Alexandria Engineering Journal (2023) 65, 909–919

HOSTED BY

ELSEVIER

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Effective ML-based quality of life prediction approach for dependent people in guardianship entities

Gaurav Kumar Yadav [a,b], **Benigno Moreno Vidales** [d], **Hatem A Rashwan** [a], **Joan Oliver** [d], **Domenec Puig** [a], **G.C. Nandi** [b], **Mohamed Abdel-Nasser** [c]

[a] *Departament d'Enginyeria Informatica i Matemátiques, Universitat Rovira i Virgili, Tarragona 43007, Spain*
[b] *Department of Information Technology, Indian Institute of Information Technology Allahabad, 211012 Prayagraj, India*
[c] *Department of Electrical Engineering, Aswan University, 81542 Aswan, Egypt*
[d] *Instituto de Robótica para la Dependencia (IRD), Sitges 08870, Barcelona, Spain*

Received 11 June 2022; revised 12 September 2022; accepted 9 October 2022
Available online 22 October 2022

KEYWORDS

Quality of life;
Support intensity scale;
Intellectual disability;
Priority of care;
Machine learning

Abstract This paper proposes an effective approach for predicting quality of life (QoL) for dependent individuals in guardianship entities. In addition, it aims to improve the QoL of people with intellectual disabilities. The proposed QoL prediction approach employs machine learning (ML) techniques to model the relationship between eight aspects of QoL and the corresponding QoL index. It determines whether or not a person needs assistance based on the index value. The proposed approach determines the priority of care (PoC) value for each aspect of a person. Based on PoC, the deficit aspect is determined, followed by the type of assistance a person requires, based on the decision priorities. It also generates a support report with suggested actions to highlight the level in that aspect. In addition, we train multiple ML models to predict the standard score (SS), which represents the support value related to the eight aspects of QoL. Finally, we use SS values to train an ML model to predict the support intensity scale (SIS). On a dataset compiled from guardianship entities, the proposed approach is validated. The QoL index, SS, and SIS prediction models achieve average R^2 values of 0.9897, 0.9998, and 0.9977 with a standard deviation of 0.0051, 0.0002, and 0.0007, respectively.

1. Introduction

Due to the increase in cases, QoL assessment of a dependent individual, particularly one with an intellectual disability (ID), attracts significant research interest. ID is typically defined as intellectual and adaptive functioning deficits determined by standardized testing (e.g., an IQ score of less than 70). It demonstrates a person's inability to avoid performing his socially expected functions, responsibilities, and tasks. The disabilities appear during the developing period and cause daily limits that require continuous support. These deficiencies have an impact not only on autonomous functioning at home

https://doi.org/10.1016/j.aej.2022.10.028
1110-0168

but also on involvement in the community, social, and academic activities [1,2]. According to the articles of the United Nations Convention on the rights of persons with disabilities, people with intellectual disabilities should be able to live as independently as possible, which is something that many people with intellectual disabilities want [3–5]. Improving their ability to manage their affairs independently could improve their QoL and engagement in the community [6]. Individuals with ID can encounter numerous difficulties in their daily lives, but interventions and support can alleviate them [7]. Typically, family members and professionals feel overburdened when supporting an individual with IDs [8–10]. Consequently, interventions encouraging people with IDs to handle their affairs are necessary.

Over the past decades, efforts have been made to reconcile the various aspects that directly influence a person's behavior to achieve the desired outcome [11]. Living life based on one's preference leads to autonomy, which also seeks to provide for those dependent on others [12]. Recently, many support instruments have been developed to aid dependent individuals in living normal life with their families. The recent development in intellectual development disorder (IDD) reveals methods to enhance the QoL using support paradigms. Various countries have recognized a different number of aspects to represent the QoL. In Spain, for instance specialists have adopted eight aspects to evaluate the QoL of people with ID [13]. These eight aspects are emotional well-being (EW), interpersonal relations (IR), material well-being (MW), personal development (PD), physical well-being (PW), self-determination (SD), social inclusion (SI), and rights (RI) [14]. The assessment of the QoL in the eight aspects covers the overall domain. Enhancing QoL means improving these eight areas of an individual with ID. These eight aspects encompass all the necessities of a dependent individual to live life equally as a normal person. The support paradigm helps make a required support plan to improve the QoL aspects for an individual. Integrating the support with QoL generates a new paradigm QoL support model (QoLSM) [15], whereas integrating the support paradigm with QoL indexes enhances the lifestyle of the dependent individuals. For instance, Gómez et al. [15] demonstrate the way to evaluate the model effect on individuals and organizations and improve its performance. Therefore, Verdugo et al. [12] reviewed the recent works in this field based on measurement tools, descriptive correlation studies, predictive studies, and interventions. These four criteria cover the current investigation of intellectual and developmental disabilities. Recent research focuses on various QoL-related factors, such as the use of technology [16,17], prompting [18], employment [19,20], and health behavior [21,22]. Due to the studies' narrow focus and methodological limitations, it is difficult to draw definitive conclusions about the efficacy of self-management interventions. Nevertheless, the majority of previous research has revealed positive results. (See Fig. 1).

Indeed, ML techniques can assist in analyzing the patient's report to predict the patient's potential need for support. This paper proposes an effective approach for predicting QoL for dependent people in guardianship entities. Our approach has numerous stages, from accepting input in the form of eight aspects of QoL to automatically generating a report on necessary assistance without the involvement of psychologists, specialists, physicians, or other professionals. Specifically, we employ ML techniques to build QoL index, SS, and SIS prediction models. We also determine the PoC value for each aspect of a person. This study is based on the Newton-One dataset, a private dataset of 26 subjects. Each subject contains eight aspects of QoL and the QoL index value, with eight to nine questions. A total of eight aspects represent the QoL. The proposed approach can contribute to understanding different aspects of QoL and fulfill the requirement of self-management to reduce reliance on family members and professionals. The main contributions of this paper are listed below:

- Proposing an efficient ML approach for predicting dependent people QoL in guardianship entities.
- Presenting a new ML-based method for determining the PoC for each quality aspect to decide the QoL aspect that needs support.
- Proposing an ML-based method for predicting the SIS value to integrate this value with the other sensory information inputs to strengthen the QoL of dependent people.
- Analyzing the efficacy of different ML techniques to model the QoL to help professionals in decision-making and patients improve their QoL.

The structure of the paper is as follows: The literature review on this topic of intellectual disability is presented in Section 2. Section 3 illustrates the proposed approach and dataset details. The results are presented in Section 4. Finally, Section 5 provides the conclusion and the future work.

