Original research

DIGITAL HEALTH

DIGITAL HEALTH
Volume 9: 1–22

DOI: 10.1177/20552076231178425
journals.sagepub.com/home/dhj
Sage

Developing a prediction model for successful aging among the elderly using machine learning algorithms

Maryam Ahmadi, Raoof Nopour and Somayeh Nasiri

Abstract

Objective: The aging phenomenon has an increasing trend worldwide which caused the emergence of the successful aging (SA)[1] concept. It is believed that the SA prediction model can increase the quality of life (QoL)[2] in the elderly by decreasing physical and mental problems and enhancing their social participation. Most previous studies noted that physical and mental disorders affected the QoL in the elderly but didn't pay much attention to the social factors in this respect. Our study aimed to build a prediction model for SA based on the physical, mental, and specially more social factors affecting SA.

Methods: The 975 cases related to SA and non-SA of the elderly were investigated in this study. We used the univariate analysis to determine the best factors affecting the SA. AB[3], XG-Boost J-48, RF[4], artificial neural network[5], support vector machine[6], and NB[7] algorithms were used for building the prediction models. To get the best model predicting the SA, we compared them using positive predictive value (PPV)[8], negative predictive value (NPV)[9], sensitivity, specificity, accuracy, F-measure, and area under the receiver operator characteristics curve (AUC).

Results: Comparing the machine learning[10] model's performance showed that the random forest (RF) model with PPV = 90.96%, NPV = 99.21%, sensitivity = 97.48%, specificity = 97.14%, accuracy = 97.05%, F-score = 97.31%, AUC = 0.975 is the best model for predicting the SA.

Conclusions: Using prediction models can increase the QoL in the elderly and consequently reduce the economic cost for people and societies. The RF can be considered an optimal model for predicting SA in the elderly.

Keywords

Successful aging, data mining, prediction model, quality of life, elderly

Submission date: 8 October 2022; Acceptance date: 10 May 2023

Highlights

- The machine learning techniques are beneficial in predicting the successful aging with high performance in classification process.
- The ensemble algorithm is an effective solution in predicting the elderly's successful aging with pleasant generalizability.
- By considering more social factors pertained to successful aging, the generalizability of the machine learning models would be increased.

Health Management and Economics Research Center, Health Management Research Institute, Iran University of Medical Sciences, Tehran, Iran

Corresponding author:
Raoof Nopour, Health Management and Economics Research Center, Health Management Research Institute, Iran University of Medical Sciences, School of Health Management & Information Sciences, No. 6, Rashid Yasemi St. Valieasr Ave, Tehran 1996713883, Iran.
Emails: raoof.n1370@gmail.com; nopour.r@iums.ac.ir

Introduction

Aging is a phenomenon that is currently increasing worldwide.[1] So far, this phenomenon and the precise amount of age at which the elderly period begins have not been well defined,[2,3] and there is a difference among nations in this respect. The age of people at which the phenomenon of aging begins has not been introduced by pundits, and they have different viewpoints on the ages beginning with the elderly. In some studies, the ranges of age associated with the elderly are 60, 65, or even 80 years old to higher, in which the phenomenon named fall occurs.[4–6] This phenomenon deteriorates individuals' physical and cognitive health, increases the degree of dependence, and declines the aged group's quality of life (QoL) and self-confidence.[7,8] In recent years, the population's life expectancy has increased due to the decrease in birth rate, improvements in training people, medical and health conditions, technology status, and nutrition betterment.[9,10] It is estimated that the number of elderly around 80 years old will increase from 143 million in 2019 to 426 million by 2050.[11] Also, the life expectancy will increase to 82 years in 2100.[12] Due to the increased elderly population and life expectancy and attention to support the elderly from various physical, mental, and social factors, the successful aging (SA) concept is widespread worldwide by humans leading to a better QoL among the elderly.[13] This trend has raised much attention to understanding this concept and improving the QoL among the elderly.[14]

