What is your dependency level? A system to evaluate individual dependency levels depending on the social context and physical activity*

Cristobal Martin[1], Joaquín Ballesteros[2], Arturo Cosano[3], and Luis Miguel Rondón[3]

[1] Universidad de Málaga, Málaga, Spain crimarare@uma.es
[2] ITIS Software, Universidad de Málaga, Málaga, Spain jballesteros@uma.es
[3] Departamento de Psicología Social, Trabajo Social, Servicios Sociales y Antropología Social. Universidad de Málaga, Málaga, Spain arturo@uma.es, luirongar@uma.es

Abstract. Dependency, defined as the need for assistance in daily activities due to physical or cognitive limitations, is a crucial concern in healthcare and social services. Accurately appraising dependency levels is essential for allocating resources, planning care, and enhancing the quality of life for individuals needing support. This work details developing a system to assess and predict individuals' dependency levels through an optimized survey. Initially, a comprehensive scale with 26 items was created by social work experts and tested with 210 users with three different dependency levels recognized and 40 without any dependency. A machine learning technique has been used to find a reduced set of items that predict the dependency level accurately. The results classify the fourth classes with 91 % accuracy using only nine items. This system has been deployed and tested in a web application, so other systems or persons can use it to predict the individuals' dependency levels online.

Keywords: Dependency assessment · Machine learning · Resource Allocation

1 Introduction

As global populations continue to age, the issue of elderly dependency has become increasingly critical in healthcare and social services [14]. Although dependence is not a unique characteristic of old age, the ageing process entails a series of changes that are prone to the person falling into this situation. In reality, it is expected that by the year 2050, the population over 65 will exceed 56.9% of the population in Spain [15]; we must understand that situations of dependency will be abundant in the future. With age, physical and cognitive abilities often

* Work supported by the projects *IRIS* PID2021-122812OB-I00 (co-financed by FEDER funds); *CAMPERO* TED2021-131739B-C21; B.3 aid from the University of Málaga; and *Cátedra de I+D+I para la prevención de la dependencia*

decline, leading to a greater need for daily assistance. This growing dependency places a significant burden on care systems, underscoring the need for accurate methods to assess and predict the support required by elderly individuals.

The study of dependency is a challenge today [6], as it constitutes a multifactorial concept that encompasses the permanent situation in which a person finds themselves due to a disability, age, or illness, which requires the help or attention of a third person to be able to carry out daily activities. Caring for people in a situation of dependency is complex due to the social changes driven by increased longevity and legislative advances that recognize these third-generation social rights. However, research and empirical practice in this area have focused on the study of dependency, evaluating it reactively without addressing the preventive dimension, which is the fundamental cornerstone to mitigate these situations and address them effectively in the future.

The research, development, and innovation chair for the Prevention of Dependency[4] promotes spaces for research and technological and social innovation focused on preventing dependency and personal promotion. They have created a scale that provides researchers with a measurement tool for dependency as an interdependent vector that articulates the triad of social inclusion, health, and quality of life. The general objective is to analyze the effectiveness of prevention in improving the quality of life of dependent individuals, focusing on two aspects:

1. Selection strategy based on criteria with predictive evidence of deterioration: This involves selecting older adults with certain factors with consistent predictions for functional loss or adverse health outcomes according to available literature to evaluate and specifically intervene in them.
2. Selection strategy based on loss of functionality: In this case, the detection of individuals at risk of dependency is based on analyzing their physical-functional capacity before they reach a state of dependency that prevents them from performing activities of daily living (ADL).

The current scale consists of 26 Likert-type items. This instrument is divided into four major blocks, which are further subdivided into the variables that comprise them:

- Health construct: physical, mental, and social health, referring to the general quality of life of the person.
- Social construct: social relationships, architectural barriers, loneliness, and social participation.
- Autonomy construct: degree of personal autonomy, self-determination, and community participation.
- Support construct: institutional support, general support, and social resources available to the person.

Evaluating those 26 items that form the variables of these constructs results in a lengthy evaluation process for a person, with the added difficulty of older

[4] https://catedradependenciauma.es

dependent individuals. For this reason, finding ways to optimize the scale is crucial, and we need to reduce the number of items as much as possible without losing information or compromising the quality of the scale.