2. Related works

2.1. QoL's aspects and support paradigm

QoL refers to an individual's general well-being while meeting fundamental needs. Independently enjoying a high quality of life is difficult for individuals with an ID. People with ID who have IQs below 70 have difficulty performing personal and social, and behavioral activities effectively [2]. Therefore, improving the QoL of these individuals is a fundamental challenge for the researchers. QoL encompasses three aspects of life— the personal, the social, and the judicial. Improving

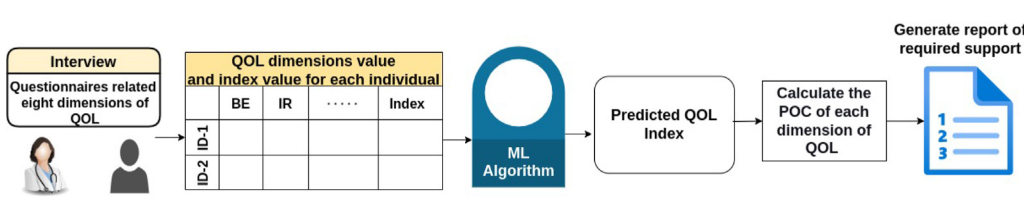

Fig. 1 Schematic diagram of the proposed method for predicting dependent people QoL in guardianship entities.

QoL entails focusing on these three aspects of an individual's identity. These three areas are further subdivided into eight aspects of QoL [23,24], as shown in Table 1. These eight aspects satisfy all necessary substitutes for a person's QoL.

The type and intensity of support necessary to complete a specified task determine the support need [14]. The support paradigm's major goal is comprehending a person based on their basic assistance needs. Using this paradigm of support, organizations assist families of individuals with intellectual disabilities with inclusive education, independent living, supported employment, and other rights using this paradigm of support [25,26]. The capability of the individual with ID to operate in their environment will be improved so that the assistance provided meets their needs, goals, and preferences. An individual's support plan is used to calculate support needs, which are then aggregated as data. Utilization of combined data to improve organizational efficiency and research allocation [26,27]. The purpose of support needs assessment is to develop an individualized and generalized support system to help people with ID, and their families improve their QoL over time. Therefore, the best practice in intellectual impairment is to assess support needs. There are numerous ways to assist with these support requirements. In the SIS column of Table 1, each QoL aspect is associated with support requirements. The area indicates the required area where support is required, corresponding to each aspect. These support areas are health and healthcare, protection and defense, Behavioral support need, social activities, employment activities, home life activities, lifelong learning, exceptional medical needs, and community life activities.

2.2. GENCAT scale

The GENCAT scale is a tool created by INICO. It provides an impartial assessment of the QoL of an individual [28,29]. In this assessment, the user must reply to the 69 objective questions using a frequency scale based on their observation. Fig. 2 presents a fragment of questionnaires used to collect the dataset. The GENCAT scale has been developed and validated using Schalock and Verdugo's multidimensional model. The scale generates valid and trustworthy scores for each of the eight aspects of QoL and a global QoL index based on the multidimensional model.

3. Methodology

The relevance of this domain in society is demonstrated in the literature. Learning algorithms aid the psychologist in analyzing the patient's situation during professional interviews. A set of questionnaires containing 69 questions covering all possible aspects of an intellectually disabled person's life is administered by professionals. Professionals record the answers to the questions based on four points frequency scale. Professionals use a tool made by INICO known as the GENCAT scale [28], which transforms 69 questions into eight aspect score values and corresponding index values. We interviewed 26 subjects and prepared our dataset using the GENCAT scale tool. Hence, the initial dataset contains 26 rows and 9 columns.

In this study, we train three ML-based models: the QoL index, the SS, and the SIS. The flowchart for constructing the three models is depicted in Fig. 3. First, we divide the dataset threefold. Each fold has the train and test sets. After that, the SMOTE-R algorithm is used to augment the training dataset of each fold. We train the index model using the augmented dataset. The index model predicts the QoL index value, which decides whether either beneficiary needs support or not. If the QoL index value is more significant than 100, the beneficiary does not need support, and the process stops. The process moves forward if the index value is less than or equal to 100.

Furthermore, we calculate the PoC value for each aspect and beneficiary and incorporate it into the augmented dataset. We also determine the SS value by summing the PoC value of each aspect of the beneficiary's QoL (detailed in Section 3.4). We train the SS model using this modified dataset. The SS model primarily calculates the PoC score value for each aspect and predicts the SS value of support for the beneficiary. Finally, we train the SIS model using another dataset that con-

Table 1 Details of aspects of QoL and support intensity scale metric with the maximum value of support, and also with several activities.

QoL aspects	Support intensity scale metric			
	AreaofSIS	*Numberofsupportactivities*	*Maximumvalueofsupport*	*Areaofsupport*
Emotional well-being (EW)	*Healthandhealthcare*(*S*1*E*)	8	94	
	Protectionanddefense(*S*2)	8	94	
	Behavioralsupportneed(*S*3*B*)	13	26	
Personal development (PD)	*Homelifeactivities*(*S*1*A*)	8	92	
	Lifelonglearning(*S*1*C*)	9	104	
				Personalarea
Physical well-being (PW)	*Healthandhealthcare*(*S*1*E*)	8	94	
	Exceptionalmedicalneed(*S*3*A*)	16	32	
Self-determination (SD)	*Protectionanddefense*(*S*2)	8	94	
Interpersonal relation (IR)	*Socialactivities*(*S*1*F*)	8	93	
Social inclusion (SI)	*Communitylifeactivities*(*S*1*B*)	8	91	*Socialarea*
	Socialactivities(*S*1*F*)	8	93	
Material well-being (MW)	*Employmentactivities*(*S*1*D*)	8	87	
Rights (RI)	*Protectionanddefense* (S2)	8	94	*Judicialarea*
	Healthandhealthcare(*S*1*E*)	8	94	

Part –B Life activities in the community	frequency					Daily support time					Type of support					Direct Score
1. Moving from one place to another throughout the community (Transportation)	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
2. Participate in recreational and leisure activities in community settings	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
3. Access buildings and public environments	0	1	2	3		0	1	2	3	4	0	1	2	3	4	7
4. Go to visit friends and family	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
5. Participate in preferred community activities (parish, volunteering..)	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
6. Go shopping and buy groceries and services.	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	8
7. Interact with members of community	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	6
8. Access public buildings and environments	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	7

Score direct total Life activities in the community	64

Fig. 2 A fragment of questionnaires used to collect the dataset. Each question has three parameters(frequency, daily support time, and type of support). The beneficiary must select the appropriate response from the scale provided. The brown color number indicates the responses that have been selected, while the black color alternatives indicate that there is no option to choose this number. The direct score is the sum of the question's three parameters, and the sum of all direct scores for all questions is the direct score total.

tains the SS value and corresponding SIS value. The SIS model predicts the SIS value and finally takes the SS value of support. The process ends once the SIS value has been predicted.