Some theories that emerged from the 1960s to 1970s described the SA concept. The disengagement theory, introduced by Cumming and Henry, characterized SA as the increased participation of the elderly in social activities.[15] The activity theory discussed engaging the elderly in community relationships was represented by Havighurst. Atchley has defined the continuity theory as describing the ability of the elderly to do personal and social activities in the latest duration of their life, similar to early.[16] A successful aging pattern and tracing its peculiarities are qualitative approaches to the aging period.[17] Rowe and Kahn have introduced that the SA model as a comprehensive approach pertains to people's lifestyles in three dimensions: (a) the low possibility of physical disease and disease-associated conditions; (b) high mental and physical performance capacity; and (c) active participation in social life.[18] They believed that genes do not influence biological aging, and lifestyle has also potentially affected aging.[19] This idea also has been scrutinized and confirmed by the world's great gerontologists.[20–22] Therefore, environmental conditions and focus on the people's lifestyle can be considered predictive factors in determining the SA, leading to a better QoL in the elderly.[18] Using a model for predicting SA in the early stages of the elderly's life or even in middle-aged people leads to higher physical performance, lower physical and mental disorders, and better social situations increasing success in these people's last years.[23] Several studies have shown that predictive models have a potential role in different medical conditions, such as prognosis, diagnosis, and treatment of physical, cognitive, and mental disorders.[24–26] The predictive models' machine learning (ML) can forecast SA betterment more efficiently. So far, little research conducted in the area of ML algorithms application in the SA among the elderly. One study performed by Cai et al. was associated with leveraging the ML approach in predicting the SA. In their study, they focused on developing a predictive model to predict SA among the elderly based on physical fitness tests. In other words, they have focused on the predictive models in the SA in the elderly based on more physical factors and paid less other factors affecting it specially attention to the social factors in this respect.[27,28] Social interactions for the elderly are crucial and should be considered in SA.[29] Therefore, this study aimed to develop the prediction model using ML algorithms for SA among the elderly by embedding social factors in this respect. To this end, we first train selected ML algorithms using various factors especially social factors and compare their performance to get the best model to predict the SA among the elderly. Also, the importance of various factors such as social interactions are investigated using the importance scores obtained by best ML-trained algorithm associated with the SA prediction.

Machine learning technique

ML technique is one of the subfields of artificial intelligence trying to imitate human measurements. This technique deals with integrating the sciences of mathematics, statistics, cognitive sciences, and the computer field to build intelligent systems in this regard. The ML algorithms use past data to learn about the past events that occurred and try to forecast future trends through the pattern that they obtained from the past data[30,31] They have a significant role in real-world applications in which one of them can be considered as predicting the health conditions and diagnosing the diseases by reducing the error rate in performing the physicians' activities.[32] ML algorithms have a beneficial role in treatment plans, for example, can be applied for the electrocardiogram waveform for analyzing, discovering, and classifying purposes.[33] In preprocessing the databases, especially in large-scale data sets, the ML algorithms are applied in analyzing the data set in this step using the feature extraction process, in addition to the pattern that can be extracted from the database.[32,34,35] Deep learning technique is one of the ML subfield named artificial neural network (ANN) and also the best technology that uses the deep neural configuration simulated from the human brain. Hence, they have structures specified by numerous hidden layer performing the extraction of the features and abstraction process at different levels. This technique is regarded as the unsupervised method in which applies the data without any class labels to train the algorithm one layer and the adjacent ones, respectively.[36,37] Deep learning techniques are mainly specified and highlighted by building various architectures including tuple and sequential layers in which the next steps of input processing are done.[38] In the way that ML deals with computers in acting and thinking, deep learning focuses on the computer learning to think using the structures formed similar to the human brain. Although ML needs less calculation power, deep learning requires human interventions. Analyzing unstructured data such as videos, images, sounds, signals, and graphs is more easily done by deep learning approaches. The deep learning techniques are very practical with highest predictive strength in the situation where the number of large data sets are increasing in recent years.[39–43]

Methods

Study roadmap

This study was a descriptive and applied study performed in three steps: data set preprocessing, model development, and assessment. The roadmap of the study is shown in Figure 1.

