrera, H., Cartlidge, J., Toumpakari, Z., Western, M., & Palomares, I. (2022). Evorecsys: Evolutionary framework for health and well-being recommender systems. *User Modeling and User-Adapted Interaction, 32*(5), 883–921.

Alipour Sarvari, P., Nozari, M., & Khadraoui, D. (2019). The potential of data analytics in disaster management. In *Industrial engineering in the Big Data Era: Selected papers from the global joint conference on industrial engineering and its application areas* (*GJCIE 2018*) (pp. 335–348), June 21–22, 2018, Nevsehir, Turkey. Springer.

Cao, H., Wang, M., Su, S., & Kang, M. (2022). Explicit quantification of coastal cultural ecosystem services: A novel approach based on the content and sentimental analysis of social media. *Ecological Indicators, 137*, 108756.

Chang, V., Ganatra, M. A., Hall, K., Golightly, L., & Xu, Q. A. (2022). An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. *Healthcare Analytics, 2*, 100118.

Chowdhury, G. G. (2010). *Introduction to modern information retrieval.* Facet Publishing.

Cutter, S. L., Ash, K. D., & Emrich, C. T. (2014). The geographies of community disaster resilience. *Global Environmental Change, 29*, 65–77.

Dara, R., Kremer, S. C., & Stacey, D. A. (2002). Clustering unlabeled data with SOMs improves classification of labeled real-world data. In *Proceedings of the 2002 international joint conference on neural networks (IJCNN'02)* (pp. 2237-2242). (Cat. No.02CH37290), Neural Networks, 2002. 3. https://doi.org/10.1109/IJCNN.2002.1007489

Flesia, L., Monaro, M., Mazza, C., Fietta, V., Colicino, E., Segatto, B., & Roma, P. (2020). Predicting perceived stress related to the covid-19 outbreak through stable psychological traits and machine learning models. *Journal of Clinical Medicine, 9*(10), 3350.

Graugaard, J. D. (2012). A tool for building community resilience? A case study of the lewes pound. *Local Environment, 17*(2), 243–260.

Homeland Security (2023). National mitigation framework. [Online; accessed 26 Jun 2023].

Jaidka, K., Giorgi, S., Schwartz, H. A., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2020). Estimating geographic subjective well-being from twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences, 117*(19), 10165–10171.

Jenkins, W. O. (2011). *Emergency management: preliminary observations on FEMA's community preparedness programs related to the national preparedness system: Congressional testimony*. DIANE Publishing Inc.

Kankanamge, N., Yigitcanlar, T., Goonetilleke, A., & Kamruzzaman, M. (2020). Determining disaster severity through social media analysis: Testing the methodology with south east Queensland flood tweets. *International Journal of Disaster Risk Reduction, 42*, 101360.

Khalili, S., Harre, M., & Morley, P. (2015). A temporal framework of social resilience indicators of communities to flood, case studies: Wagga Wagga and Kempsey, NSW, Australia. *International Journal of Disaster Risk Reduction, 13*, 248–254.

Kido, T., & Swan, M. (2016). Machine learning and personal genome informatics contribute to happiness sciences and wellbeing computing. In *AAAI Spring Symposium Series*.

Klainin-Yobas, P., Vongsirimas, N., Ramirez, D. Q., Sarmiento, J., & Fernandez, Z. (2021). Evaluating the relationships among stress, resilience and psychological well-being among young adults: A structural equation modelling approach. *BMC Nursing, 20*, 1–10.

Kumar, S., Wajeed, M. A., Kunabeva, R., Dwivedi, N., Singhal, P., Jamal, S. S., & Akwafo, R. (2022). Novel method for safeguarding personal health record in cloud connection using deep learning models. *Computational Intelligence and Neuroscience*.

Kutty, A. A., Wakjira, T. G., Kucukvar, M., Abdella, G. M., & Onat, N. C. (2022). Urban resilience and livability performance of European smart cities: A novel machine learning approach. *Journal of Cleaner Production, 378*, 134203.

Leaning, J., & Guha-Sapir, D. (2013). Natural disasters, armed conflict, and public health. *New England Journal of Medicine, 369*(19), 1836–1842.