The flowchart of the process for generating the required support report is demonstrated in Fig. 4. After calculating the PoC value in the SS model, a support report is generated based on the PoC value. If the PoC value is above 25, the beneficiary is suffering in this aspect of life, and needs immediate support to improve it. After that, if the PoC value lies between 15 to 25, the beneficiary does not need direct support but requires support to perform better in this aspect of life. Finally, if the PoC value is between 10 to 15, the beneficiary only needs optional support as a precaution. In contrast, the beneficiary does not require any support in this aspect.

3.1. Dataset preparation

We own the newton-One dataset, as we collected the information by interviewing various individuals. Based on questionnaires interviewer asked questions to each individual. The dataset includes eight QoL aspects, each with eight to nine questions. Professionals pose these questions to individuals and collect the answers based on a four-point frequency scale [30]. Following the interview, professionals use the GENCAT scale to transform the answers to each aspect question into the aspect's QoL value. There are eight aspect values in the dataset. These eight aspect values are used to generate the QoL index value based on the GENCAT Scale method [31]. The beneficiary's QoL index value indicates whether or not the beneficiary needs assistance.

The value of the QoL index varies according to the Gaussian distribution. The maximum populations are found on the Gaussian curve. The minimum value starts at 68, and the maximum reaches 130. The average of this distribution is 100. Therefore, if a person's QoL index value is above 100, she/he does not need any support. Whereas if the person's QoL index value is equal to or less than 100. Consequently, this person needs support in relation to each aspect. We have a dataset containing the information of 26 individuals affected with ID. The original dataset includes 26 subjects, eight QoL aspects, and a QoL index value. The mean and standard deviation of our dataset for each aspect of QoL and index value are depicted in Table 2. It also contains the minimum and maximum values of the dataset.

3.2. ML techniques

We employ various ML techniques in this study, including regression tree (RT) [32], random forest (RF) [33], gradient boosting (GB) [34], multiple linear regression (MLR) [35], multilayer perceptron regressor (MLPRegressor) [40], and adaptive neuro-fuzzy inference system (ANFIS) [41]. These

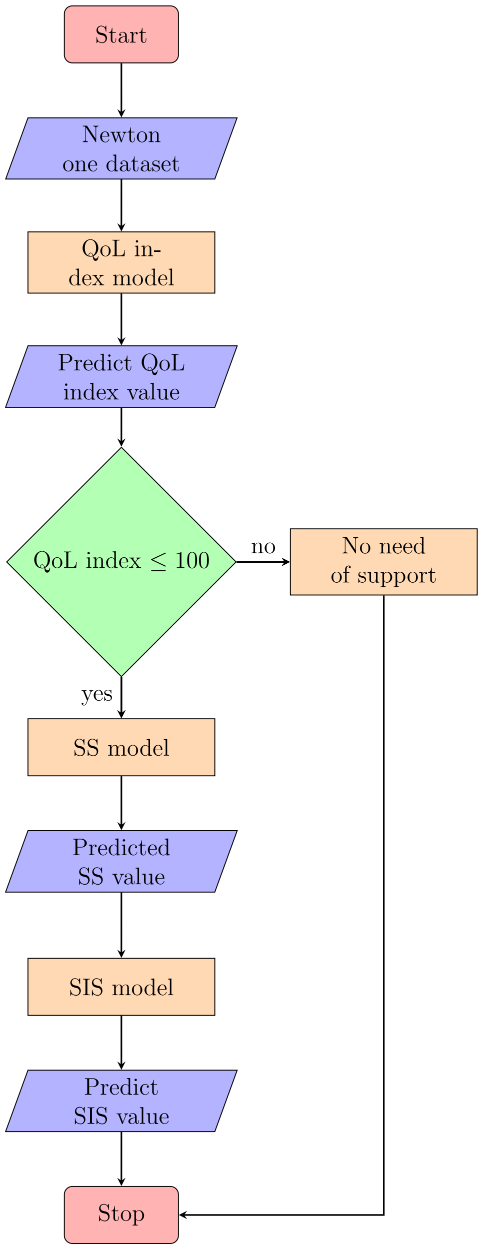

Fig. 3 Flowchart of constructing the QoL index, SS, and SIS models.

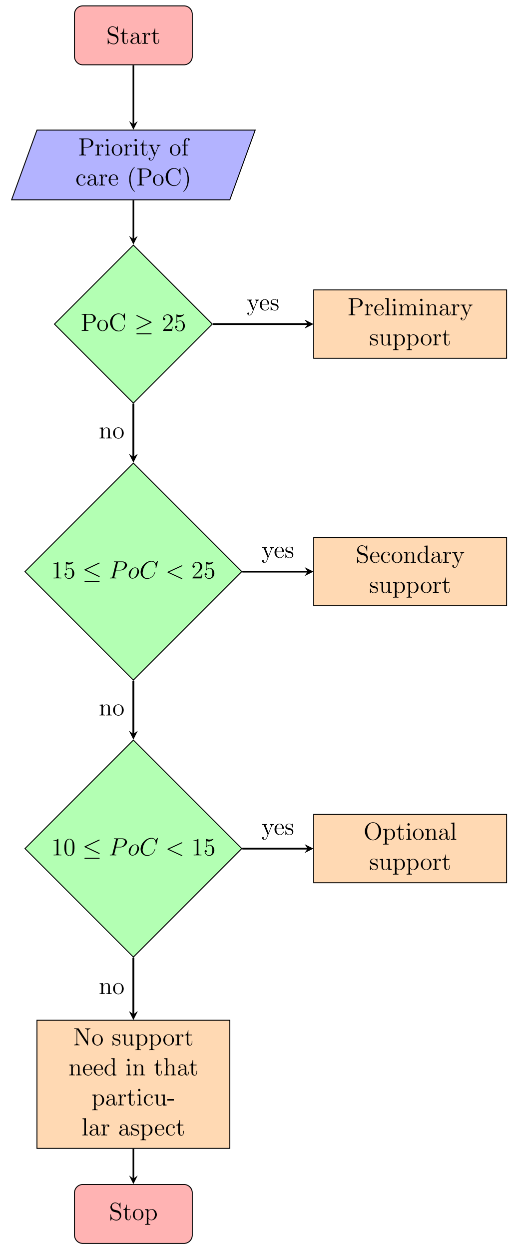

Fig. 4 Flowchart of generating required support reports.