Recent advancements in technology, particularly in machine learning, offer promising solutions for improving the assessment of elderly dependency. In other related areas, it has been used to reduce some large questionaries successfully [12, 7]. In this paper, we utilized this initial 26-item scale and employed a genetic algorithm to identify a condensed set of 9 questions capable of accurately predicting dependency levels. We have developed an estimator using a regression tree that notably reduces the time to assess user dependency from around half an hour to some minutes. In addition, this estimator has been integrated into a user-friendly web application, allowing caregivers and healthcare professionals to receive instant predictions regarding an individual's dependency level. This streamlined system ensures the assessment process is easily accessible and reliable, ultimately facilitating more effective resource allocation and care planning.

The rest of the paper is organized as follows: The next section describes the scale reduction and selected items. Then, section 3 presents the web-responsive application of the estimation model and condensed scale into it. Finally, section 4 presents the conclusion and future work.

2 Reducing the number of items in the scale: from 26 to 9 items

During the methodological process of the 26-item scale, which has lasted several months of inferential statistical analysis, no empirical evidence has been found from either inferential or multifactorial analysis that definitively validates the results. It has also not been possible with the structural equation model, which has yet to establish causality between dependency and the indicated variables due to, among other reasons, its multifactorial nature and the large number of items on the scale.

There are several traditional strategies to reduce the number of questions in a survey, ranging from traditional analytical techniques to more advanced machine-learning approaches. In traditional statistics, the dependency between two variables can provide very interesting information for reducing the burden of a questionnaire. Measuring the degree of association and understanding how they relate is fundamental. Primarily, correlation is used to understand this degree of association between components. Analytical methods include factor analysis [16], which identifies underlying relationships between questions and groups of highly correlated items, allowing for the elimination of redundant questions. Reliability analysis, such as Cronbach's alpha [3], evaluates the internal consistency of the survey, allowing for the elimination of the less significant questions. Another example is Pearson's linear correlation coefficient [13], which helps measure the degree of linear association between two variables [9].

Although traditional methods assume linear relationships between variables and can only identify locally optimal solutions, they may overlook superior alternatives. Additionally, they may be less effective in handling complex and non-linear interactions and require strict statistical assumptions that are only sometimes met. The machine learning approaches are based on finding a good question set that provides high-accuracy estimations in a classification model, allowing non-linear relations. In this work, we use a evolutionary algorithm to find a good question set guided by a fitness function that uses a classification tree created in Scikit-learn [10] to estimate how those questions predict the dependency level.

2.1 Dataset

The dataset was obtained from a non-probabilistic quantitative sampling, with a confidence level of 95% and a margin of error of 2.5%. As it was necessary to know the responses of dependent people, an additional selection criterion was that respondents must have a recognized degree of dependency and sufficient cognitive abilities. These surveys were conducted with dependent people over 55 years old in day centers and nursing homes, 210 in total. Additionally, to include a non-dependent population, surveys were also conducted 40 people who had no degree of dependency, who were interviewed at Active Participation Centers in the Province of Málaga.

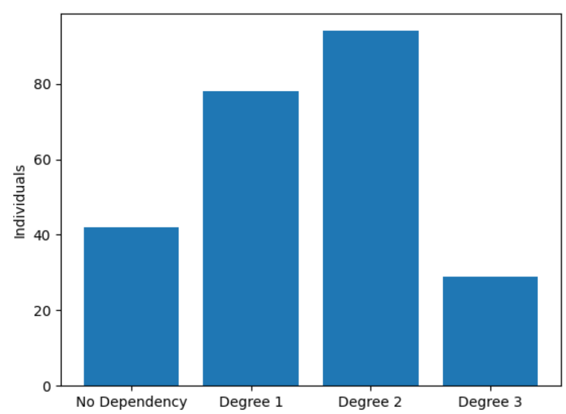

Fig. 1. Number of individuals per dependency level class.