In this study, we first prepared our data set for analysis. In this respect, first, noisy and redundant data were deleted. After removing the noisy and redundant data, we investigated our data set concerning existing missing values and used the imputation method to replace the missing values with the values predicted by the algorithm. The cases with high-rate data were removed from the study. In the next step, we developed ensemble and non-ensemble ML algorithms to build the prediction model for SA. Finally, the best prediction model for SA was obtained by comparing various performance criteria.

Sample characteristics

We used the 975 records from January 2019 to January 2021 in three elderly centers of Hamdeli, Mehrjooyan, and Hasti in Ahvaz city to investigate the most important factors influencing the SA and build the prediction model. There were 751 and 229 cases that were associated with non-SA and SA cases, respectively. It included 515 and 460 records related to females and males in the database gathered. The sample's specifications are shown in Table 1, describing the frequency of each predictor affecting the SA in each SA and non-SA classes.

Measuring methods of some predictor features

Perceived health condition. This factor assesses the situation of elderly health by the question of how you measure your health. The response to this question can be categorized into three classes: very bad/bad, moderate, and good/ very good, associated with the scores of 0, 1, and 2, respectively.

Official social relationship. It is associated with formal social relationships such as governmental and organizational meetings. It can be measured as five-Likert scales including never = 0, one time in year = 1, one time in month = 2, one time in week = 3, and each day = 4.

Non-official social relationship. This variable deals with informal social relationships. For example, it can be included in friendships, neighbors, religious associations, and meetings. This variable is measured by the five-choice

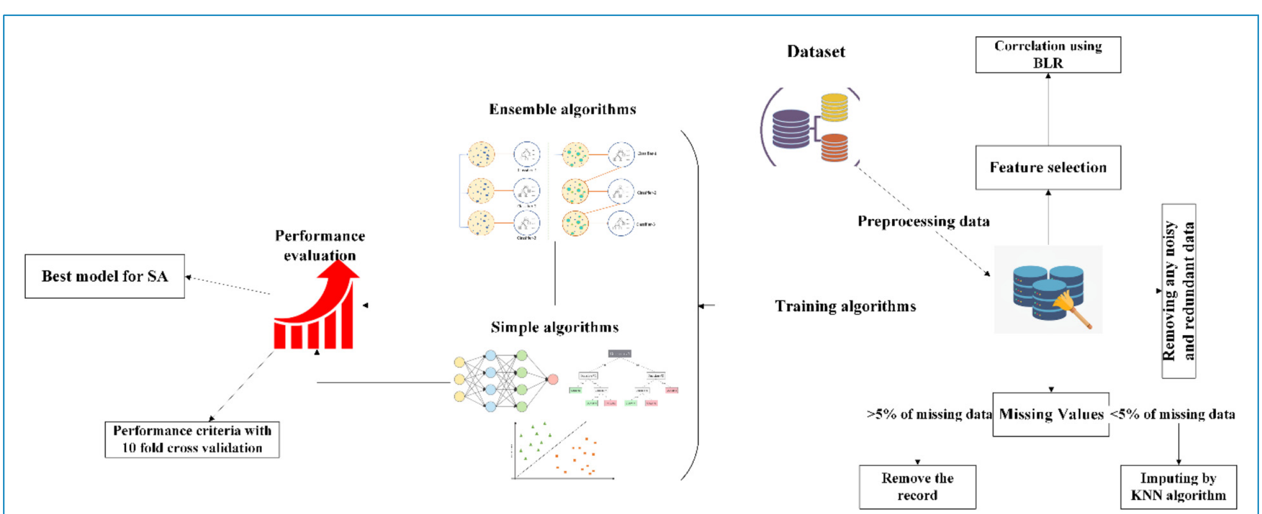

Figure 1. The study's methodology map.

Likert, including never = 0, one time in year = 1, one time in month = 2, one time in week = 3, and each day = 4.