algorithms are popular for dealing with linear as well as non-linear datasets. MLR determines the relationship between one dependent and two or more two independent variables. MLR follows some assumptions such as independence of observations, there should be no hidden relation among variables, and data follow normality and linearity.

$$Y^i = W^i_0 + W^i_1.X^i_1 + W^i_2.X^i_2 + \ldots.. + W^i_n.X^i_n, \quad (1)$$

where W_0 is the bias term, and W_1 to W_n is the weight coefficient. X_1 to X_n are the input features, and Y is the output. i represent the number of the dataset. Here, we primarily have eight aspects of QoL as input to the model, in which X_1 to X_8 and QoL index is the output, representing Y.

The random forest regression algorithm, is a supervised learning approach for regression that uses the ensemble learning method. The ensemble learning method combines predictions from several ML algorithms to produce a more accurate forecast than a single model [36]. MLPRegressor [40] is a neural network that contains an input layer, hidden layers, and an output layer. The MLPRegressor algorithm is adjustable and can commonly be used to learn a mapping from inputs to outputs, meaning it can solve complex linear and nonlinear regression problems. The ANFIS integrates the benefits of artificial neural networks and fuzzy inference systems [41]. ANFIS is a reliable technique for constructing complex and nonlinear relationships between input and output data sets.

3.3. Detail of experiments

The algorithm must first examine the QoL index value after receiving the dataset. The QoL index ranges from 68 to 130.

Table 2 Mean, standard deviation, minimum, and maximum values of each aspect of the Newton-One dataset.

	EW	IR	MW	PD	PW	SD	SI	RI	QoL Index
Mean	101.192308	102.076923	92.038462	104.461538	81.346154	114.730769	97.961538	94.384615	97.461538
std	14.759456	15.478819	17.340083	15.862486	15.184050	8.951235	20.588309	17.024869	16.322330
min	81.000000	81.000000	68.000000	68.000000	68.000000	94.000000	68.000000	68.000000	72.000000
max	126.000000	130.000000	115.000000	130.000000	119.000000	130.000000	126.000000	115.000000	125.000000

If the index value is greater than 100, indicating the typical distribution, the beneficiary does not need assistance. Only data with a QoL index larger than 100 is eliminated. Once we have obtained data with QoL index values less than or equal to 100, the dataset is divided into a train and validation fold. In this study, the dataset is small enough to train an ML system. As a result, we use the synthetic minority oversampling strategy for regression to augment the train data. The number of train data points used to train our algorithm is increased. After augmentation, we acquire a large number of distinct train datasets. Nothing is added to the validation data set to validate our trained model on the original dataset. The results reveal that the trained model is accurate. After receiving the augmented train dataset and original validation dataset, the algorithm calculates the SS value of support in the case of the SS model for both augmented train and validation data. We divided the original data by three and performed the aforementioned steps for each fold. We train our data using MLR, RT, RF, GB, MLPRegressor, and ANFIS models for all threefold. For every three models, we train separately based on the performance in each case, and we select the model accordingly. In the case of the MLR model, we evaluated the trained model and obtained an excellent R^2 score for the training testing case.

Algorithm 1. Priority of care for each QoL aspect algorithm

Require: Data, Dict = {}, List = []
Ensure: (Dict contains the maximum value of each support intensity scale, and the list contains no values initially)

function LOOP(Data)
    for $j \leftarrow 0$to $D-1$ **do** ▷D = Number of aspects of QoL
        if $j = 0$ **then** Calculate PoC_{EW} score value by averaging the three support metric values.
        else if $j = 1$ **then**
Calculate PoC_{IR} score value by averaging the corresponding support metric values.
        else if $j = 2$ **then**
Calculate PoC_{MW} score value by averaging the corresponding support metric values
        else if $j = 3$ **then**
Calculate PoC_{PD} score value by averaging the corresponding support metric values.
        else if $j = 4$ **then**
Calculate PoC_{PW}score value by averaging the corresponding support metric values.
        else if $j = 5$ **then**
Calculate PoC_{SD} score value by averaging the corresponding support metric values.
        else if $j = 6$ **then**
Calculate PoC_{SI} score value by averaging the corresponding support metric values.
        else if $j = 7$ **then**
Calculate PoC_{RI} score value by averaging the corresponding support metric values
        end if
    end for
    $SS \leftarrow Sum(PoC_{EW}, PoC_{IR}, PoC_{MW}, PoC_{PD}, PoC_{PW}, PoC_{SD}, PoC_{SI}, PoC_{RI})$
    $List \leftarrow SS$
end function

3.4. Standard score of support corresponding QoL

We calculate the PoC value corresponding to each aspect of QoL for every beneficiary. Initially, we calculate PoC for an area of action separately, and each aspect has corresponding support group actions as depicted in Table 1 in the second column. Table 1 illustrates that, the first aspect of QoL, EW has three support areas known as S1E, S2, and S3B. The PoC value for every three areas is calculated separately using (2).

$$PoC = \frac{A - E}{A} \times I, \tag{2}$$

where A is the average QoL index value, E is the aspect value, and I is the maximum support need value. The values of I are listed in Table 1, A value is 100, and the E value is given by the beneficiary.

We calculate the final PoC_{EW} value for the EW aspect by averaging the PoC values for each of the three areas for the EW aspect. Likewise, we calculate the PoC value of other aspects. After obtaining all the PoC values for each aspect, we compute the SS by summing them as shown in Eq. (3).

$$SS = PoC_{EW} + PoC_{PD} + PoC_{PW} + PoC_{SD} + PoC_{IR} + PoC_{SI} + PoC_{MW} + PoC_{RI}, \tag{3}$$

Algorithm 1 presents the steps to calculate the PoC value for each aspect by following Eqs. (2) and (3). Once the PoC value has been calculated, they are summed to obtain the SS value. SS value represents the support score value corresponding to eight aspects of QoL. After that, we add the new column to the original data for the training of the SS model. We calculated the SS value of each augmented data threefold. The training of the SS model takes eight aspects as input and SS value as a target. Once training was completed, we validated our trained model using the original validation dataset.