The dataset is not balanced. Fig. 1 shows its distribution. To balance this imbalanced dataset, we have upsampling [8] the minority classes to match the size of the majority class. Three levels of dependency are measured in this work, each with specific characteristics and criteria for support and care. The first level, Moderate Dependency (Degree I), includes individuals who require assistance to perform several basic daily activities at least once a day but who do not need continuous support. The second level, Severe Dependency (Degree II), encompasses individuals who need help two or three times a day to perform several basic daily activities but do not require constant supervision. The third and highest level, Great Dependency (Degree III), is for individuals who need help multiple times a day to perform various basic activities and require continuous presence and support from caregivers.

2.2 Evolutionary algorithm approach

We employed the eaMuPlusLambda evolutionary algorithm from the DEAP library [5] to optimize the selection of survey items. The algorithm was configured with a population size (mu) of 1000 and a lambda parameter of 2000, reflecting the number of offspring produced each generation. The crossover probability (cxpb) and mutation probability (mutpb) were set to 0.7 and 0.3, respectively, to balance exploration and exploitation during the evolutionary process. This setup allowed for effective optimization of the survey item subset, leveraging the evolutionary algorithm's strengths in navigating complex solution spaces.

The individual are vectors with the exact dimensions as the number of items in the survey (26). This binary vector takes a value of one if the item is used to train a model and zero if not. The fitness function in this study is crafted to assess the quality of each candidate solution based on two distinct criteria: the accuracy of a classification tree [1] and the number of selected items. The accuracy is the mean of a 10-fold cross-validation, so the overfitting is reduced. The evolutionary algorithm aims to maximize average accuracy while minimizing the number of items, thus ensuring that the resulting feature subset is highly effective and minimally burdensome.

The evolutionary algorithm found nine items with an 91% average testing accuracy. Fig. 2 shows one of the confusion matrix in the cross-validation. It can be observed that a user with no dependency can be misled by dependency 1. The same occurs for dependency levels 1 to 3, but the miss-classification is always with a class close to the actual value.

The nine items selected are presented in table 1. Fig. 3 shows the decision tree with all the individuals and items selected by the evolutionary algorithm.

2.3 Selected items

The algorithm has been able to decrease the number of items from 26 to 9, prompting the question of whether these remaining items are the most suitable. In the upcoming subsections, we will analyze the influence of each of these items on the individual's dependency.

Could you please indicate what is your health condition? This item is more than obvious that it has come up. The state of health is key to understanding dependency. Health is a binomial associated with the concept of dependency,

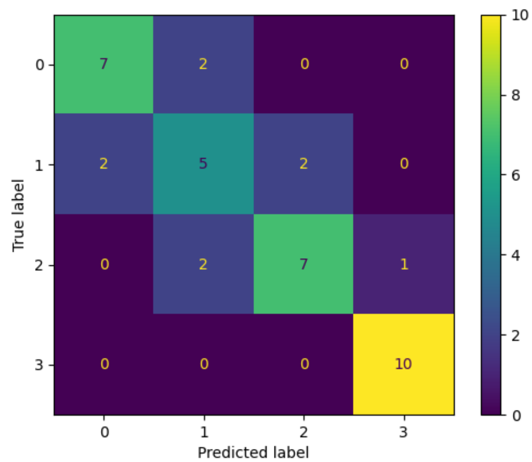

Fig. 2. Confusion matrix of a iteration of the cross-validation loop. Zero is no dependency, and 1 to 3 is the degree of dependency.

the fact of being well. It is understood here as a complete state of physical, mental, and social well-being in addition to the absence of disease (WHO). It should be broken down into two dimensions, physical and mental health, because the social dimension is addressed in the rest of the items.

How many times have you left home in the last month? Dependent people have many difficulties moving outside the home; in fact, it is one of the variables measured to calculate dependency, so it is a good sign that it appears. Social relationships are understood as a fundamental vector in the quality of life that prevents dependency and health or death situations, as is agreed upon in the latest research trends in the field of gerontology and geriatrics.

Is your home in a condition to ensure your well-being? The home is the place where a person develops their life. Being in good condition not only prevents dependency but also ensures a good quality of life for both dependents and non-dependents. A home with poor accessibility conditions and not adapted limits their independence and the ability to carry out daily life activities (such as bathing, eating, dressing, etc.) as well as access to the outside community life.