Quality of life. This variable was measured using a self-assessment questionnaire with a qualitative approach. The questionnaire is divided into three types of factors measuring the QoL among the elderly. The factors including physical, mental, and social factors that affect the QoL among the elderly. Then, the elderly declare their response to the question "What is the rate of your quality of life" through the five-choice Likert including "very bad", "bad", "moderate", "good", or "very good". Finally, the five-choice Likert scales were transformed into two states of "pleasant" and "unpleasant" and registered in the database based on the gerontologists' opinions.[44]

Individual independence. It poses 10 questions and computes people's physical health level. This variable specifies each person's ability in various domains of daily functions on one scale ranging from 0 to 100. The 0 to 20 pertain to entirely dependent people, 20 to 60 were very dependent, 61 to 90 were medium dependent, 99 to 91 were little dependent, and the grade of 100 was associated with the fully independent person as Barthel score.[45]

Satisfaction with life scale (SWLS). This criterion was introduced by Diener et al. It included five items assessing the cognitive parts of health and well-being. Each condition consists of seven options and ranges from 1 to 7, from highly disagree to agree strongly, in which a higher score shows better life satisfaction. The validity of this questionnaire was justified by Bayani et al.[46] This variable had two states of pleasant and unpleasant stored in the database.

Lifestyle. Determining the lifestyle depends on the whole grade obtained. It can be obtained by specifying a rate ranging from 42 to 98 associated with the unpleasant lifestyle, 99 to 155 for the medium lifestyle, and 156 to 211 for the pleasant lifestyle. It evaluates exercise, healthy eating, stress management, social and interpersonal relationships, physical activity, and recreation.[47]

Outcome variable. According to the scales defined in the studies and evaluation of gerontology specialists, a successful older adult should have an optimal level of personal independence (getting a score of 90–100 on the independent living score), the highest life satisfaction (getting a score between 20 and 35 on the scale from the Diener's life satisfaction, which has a score between 5 and 35), favorable QoL, and moderate to high lifestyle (scores greater than 99 from lifestyle questionnaire). All the scales of these determinant factors affecting the SA are presented in Table 1; in other words, based on Table 1, an older adult with a little or non-independent situation for individual independence, a pleasant life satisfaction, high QoL, and medium and desirable lifestyles. Also, the condition that an older adult had acquired three criteria of SA from these four ones and another measure was not in a bad situation, for example, the Barthel score was close to 90, was considered as an SA case by gerontologists, and stored in the data set. To investigate other factors affecting the SA among the elderly and extend the SA pattern in this study, we used physical, mental, and specially more social variables affecting the SA in the data set for SA prediction.

Imputation of missing values in the data set. We prepared our data set for establishing the ML models in this study phase. The data set with the rate of missing values was excluded from the study. In this respect, the cases pertained to successful and unsuccessful cases with more than 5% missing values were removed from the data set. For the missing cases with lower than 5% missing values, we used the KNN algorithm with specific K = 1, 3, 5, and more for imputing the missing data in this regard.

Feature selection. To better analyze the data set and reduce the data dimensions, we used the feature selection (FS) method. FS is equaled to select the best variables and decrease the data set dimensions. This process has potential benefits such as removing irrelevant data, precluding form overfitting of the algorithm, increasing the ML training speed, reducing the data redundancy, and higher learning accuracy.[48–50] In this study, we used the univariate regression analysis to select the best factor influencing the SA prediction. Due to the numerous feature's existing in the data set, we set the higher confidence than 95% for selecting the feature affecting the SA. So the $P < 0.05$ were considered statistically meaningful in this regard.

Model development. We used the Weka V 3.9 software to develop the prediction for SA. The ensemble ML algorithms, including random forest (RF), XG-Boost, and Ada-boost, and simple algorithms, including J-48, ANN, support vector machine (SVM), and Naïve Bayes (NB), were used in this respect. All hyperparameter combinations of each algorithm were tested during the models' development, and the best ones with the highest ML performance were selected. The ML algorithms used in this study are described below.

- RF: In this algorithm, as a bagging method, the splitting process occurs through the random selection of the variables. This feature gave the RF a high capability for sample classification, especially in large data set dimensions. This algorithm is suitable for the high-dimensional data set with many variables and uses the voting process to classify the samples with high accuracy. The algorithm performance is shared in its most subtrees with the highest classification capability. By

Table 1. Samples' characteristics associated with each class.