3.5. Data augmentation using SMOTE-R

The synthetic minority over-sampling technique for regression SMOTE-R [37] is an oversampling technique to increase the imbalanced dataset. The SMOTE concept is proposed to reduce the imbalanced distribution of a dataset for classification tasks [38]. It utilizes the Gaussian noise concept to maximize the amount of data. In regression, entering the data and the corresponding column helps determine the minimum value of the column. It is a sampling method to address with imbalanced class distribution by undersampling the regular classes and oversampling the rare category. It generates synthetic data corresponding to minority target values. Algorithm 2 presents the detailed working process of SMOTE-R [37].

Algorithm 2. SMOTE for regression algorithm

function SMOTE-R(D, thr, o, u, k) ▷D - dataset,
thr- The value of the target variable's relevance threshold,
o and u represent the percentage of over and under-sampling,
k -In case of generation, the number of neighbors used.
$rearL \leftarrow \{(x, y) \in D : \phi(y) > thr \bigwedge y < \hat{y}\}$ ▷$\hat{y}$ is the median of the target Y
$newCasesL \leftarrow GenSynthCases(rareL, \%o, k)$ ▷generate synthetic cases for rareL
$rareH \leftarrow \{(x, y) \in D : \phi(y) > thr \bigwedge y > \hat{y}\}$
$newCasesH \leftarrow GenSynthCases(rearH, \%o, k)$ ▷generate synthetic cases for rareH
$newCases \leftarrow newCasesL \cup newCasesH$
$nrNorm \leftarrow \%uof|newCases|$
$normCases \leftarrow sampleofnrNormcases \in D\{rareL \cup rareH\}$ ▷under-sampling **return** $newCases \cup normCases$
end function

3.6. Developing the prediction models

The three prediction models are the index, SS, and SIS models. The index model is the first to determine if a person requires assistance, and it predicts the QoL index value using eight aspects of the QoL value as input. The SS model is trained to predict the SS value. It estimates the SS value based on eight aspects of QoL. As a result, it is the most common model. The SIS model uses the output of the SS model to predict the SIS value. The SS model's output is used to predict the SIS value by the SIS model. These three models accomplish their tasks sequentially. A multivariate linear regression approach is used to train all three models. The second branch is followed to provide a beneficiary with a support report corresponding to the requirement in a particular aspect. The second branch uses the index and SS models, which estimate the PoC for each beneficiary's eight aspects. The approach to computing the PoC for each aspect is illustrated in Algorithm 1. The value of PoC determines the level of support.

As demonstrated in Fig. 4, it contains three conditions. If an aspect's PoC value is greater than or equal to twenty-five, the beneficiary needs an immediate action plan in relation to that QoL aspect. Furthermore, if the PoC of an aspect is more significant than fifteen but less than twenty-five, a person requires secondary-level assistance. When a person's PoC of an aspect is more significant than ten or equal to ten but less than fifteen, the recipient will only require partial assistance, which is not mandatory. Furthermore, if a person's PoC in any aspect is less than ten, no assistance is required in that aspect. Our model generates all three sorts of support for each aspect of each person. Beneficiaries can use these tips without the assistance of a professional, and they will improve their lives.

3.7. Validation and evaluation metrics

In this study, the k-fold cross-validation technique [35] is used to cross-validate the model performance of the random train and validation case data in the K fold. It shows the performance of the model on the new dataset. It is generally used in the case of a limited data sample. This procedure has a single parameter, K, which depicts how many folds we need to split the original dataset. The data are randomly divided into the fold. In this work, we split our dataset threefold. We get to train and validate data in each fold differently. Now we use the SMOTE-R algorithm to increase the train data in each fold and keep validation data to verify the model accuracy.

In this study, we use three evaluation metrics to evaluate our performance of the trained model, namely, mean absolute error (MAE), root mean square error (RMSE), and R^2. MAE calculate the absolute error between the predicted and actual values. MAE can be formulated as follows:

$$MAE = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n}, \tag{4}$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value, and n is the total number of data points.

RMSE measures how far the data points are from the regression line; RMSE measures how to spread out the residuals. RMSE can be expressed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{n}}, \tag{5}$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value, and n is the number of data points.

Finally, we calculate the R^2 score value, also known as the coefficient of determination. It is a statistic that indicates how well a model fits the data. R^2 can be defined as follows:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}, \quad (6)$$

where $\sum (y_i - \hat{y}_i)^2$ is the sum of squares of residuals, and $\sum (y_i - \bar{y}_i)^2$ is the total sum of the square.

4. Results and discussion

Tables 3–5 display the results of the index, SS, and SIS models. Table 3 presents the results for the first case, where input is the eight aspects of QoL, and the target is the QoL index. Results are shown here in the form of various metrics. Six regression models we trained for three different folds. Table 3 contains the best results threefold for all six algorithms. MAE, RMSE, and R^2 scores are the fundamental metrics representing the regression performance. It should be noted that we use a three-fold cross-validation technique in our experiments. Each fold of the cross-validation produces MAE, RMSE, and R^2 values. In all Tables, we report the mean and standard deviation (SD) of the evaluation metrics values of the three folds. R^2 score demonstrates the closeness of the predicted value with the actual value, and it varies between zero and one. One demonstrated the overlap between the predicted value with the regression line. Therefore, a higher value of R^2 shows the model's goodness. Both MAE and RMSE measure the accuracy of the predictions of the ML regression models and demonstrate the amount of deviation from the actual values. Table 3 presents the index model's mean and SD of MAE, RMSE and R^2 values. MLPRegressor and ANFIS produce MAE of 4.1344 and 2.6936, respectively. RT, RF, and GB produce MAE and RMSE values higher than 5. MLR achieves the smallest MAE and RMSE values (1.2247 and 1.5236) and the highest R^2 score value than other regression techniques results for the index model.

Table 4 presents the results for the SS model. Here the input parameters are the eight aspects of QoL, and the corresponding target is the PoC value. RT, RF, GB, MLPRegressor, and ANFIS produce an average MAE higher than 0.6. MLR achieves the highest R^2 score value (0.9977) and the lowest MAE value (0.7702). Therefore, we use MLR trained model for future integration with a SS model to predict the SS value.

Table 5 presents the results of the SIS model. It takes input as SS value and has a corresponding SIS value, representing the final support metric for a person's fundamental aspects of QoL. MLR achieves MAE, RMSE, and R^2 of 0.7702, 1.3403, and 0.9977, respectively, which are much better than the results of RF, GB, MLPRegressor, and ANFIS. However, RT produces RMSE values lower than MLR, and the MAE and R^2 values of MLR are better than RT. Consequently, we selected MLR to build the SIS model.