What is your level of participation in innovative activities? Participating in innovative activities has many physical and mental benefits. Therefore, the Active Participation Centers conduct workshops on innovative activities as prevention mechanisms. These activities involve cognitive stimulation, psychomotor

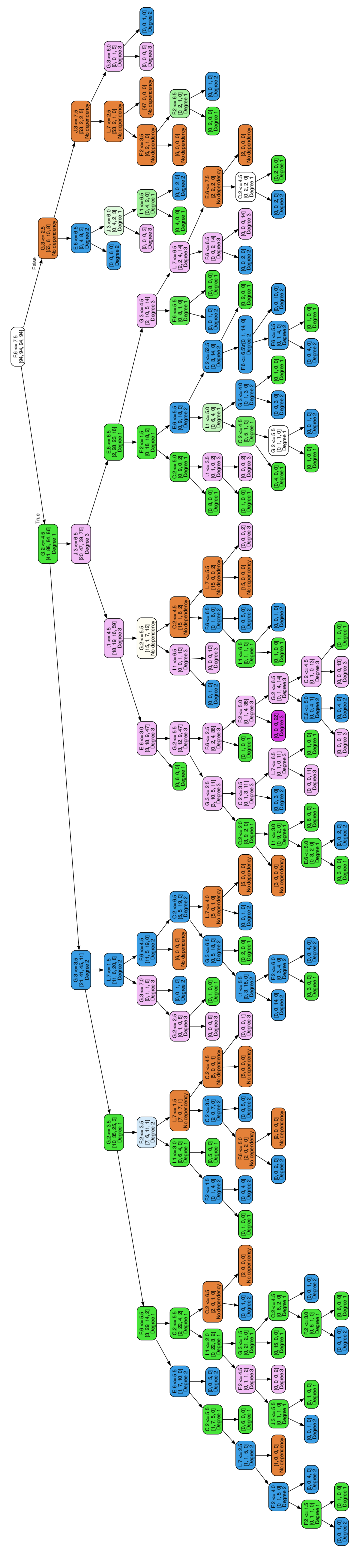

Fig. 3. Decision Tree. Each node shows the condition to go left(true) or right (false), the individuals of each class (from no dependency, class 1 to 3) and the estimated class.

ID	Feature Name
C.2	Could you please indicate what is your health condition?
I.1	How many times have you left home in the last month?
J.3.	Is your home in a condition to ensure your well-being?
L.7	What is your level of participation in innovative activities?
G.2	Indicate the level of care you require to perform the basic necessities of life
E.6	Rate the extent to which they are paying attention to your personal autonomy
G.3	How often do you engage in leisure and cultural activities?
F.6	Indicate the level of involvement of the caregivers.
F.2	Indicate the support you receive from your second-degree family

Table 1. IDs and Likert-type items selected by the evolutionary algorithm.

skills work, mobility, and improvement of social skills. They are usually from the perspective of an active and productive aging approach, adapted to each person's abilities. Innovating certainly involves engaging in creative activities.

Indicate the level of care you require to perform the basic necessities of life Activities of daily living (ADLs) are essential tasks that enable adults to live independently and participate in society. These include basic activities like bathing, dressing, eating, hygiene, personal care, and sleeping, as well as instrumental activities involving social interaction, such as using communication systems, shopping, and managing household affairs. According to Law 39/2006 [4] on the care of people in situations of dependency , the level of dependency is determined by the frequency of required personal autonomy care, with higher dependency levels indicating more assistance needed for ADLs. Those who can perform ADLs without assistance are not considered dependent. ADLs are crucial for the social inclusion and identity of the elderly.

Rate the extent to which they are paying attention to your personal autonomy Regarding the previous item, personal autonomy is the goal of all promotion, care, and prevention of dependency. The resources and services offered to the user aim to guarantee or promote as much as possible the personal autonomy of the person in a situation of dependency. Additionally, those who are not dependent wish to extend their full personal autonomy as much as possible, and prevention is the key element for this. On the other hand, the services provided by the Dependency Law, through social resources (home help, day centers, residences), are essentially aimed at personal autonomy. All surveyed users have these needs met through these resources, and therefore, the deterioration of dependency is prevented.