Variables	Frequency for successful cases (%N)	Frequency for unsuccessful cases (%N)	P-value
Age	60–70 (45%) 70–80 (35%) >80 (20%)	60–70 (42%) 70–80 (36%) >80 (22%)	0.2
Sex	Female (57%) Male (43%)	Female (52%) Male (48%)	0.15
BMI	<18.5 (52%) 18–25 (37%) >25 (11%)	<18.5 (47%) 18–25 (38%) >25 (15%)	0.12
Educational level	No literacy (47%) Elementary (30%) Diploma (15%) Academic (8%)	No literacy (51%) Elementary (32%) Diploma (12%) Academic (4%)	0.106
Marital situation	Married (7%) Single (35%) Divorced (8%) Widowed (50%)	Married (5%) Single (37%) Divorced (15%) Widowed (43%)	0.568
Past occupation type	No job (35%) Housekeeper (38%) Retired (20%) Self-employment (7%)	No job (40%) Housekeeper (32%) Retired (18%) Self-employment (10%)	0.378
Income level	Under poverty line (68%) On poverty line (32%)	Under poverty line (76%) On poverty line (24%)	0.17
Number of children	No (12%) 1 (15%) 2–3 (36%) >4 (37%)	No (15%) 1 (15%) 2–3 (41%) >4 (29%)	0.23
Family payment support	Have (41%) Haven't (59%)	Have (28%) Haven't (72%)	0.09
Past family structure	Nuclear (38%) Joint (62%)	Nuclear (41%) Joint (59%)	0.16
Insurance situation	Have (51%) Haven't (49%)	Have (37%) Haven't (63%)	0.04
Arterial blood pressure and cardiac diseases	Have (35%) Haven't (65%)	Have (48%) Haven't (52%)	<0.001
Habit of smoking	Smoking (46%) Non-smoking (54%)	Smoking (58%) Non-smoking (42%)	0.06
Habit of alcohol	Alcoholic (11%) Non-alcoholic (89%)	Alcoholic (15%) Non-alcoholic (85%)	0.231
Medical treatment	Yes (35%) No (65%)	Yes (45%) No (55%)	0.1

(continued)

Table 1. Continued.

Variables	Frequency for successful cases (%N)	Frequency for unsuccessful cases (%N)	P-value
Governmental subsides	Yes (26%) No (74%)	Yes (18%) No (82%)	0.13
Family visits	3 times or more (23%) 1–2 times (38%) Less than 1 time (39%)	3 times or more (18%) 1–2 times (40%) Less than 1 time (42%)	0.08
Family emotional support	Have (55%) Haven't (45%)	Have (43%) Haven't (57%)	0.03
Ability of emotional role	Have (48%) Haven't (52)	Have (37%) Haven't (63%)	0.02
Liver diseases	Have (27%) Haven't (73%)	Have (36%) Haven't (64%)	0.037
Arthritis	Have (3%) Haven't (97%)	Have (5%) Haven't (95%)	0.21
Respiratory diseases	Have (31%) Haven't (69%)	Have (35%) Haven't (65%)	0.09
Digestive diseases	Have (35%) Haven't (65%)	Have (45%) Haven't (55%)	0.06
Musculoskeletal diseases	Have (26%) Haven't (74%)	Have (34%) Haven't (66%)	<0.001
Depression	Have (31%) Haven't (69%)	Have (56%) Haven't (44%)	0.015
Allergy	Have (25%) Haven't (75%)	Have (26%) Haven't (74%)	0.13
Stroke	Have (16%) Haven't (84%)	Have (27%) Haven't (73%)	0.07
Convalescences	Have (77%) Haven't (23%)	Have (45%) Haven't (55%)	0.118
Eye disease	Have (12%) Haven't (88%)	Have (25%) Haven't (75%)	0.057
Skin diseases	Have (16%) Haven't (84%)	Have (23%) Haven't (77%)	0.11
Hearing diseases	Have (33%) Haven't (67%)	Have (45%) Haven't (55%)	0.08
Diabetes	Have (31%) Haven't (69%)	Have (51%) Haven't (49%)	<0.001
Cancer	Have (13%) Haven't (87%)	Have (15%) Haven't (85%)	0.07

(continued)

Table 1. Continued.