Our entire model has two branches. First, it contains a trained MLR model that considers eight aspects of QoL and predicts the QoL index value. The algorithm determines whether or not the patient needs support based on the QoL index value. The algorithm is trained to pass only those patient information to further for an analysis whose QoL index value is less or equal to 100. After detecting the QoL index value, it then goes through the second trained model after detecting the QoL index value. During testing, the SS model receives input from the beneficiary and predicts the SS value. The third trained SIS model receives the projected value as input and predicts the SIS value after the SS model predicts the SS value,

Table 3 The MAE, RMSE, and R^2 values of the six ML techniques used for building the index model.

Algorithms	MLR	RT	RF	GB	MLPRegressor	ANFIS
metrics	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD
MAE	1.2247 ± 0.0904	9.2030 ± 3.0083	5.5271 ± 2.1788	5.9723 ± 1.2817	4.1344 ± .4683	2.6936 ± 0.1325
RMSE	1.5236 ± 0.1644	11.6664 ± 4.2920	6.5353 ± 2.4031	7.2981 ± 1.6082	5.1202 ± 0.3524	2.8361 ± 0.1722
R^2 Score	0.9897 ± 0.0051	0.4213 ± 0.3698	0.8227 ± 0.0935	0.7650 ± 0.1257	0.8698 ± 0.0095	0.9259 ± 0.0619

Table 4 The MAE, RMSE, and R^2 values of the six ML techniques used for building the SS model.

Algorithms	MLR	RT	RF	GB	MLPRegressor	ANFIS
metrics	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD
MAE	0.2260 ± 0.2655	39.9563 ± 2.3875	32.8962 ± 2.4674	34.7427 ± 1.6301	7.8524 ± 0.3678	0.7362 ± 0.0785
RMSE	0.2765 ± 0.3173	47.3084 ± 3.9119	39.2744 ± 1.4503	40.0628 ± 0.6041	9.5437 ± 0.2537	0.9821 ± 0.1080
R^2 Score	0.9998 ± 0.0002	−0.7837 ± 0.4098	−0.2575 ± 0.4463	−0.3001 ± 0.4129	0.9341 ± 0.7849	0.9632 ± 0.0445

Table 5 The MAE, RMSE, and R^2 values of the six ML techniques used for building the SIS model.

Algorithms	MLR	RT	RF	GB	MLPRegressor	ANFIS
metrics	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD	Avg ± STD
MAE	0.7702 ± 0.1025	0.9091 ± 0.1285	0.6657 ± 0.03816	0.9364 ± 0.1030	1.3324 ± 0.0968	1.7536 ± 0.5361
RMSE	1.3403 ± 0.4674	1.1692 ± 0.0253	1.02734 ± 0.4342	1.1821 ± 0.0381	1.6998 ± 0.1068	1.9642 ± 0.7385
R^2 Score	0.9977 ± 0.0007	0.9976 ± 0.0002	0.9981 ± 0.0014	0.9976 ± 0.0005	0.9950 ± 0.0008	0.9825 ± 0.0125

which shows the support value corresponding to a beneficiary's eight aspects of QoL. In the second branch, after predicting the QoL index value by the index model, the model decides whether the person needs support or not. Once the decision is made model calculates the PoC value and, based on that, generates the support report as shown in flowchart 2 in Fig. 4. Assessing the accuracy of a regression model is a tedious task. Here, we use the λ accuracy measure and threshold accuracy measure [39]. As presented in Algorithm 3, λ is an error threshold value between the predicted and actual values. If the error calculated between the actual and predicted value is less than or equal to the λ value, we consider the prediction accurate; otherwise, inaccurate. In the current research, the accuracy is directly dependent on the λ value. If the value of λ is high, accuracy is directly increased. If the λ value is small, the error acceptance is compassionate, and the accuracy may be less. We set the λ value to 0.75, yielding accuracy values of 100, 83.33, and 66.67%, with the index, SS and SIS models, respectively.

Algorithm 3. λ accuracy measure

```
λ = 0.75
if Error ⩽ λ then
True ← True + 1
else
False ← False + 1
end if
Accuracy = True / (True+False) * 100
```

We also use the threshold accuracy measure [39] to evaluate the accuracy of the ML regression models to estimate errors under different thresholds, indicating how often our estimate is correct. The threshold accuracy measure is essentially the expectation E that the predicted value error of input is lower than a threshold thr^z [39]:

$$\delta_z = E[F(max(\frac{y}{\hat{y}}, \frac{\hat{y}}{y}) < thr^z)], \tag{7}$$

where $F(.)$ is an indicator function that outputs a binary value (0 or 1), y and $\hat{y}$ are the actual and predicted values, respectively, thr is a threshold value set as 1.01. z is the parameter that controls the acceptance span. Here, the values of z range from 1 to 5 that correspond to $thr^z = \{1.01, 1.0201, 1.030301, 1.04060401, 1.0510100501\}$. As the value of z increases, the span of acceptance increases.

Fig. 5 presents the accuracy of the index, SS and SIS models with five threshold values (five z values). In the case of $z = 1$, the threshold $thr^1 = 1.01$, resulting in accuracy values of 50, 33, and 72% with the index, SS and SIS models, respectively. The accuracy of the SIS model is identical for the threshold values 1.0201, 1.030301, and 1.04060401. When we set a high acceptance span z to 5 (thr^5 = 1.05101005), it achieves the highest accuracy values of the index, SS and SIS models. For the index and SIS models when $z = 5$, the term $max(\frac{y}{\hat{y}}, \frac{\hat{y}}{y})$ of all test samples is lower than 1.05101005, and thus the accuracy is 100 %. While in the case of the SS model, the term $max(\frac{y}{\hat{y}}, \frac{\hat{y}}{y})$ of all test samples is not lower than 1.05101005.