How often do you engage in leisure and cultural activities? This item is interesting because it hides many nuances. Participation in cultural and leisure activities is synonymous with being able to perform your functions fully, with active inclusion and community participation. This is also a determining factor for the quality of life and the feeling of well-being to prevent health and/or dependency situations. Additionally, it is implied that it is a way to interact with others. Not engaging in activities or maintaining social relationships can lead to unwanted loneliness, which is a powerful agent of deteriorating quality of life and being at risk of suffering dependency.

Indicate the level of involvement of the caregivers A person in a situation of dependency, if properly attended to, will see their quality of life greatly improved. Conversely, if their needs are not met, they risk increasing their degree of dependency. Likewise, a person who is not in a situation of dependency also benefits from proper attention. Appropriate treatment, social skills, and professional care are essential for improving the quality of life of dependent people. It has therefore been verified that social resources for dependency care adequately fulfill these social functions in line with the Law.

Indicate the support you receive from your second-degree family Family support is one of the essential aspects of the construct of social support. In this case, it refers to the bidirectional support in daily living activities, psychological and emotional support, and support for family caregivers who attend to dependent people, understanding the family as an interdependent system among all members and a part of social intervention. In this context, it refers to second-degree relatives. Although first-degree relatives often have a more significant role in the care and attention of family members, this is often assumed to be the case. The support of second-degree relatives can make a difference in better care for the person. Additionally, people who have more contact with their relatives and close ones will be less prone to suffer from unwanted loneliness and its consequences. This idea connects with the latest scientific publications in gerontology and geriatrics, which indicate that second-degree relatives provide greater well-being to the elderly by focusing on sharing family rituals, meeting and leisure spaces, and sharing emotions. These positive symbols are providers of social happiness. First-degree relatives are usually more focused on basic daily living activities and assuming the functions of family leadership, and therefore, do not always contribute proportionally to happiness.

3 Web application

Since the application was meant to be used by a wide spectrum of end users, it was vital to guarantee that anyone could use the app without any impairments. To achieve this, the overall design was kept as simple as possible so anyone would be able to easily understand it at a glance (see Fig. 4)[5]. We followed Nielsen's usability principles [11], emphasizing a user-centered design approach to enhance ease of use, efficiency, and satisfaction for all users. This comprehensive focus

[5] https://github.com/joaquinballesteros/Dependency

on simplicity, clarity, and accessibility aims to create an inclusive application that meets the diverse needs of our user base. Additionally, we have follow the guidelines for to Web content accessibility guidelines (WCAG) 2.0 standard for accessibility to ensure that the app is accessible to individuals with disabilities [2].

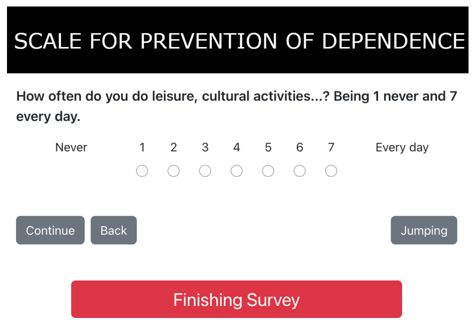

Fig. 4. Web application for evaluate the dependency level

We conducted four usability tests using the think-aloud method to identify issues in the application and improve accessibility. In the first version, two main problems were found. First, users were confused about where to finish their tasks, which we addressed by adding a button clearly indicating the end point. Second, users were uncertain whether they had completed the scale, so we added text to clearly indicate when the scale was finished. These changes aim to enhance the overall user experience and accessibility of the application.

The decision tree that governs what items to show and decides the dependency level of the user taking the survey was created using Python and Scikit-learn. However, this model needed to be translated into a format the application could understand. This format consisted of nodes that kept track of what node to traverse to for each answer, the ID of the item the node refers to, and the current predicted result. Using an abstraction separated from the Python implementation allows us to change the model to a completely separate one if needed.