Lei, Y., Wang, J., Yue, Y., Zhou, H., & Yin, W. (2014). Rethinking the relationships of vulnerability, resilience, and adaptation from a disaster risk perspective. *Natural Hazards, 70*, 609–627.

Linardos, V., Drakaki, M., Tzionas, P., & Karnavas, Y. L. (2022). Machine learning in disaster management: Recent developments in methods and applications. *Machine Learning and Knowledge Extraction*, *4*(2).

Makridis, C. A., Zhao, D. Y., Bejan, C. A., & Alterovitz, G. (2021). Leveraging machine learning to characterize the role of socio-economic determinants on physical health and well-being among veterans. *Computers in Biology and Medicine, 133*, 104354.

Mental Health Commission of NSW, (2023). Building community wellbeing. [Online; accessed 6 Oct].

Mukherjee, S., Frimpong Boamah, E., Ganguly, P., & Botchwey, N. (2021). A multilevel scenario based predictive analytics framework to model the community mental health and built environment nexus. *Scientific Reports, 11*(1), 17548.

Ogie, R. I., Rho, J. C., & Clarke, R. J. (2018). Artificial intelligence in disaster risk communication: A systematic literature review. In *2018 5th International conference on information and communication technologies for disaster management (ICT-DM)* (pp. 1–8). IEEE.

Pabreja, K., Arya, S., & Madnani, P. (2022). An ensemble machine learning model for automatic prediction of perceived personal well-being of Indian university students during covid-19 lockdown. *MIER Journal of Educational Studies Trends and Practices*, 301–319.

Powers, C. J., Devaraj, A., Ashqeen, K., Dontula, A., Joshi, A., Shenoy, J., & Murthy, D. (2023). Using artificial intelligence to identify emergency messages on social media during a natural disaster: A deep learning approach. *International Journal of Information Management Data Insights, 3*(1), 100164.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web* (pp. 285–295).

Schirmer et al. (2021). Community resilience, wellbeing and recovery project. [Online; accessed 6 Oct 2023].

Sherrieb, K., Norris, F. H., & Galea, S. (2010). Measuring capacities for community resilience. *Social Indicators Research, 99*, 227–247.

UNDRR (2020). Undrr annual report 2020. [Online; accessed 6 Oct 2023].

Wang, Y., Li, C., Tang, J., and Sun, D. (2018). Learning collaborative sparse correlation filter for real-time multispectral object tracking. In *International conference on brain inspired cognitive systems* (pp. 462–472). Springer.

Widiyaningtyas, T., Hidayah, I., & Adji, T. B. (2021). User profile correlation-based similarity (UPCSim) algorithm in movie recommendation system. *Journal of Big Data, 8*, 1–21.

Wikipedia contributors (2023). Grand challenges. [Online; accessed 6 Jun 2023].

Wu, D., & Cui, Y. (2018). Disaster early warning and damage assessment analysis using social media data and geo-location information. *Decision Support Systems, 111*, 48–59.

Zhou, Q., & Zhang, C. (2021). Breaking community boundary: Comparing academic and social communication preferences regarding global pandemics. *Journal of Informetrics, 15*(3), 101162.

Zou, L., Lam, N. S., Cai, H., & Qiang, Y. (2018). Mining twitter data for improved understanding of disaster resilience. *Annals of the American Association of Geographers, 108*(5), 1422–1441.

Authors and Affiliations

Nasrin Sultana[1] · **Md Sarwar Kamal**[2] · **Ho Leung IP**[3] · **Michael Bewong**[2] · **Md Zahidul Islam**[2] · **Azizur Rahman**[2]

✉ Azizur Rahman
azrahman@csu.edu.au

Nasrin Sultana
nasrin21@gmail.com

Md Sarwar Kamal
mdkamal@csu.edu.au

Ho Leung IP
ryan.ip@aut.ac.nz

Michael Bewong
mbewong@csu.edu.au

Md Zahidul Islam
zislam@csu.edu.au

1 RMIT University, Melbourne, Australia

2 School of Computing, Mathematics and Engineering, Charles Sturt University, Port Macquarie, NSW, Australia

3 Auckland University of Technology, Auckland, New Zealand