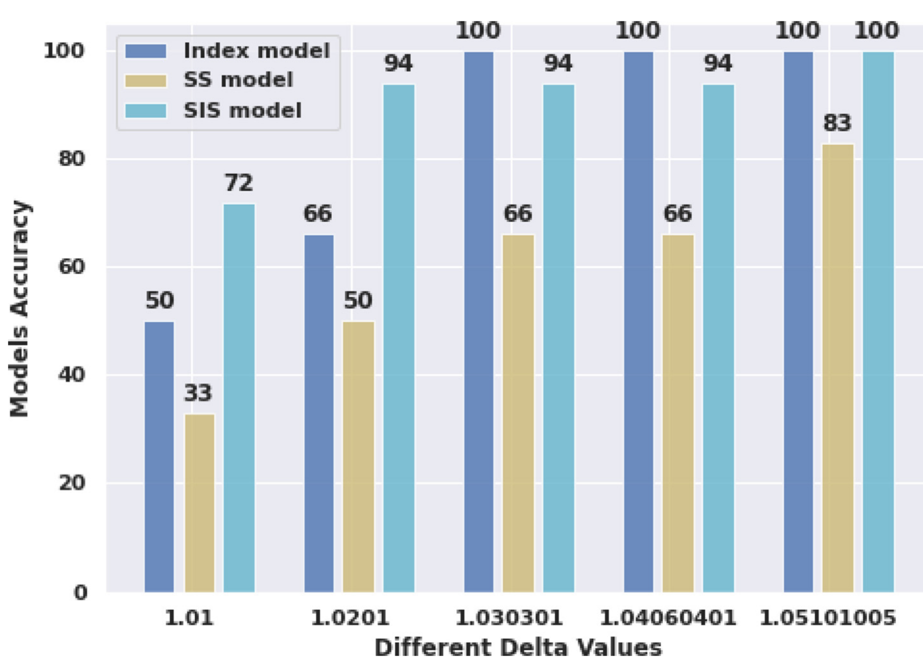

Fig. 5 The threshold accuracy measure of the index, SS and SIS models with five threshold values (five z values). The blue bars represent the index model, the yellow bars represent the SS model, and the green bars represent the SIS model.

5. Conclusion and future work

This paper proposed an efficient ML-based approach for predicting the QoL of dependent people with intellectual and developmental disabilities in guardianship entities by analyzing the various aspects of QoL. In addition, we proposed three ML models for predicting QoL index value, SS value, and SIS value, along with implementing and testing different ML techniques: MLR, RT, RF, GB, MLPRegressor, and ANFIS. Based on the QoL index, the proposed approach determines the PoC for each aspect of QoL. We validated our approach on a dataset collected from guardianship entities. We found that MLR yields the best prediction results for the QoL index, SS, and SIS. The QoL index, SS, and SIS ML models achieved MAE values of 1.2247, 0.2260, and 0.7702, respectively, and also obtained average R^2 scores of 0.9897, 0.9998, and 0.9977, respectively. The proposed ML approach can assist professionals in analyzing the QoL of a beneficiary to determine which measures are required to improve their QoL.

Future research will focus on using sensory data (e.g., data of sensors that monitor the activity of dependent individuals) collected from guardianship entities to improve the effectiveness of the proposed approach.

Data availability statement

No associated data is available.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] V.B. Pender, R.W. Brendel, M.J. Fitz-Gerald, A.P. Nelson, E.G. Aoun, J.L. Boyer, K. Certa, C. Freeman, M.H. Roessel, G.A. Martin, E.M. Plakun, M. Reid, F. Torres, S. Virani, C.D. Wills, M. Young, U. Yazdani, Diagnostic and statistical manual of mental disorders, Am. Psychiat. Assoc. 21 (2013) 591–643, https://doi.org/10.1176/appi.books.9780890425787.

[2] J. Sandjojo, A.M. Zedlitz, W.A. Gebhardt, J. Hoekman, J.A. den Haan, A.W. Evers, Effects of a self-management training for people with intellectual disabilities, J. Appl. Res. Intell. Disabil. 32 (2) (2019) 390–400.

[3] A. Haigh, D. Lee, C. Shaw, M. Hawthorne, S. Chamberlain, D. W. Newman, N. Beail, What things make people with a learning disability happy and satisfied with their lives: An inclusive research project, J. Appl. Res. Intell. Disabil. 26 (1) (2013) 26–33.

[4] N.M.J. Kuijken, J. Naaldenberg, M.W. Nijhuis-Van der Sanden, H.M.J. Van, Schrojenstein-Lantman de Valk, Healthy living according to adults with intellectual disabilities: towards tailoring health promotion initiatives, J. Intell. Disabil. Res. 60 (3) (2016) 228–241.

[5] R.J. Bond, J. Hurst, How adults with learning disabilities view living independently, Brit. J. Learn. Disabil. 38 (4) (2010) 286–292.

[6] C.A. Dollar, L.D. Fredrick, P.A. Alberto, J.K. Luke, Using simultaneous prompting to teach independent living and leisure skills to adults with severe intellectual disabilities, Res. Develop. Disabil. 33 (1) (2012) 189–195.

[7] R.L. Schalock, R. Luckasson, M.J. Tassé, M.A. Verdugo, A holistic theoretical approach to intellectual disability: Going beyond the four current perspectives, Intell. Develop. Disabil. 56 (2) (2018) 79–89.

[8] F. Dawson, S. Shanahan, E. Fitzsimons, G. O'Malley, N. Mac Giollabhui, J. Bramham, The impact of caring for an adult with intellectual disability and psychiatric comorbidity on carer stress and psychological distress, J. Intell. Disabil. Res. 60 (6) (2016) 553–563.

[9] M.A. Hermsen, P.J.C.M. Embregts, A.H.C. Hendriks, N. Frielink, The human degree of care. Professional loving care for people with a mild intellectual disability: an explorative study, J. Intell. Disabil. Res. 58 (3) (2014) 221–232.

[10] R. Vilaseca, M. Gràcia, F.S. Beltran, M. Dalmau, E. Alomar, A. L. Adam-Alcocer, D. Simó-Pinatella, Needs and supports of people with intellectual disability and their families in Catalonia, J. Appl. Res. Intell. Disabil. 30 (1) (2017) 33–46.

[11] L.E. Gómez Sánchez, M.L. Morán Suárez, S. Al-Halabí Díaz, C. Swerts, M.A. Verdugo Alonso, R.L. Schalock, Quality of life and the International Convention on the Rights of Persons with Disabilities: consensus indicators for assessment, Psicothema. 34 (2) (2022) 182–191.

[12] M.A. Verdugo, V. Aguayo, V.B. Arias, L. García-Domínguez, A systematic review of the assessment of support needs in people with intellectual and developmental disabilities, Int. J. Environ. Res. Public Health. 17 (24) (2020) 9494.

[13] M.A. Verdugo, B. Arias, L.E. Gómez, R.L. Schalock, Development of an objective instrument to assess quality of life in social services: Reliability and validity in Spain, Int. J. Clin. Health Psychol. 10 (1) (2010) 105–123.