Variables	Frequency for successful cases (%N)	Frequency for unsuccessful cases (%N)	P-value
Activity daily living	Have (87%) Haven't (13%)	Have (69%) Haven't (31%)	<0.001
Sporting exercises	Have (67%) Haven't (33%)	Have (55%) Haven't (45%)	0.03
Exercise duration	Haven't (33%), <0.5 hours (38%), >0.5 hours (29%)	Haven't (45%), <0.5 hours (53%), >0.5 hours (2%)	0.021
Exercise type	Laborious exercise (2%), soft exercise (65%), none (33%)	Laborious exercise (2%), soft exercise (53%), none (45%)	0.38
Sexual condition	Healthy (12%) Unhealthy (88%)	Healthy (10%) Unhealthy (90%)	0.14
Sexual relationships	Have (5%) Haven't (95%)	Have (7%) Haven't (93%)	0.16
Stress control capability	Low (16%) High (84%)	Low (38%) High (62%)	0.02
Food habitat	Vegetarian (56%) Non-vegetarian (35%) Occasional non-vegetarian (9%)	Vegetarian (52%) Non-vegetarian (33%) Occasional non-vegetarian (15%)	0.13
Recreational activities	Have (69%) Haven't (31%)	Have (46%) Haven't (54%)	0.04
Healthcare utilization	Have (25%) Haven't (75%)	Have (44%) Haven't (56%)	0.01
Mortality of one of friends or families	Have (44%) Haven't (56%)	Have (73%) Haven't (27%)	<0.001
Instrumental activities of daily living	Dependent (35%) Independent (65%)	Dependent (43%) Independent (57%)	0.08
Individual independence	Entirely dependent (12%) Severe dependent (15%) Medium dependent (36%) Little dependent (20%) Non-dependent (17%)	Entirely dependent (26%) Severe dependent (35%) Medium dependent (28%) Little dependent (6%) Non-dependent (5%)	<0.001
Mobility	No problem (80%) Some problem (15%) Extreme problem (5%)	No problem (47%) Some problem (36%) Extreme problem (17%)	0.03
Perceived health condition	Bad (10%) Moderate (41%) Good (49%)	Bad (31%) Moderate (46%) Good (23%)	0.02
Official social relationships	None (6%) One time per year (13%) One time per month (18%)	None (12%) One time per year (23%) One time per month (36%)	0.04

(continued)

Table 1. Continued.

Variables	Frequency for successful cases (%N)	Frequency for unsuccessful cases (%N)	P-value
	One time per week (45%) Each day (18%)	One time per week (28%) Each day (1%)	
Non-official social relationships	None (4%) One time per year (8%) One time per month (26%) One time per week (51%) Each day (11%)	None (4%) One time per year (21%) One time per month (32%) One time per week (41%) Each day (2%)	0.16
Life satisfaction	Pleasant (63%) Unpleasant (37%)	Pleasant (39%) Unpleasant (61%)	<0.001
The general explanation of lifestyle	Undesirable (7%) Medium (65%) Desirable (28%)	Undesirable (31%) Medium (55%) Desirable (14%)	0.001
Assessment of body pain	Have (6%) Haven't (94%)	Have (11%) Haven't (89%)	0.11
Physical dysfunction	Have (4%) Haven't (96%)	Have (13%) Haven't (87%)	0.01
Fatigue	Have (10%) Haven't (90%)	Have (8%) Haven't (92%)	0.162
Mental dysfunction	Have (2%) Haven't (98%)	Have (25%) Haven't (75%)	<0.001
Social dysfunction	Have (9%) Haven't (91%)	Have (21%) Haven't (79%)	<0.001
Quality of life	Low (17%) High (83%)	Low (56%) High (44%)	<0.001
Physical activity	Low (15%) High (85%)	Low (46%) High (54)	0.016
Debarment activities when occurring disease	Have (86%) Haven't (14%)	Have (59%) Haven't (41%)	0.027