To achieve this, a Python script was used that takes a file storing the model, traverses the tree, and converts each node of the tree into a node with the proper format. Then it stores all the nodes into the external database so the application can retrieve them. This process is necessary to automate the integration of the decision tree with the selected items. It also enables the system to evolve if the model changes in the future, ensuring flexibility and adaptability.

4 Conclusions and future work

In this work, we use a dependency scale composing by 26 Likert-type items. By employing genetic algorithms, explicitly using a combination of selection, crossover, and mutation, we successfully reduced these 26 items to a more manageable 9-item subset while maintaining high predictive accuracy. Additionally, a web application has been developed to facilitate the use of this estimator. This application includes a decision tree that governs what items to show and determines the user's dependency level.

The evolution play an important role, so a python script automates the conversion and storage of the decision tree nodes into an external database, enabling seamless integration and future adaptability. We conducted four usability tests using the think-aloud method to identify and rectify issues, enhancing both usability and accessibility. These efforts ensure the application is user-friendly and capable of evolving as the model changes over time.

Future work will focus on further refining the dependency estimator by exploring additional machine learning techniques and expanding the dataset to enhance predictive accuracy and generalizability. We aim to conduct extensive field testing with diverse user groups to gather comprehensive feedback, which will inform iterative improvements to the web application. Finally, we will explore the potential for integrating the system with other health and social care platforms to create a more holistic support network for individuals requiring assistance.

References

1. Breiman, L.: Classification and regression trees. Routledge (2017)
2. Caldwell, B., Cooper, M., Reid, L.G., Vanderheiden, G., Chisholm, W., Slatin, J., White, J.: Web content accessibility guidelines (wcag) 2.0. WWW Consortium (W3C) **290**(1-34), 5–12 (2008)
3. Cortina, J.M.: What is coefficient alpha? an examination of theory and applications. Journal of applied psychology **78**(1), 98 (1993)
4. España: Ley 39/2006, de 14 de diciembre, de Promoción de la Autonomía Personal y Atención a las personas en situación de dependencia. Secretaría de Estado de Servicios Sociales, Familias y Discapacidad (2006)
5. Fortin, F.A., De Rainville, F.M., Gardner, M.A., Parizeau, M., Gagné, C.: DEAP: Evolutionary algorithms made easy. Journal of Machine Learning Research **13**, 2171–2175 (jul 2012)
6. Gutiérrez-Murillo, R.S.: Envejecimiento de la población: breves comentarios gerontológicos. Atención Familiar **29**(2), 124–125 (2022)
7. Hæsum, L.K.E., Cichosz, S.L., Hejlesen, O.K.: Using machine learning to design a short test from a full-length test of functional health literacy in adults—the development of a short form of the danish tofhla. Plos one **18**(7), e0280613 (2023)
8. He, H., Garcia, E.: Learning from imbalanced data ieee transactions on knowledge and data engineering (2009)
9. Hernández, Z.: Método de análisis de datos. Apuntes, La Rioja: Universidad de La Rioja, Servicio de Publicaciones (2012)
10. Kramer, O., Kramer, O.: Scikit-learn. Machine learning for evolution strategies pp. 45–53 (2016)
11. Nielsen, J.: Ten usability heuristics (2005)

12. Panovka, P., Salman, Y., Hel-Or, H., Rosenblum, S., Toglia, J., Josman, N., Adamit, T.: Using machine learning to modify and enhance the daily living questionnaire. Digital Health **9**, 20552076231169818 (2023)
13. Schober, P., Boer, C., Schwarte, L.A.: Correlation coefficients: appropriate use and interpretation. Anesthesia & analgesia **126**(5), 1763–1768 (2018)
14. of Sciences, N.A., of Behavioral, D., Sciences, S., Division, M., on Behavioral, B., Sciences, S., on Health Sciences Policy, B., on the Health, C., of Social Isolation, M.D., in Older Adults, L.: Social isolation and loneliness in older adults: Opportunities for the health care system. National Academies Press (2020)
15. Yearbook, E.R.: Luxembourg: Publications office of the european union (2021)
16. Yong, A.G., Pearce, S., et al.: A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutorials in quantitative methods for psychology **9**(2), 79–94 (2013)