[14] L.E. Gómez Sánchez, R.L. Schalock, M.A. Verdugo Alonso, A new paradigm in the field of intellectual and developmental disabilities: Characteristics and evaluation, Psicothema. 33 (1) (2021) 28–35.

[15] L.E. Gómez, R.L. Schalock, M.A. Verdugo, A quality of life supports model: Six research-focused steps to evaluate the model and enhance research practices in the field of IDD, Res. Develop. Disabil. 119 (2021) 104112.

[16] A. Geppert, E.M. Smith, D. Haslett, J. Wong, I.D. Ebuenyi, J. Noske-Turner, M. MacLachlan, Assistive Technology to Promote Participation in Sport for People with Disabilities, pre-print, Res. Square (2022), https://doi.org/10.21203/rs.3.rs-1250239/v1.

[17] L. Desideri, G. Lancioni, M. Malavasi, A. Gherardini, L. Cesario, Step-instruction technology to help people with intellectual and other disabilities perform multistep tasks: a literature review, J. Develop. Phys. Disabil. 33 (6) (2021) 857–886.

[18] S. Walters, J. Stierle, D. Stockwell, A. Carlson, J. Ryan, Efficacy of video prompting using mobile technology to teach employment tasks to individuals with intellectual disability, J. Inclusive Postsecond. Educ. 3 (1) (2021).

[19] V. Garrels, H.M.H. Sigstad, Employment for persons with intellectual disability in the Nordic countries: A scoping review, J. Appl. Res. Intell. Disabil. 34 (4) (2021) 993–1007.

[20] H. Gjertsen, S. Hardonk, J. Ineland, Work Inclusion for People with Intellectual Disabilities in Three Nordic Countries: The Current Policy and Challenges, Scandinavian J. Disabil. Res. 23 (1) (2021) 360–370.

[21] L.E.G. Sanchez, The Environmental, Public Health, and Human Rights Impacts on Enhancing the Quality of Life of People with Intellectual Disability. International Journal of Environmental Research and Public Health, MDPI, 2021.

[22] L. Morán, L.E. Gómez, G. Balboni, A. Monsalve, M.A. Verdugo, M. Rodríguez, Predictors of individual quality of life in young people with Down syndrome, Rehabil. Psychol. 67 (2) (2022) 205–214.

[23] H. González Fernández, C. Laborda Molla, M. Jariot García, Employability skills and quality of life among employees with mild and moderate intellectual disability who attend occupation centres in Catalonia (Spain). 52(1) (2021) 45-57.

[24] M.A. Verdugo, R.L. Schalock, K.D. Keith, R.J. Stancliffe, Quality of life and its measurement: Important principles and guidelines, J. Intell. Disabil. Res. 49 (10) (2005) 707–717.

[25] J.R. Thompson, K.A. Shogren, M.L. Wehmeyer, Supports and support needs in strengths-based models of intellectual disability, in: Handbook of research-based practices for educating students with intellectual disability, Routledge, pp. 39-57, 2016.

[26] R.L. Schalock, R. Luckasson, M.J. Tassé, An Overview of Intellectual Disability: Definition, Diagnosis, Classification, and Systems of Supports, Am. J. Intell. Develop. Disabil. 126 (6) (2021) 439–442.

[27] M. Lombardi, L. Croce, C. Claes, S. Vandevelde, R.L. Schalock, Factors predicting quality of life for people with intellectual disability: Results from the ANFFAS study in Italy, J. Intell. Develop. Disabil. 41 (4) (2016) 338–347.

[28] M.A. Verdugo, R.L. Schalock, B. Arias, L.E. Gómez, B. Jordán de Urríes, Calidad de vida. MA Verdugo RL Schalock (Coords.), Discapacidad e inclusión manual para la docencia. (2013) 443–461.

[29] F.S. Carrión-Nessi, M.V. Marcano-Rojas, D.C. Freitas-DeNobrega, S.R.R. Arocha, A.W. Antuarez-Magallanes, Y.J. Fuentes-Silva, Validación del LupusQoL en Venezuela: una medida específica de la calidad de vida en pacientes con lupus eritematoso sistémico, Reumatología Clínica. 18 (6) (2022) 355–360.

[30] M.A. Verdugo, P. Navas, L.E. Gómez, R.L. Schalock, The concept of quality of life and its role in enhancing human rights in the field of intellectual disability, J. Intell. Disabil. Res. 56 (11) (2012) 1036–1045.

[31] M.A. Verdugo, B. Arias, L.E. Gómez, R.L. Schalock, Manual d'aplicació de l'Escala Gencat de Qualitat de vida. Barcelona: Departament d'Acció Social i Ciutadania, Generalitat de Catalunya 2008.

[32] K. Mahmoud, M. Abdel-Nasser, Fast yet accurate energy-loss-assessment approach for analyzing/sizing PV in distribution systems using machine learning, IEEE Trans. Sustain. Energy. 10 (3) (2018) 1025–1033.

[33] D. Liang, D.A. Frederick, E.E. Lledo, N. Rosenfield, V. Berardi, E. Linstead, U. Maoz, Examining the utility of nonlinear machine learning approaches versus linear regression for predicting body image outcomes: The US Body Project I, Body Image. 41 (2022) 32–45.

[34] C. Bentéjac, A. Csérgö, G. Martínez-Muöoz, A comparative analysis of gradient boosting algorithms, Artif. Intell. Rev. 54 (3) (2021) 1937–1967.

[35] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to linear regression analysis, John Wiley Sons, 2021.

[36] G. Biau, Analysis of a random forests model, J. Mach. Learn. Res. 13 (1) (2012) 1063–1095.

[37] L. Torgo, R.P. Ribeiro, B. Pfahringer, P. Branco, SMOTE for Regression. Portuguese conference on artificial intelligence. (2013): 378–389.

[38] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2002) 321–357.

[39] S. Abdulwahab, H.A. Rashwan, M.A. Garcia, M. Jabreel, S. Chambon, D. Puig, Adversarial learning for depth and viewpoint estimation from a single image, IEEE Trans. Circ. Syst. Video Technol. 30 (9) (2020) 2947–2958.

[40] M.I. Dutt, W. Saadeh, A Multilayer Perceptron (MLP) Regressor Network for Monitoring the Depth of Anesthesia. In 2022 20th IEEE Interregional NEWCAS Conference (NEWCAS), pp. 251–255, 2022.

[41] M. Panahi, A. Gayen, H.R. Pourghasemi, F. Rezaie, S. Lee, Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms, Sci. Total Environ. 741 (2020) 139937.