In Table 1, the frequency of variables in each mode is presented per percent in both classes. Also, the differences between each variable in successful and unsuccessful classes are presented at the statistical level.
BMI: body mass index.

distributing the samples in subtrees with different data set attributes, this algorithm poses a high accuracy with noisy data and can prevent overfitting. Generally, the RF advantages can be noted as fast and accurate in the training process with resistance to noisy data as its flexibility.[51–54]

- Ada-boost: The adaptive-boost (AB) is one type of boosting algorithm from the ensemble category using the weak classifiers to classify the cases simultaneously and identify and remove the errors in the classified cases in each classifier sequentially. Some advantages of this algorithm can be enumerated as the generalizability with high accuracy, efficient calculation ability, flexibility for various tasks with sophisticated data, easily adaptable, and the capability to integrate with other algorithms.[55,56]
- J-48: This algorithm is the newer model but similar to the ID3 algorithm. Its generic name is C4.5, but in Weka software is named J-48. This algorithm uses the info gain to split the tree using the variable with the highest entropy. The tree is built based on features capability in classifying samples with the highest entropy

recognized during the training process. This algorithm has more pruning features than others, which allows the J-48 to have less capability to be overfitted than other ML models. Generally, some beneficial algorithm features are: using the confidence factor to set the tree size and preclude the overfitting process, the capability of working with continuous variables contrary to other decision tree algorithms, extracting rules with the maximum discrimination between the output classes, and the ability to work with missing values.[57–59]

- ANN: The ANN tries to imitate human behavior, and its structure consists of neurons and weights between them to transfer messages around the network like humans. Generally, the ANN includes three layers, including input layers which receive the input data such as signals, images, or any other data type. The hidden layer is responsible for calculation in ANN, receiving data from input layers, and considering the calculation nodes that process the input data. The results of the computation process that occurred in the hidden layers will transfer to the output layers. The last layer is the output layer, and in this layer, the users can see the output of the ANN, which is represented by this layer. So far, ANN has various applications for solving highly complex computational problems in medical conditions.[60–64]
- XG-Boost: Another boosting algorithm that uses several tree algorithms for performance betterment. In this algorithm, similar to other boosting methods, the learning process occurs sequentially. The current algorithm, in this way, recognizes and deletes the incorrectly classified cases by previous cases, so the error rate is lower than the simple algorithm during the training process. Some benefits of this algorithm can be enumerated as scalability, running speed, parallel training, sparse optimization, regularization, bagging, numerous loss function, and early stopping.[65,66]
- SVM: The SVM is one of the classification and regression algorithms using the hyperplane concept for performing the classification task. This algorithm uses the mapping process to convert the low-dimensional data points to higher dimensions for classifying the cases, which can be used through the kernel functions. Different kernel types are used for various types of data sets in terms of the complexity of data so that the SVM algorithms can be categorized as linear and non-linear types.[35,67,68]
- NB: The NB algorithm is a statistical classification based on the Bayesian theory using probabilistic hypothetic concepts. It determines the probability of each sample belonging to different data classes using Equation 1. The new samples will be classified into one category having the highest chance. In this algorithm, each variable's occurrence is independent of determining the dependent variable.[31,69]

$$\mathrm{P(A/B)} = \frac{\mathrm{P(B/A)}}{\mathrm{P(B)}}(\mathrm{P(A)}) \tag{1}$$

In this equation, the P (B/A) is the posterior probability meaning the probability of occurring the tuple (B) in the condition that A as a specific class occurs, and P (A) means the probability of occurring the class of B occurs.

Performance metrics. To get the best model to determine the success among the elderly, we evaluated the performance of selected ML algorithms using the confusion matrix and positive predictive value (PPV), negative predictive value (NPV), sensitivity (Equation 4), specificity (Equation 5), accuracy (Equation 6), F-measure (Equation 7), and area under the receiver operator characteristics (ROC) curve (AUC) obtained from the confusion matrix. The true positive (TP) and true negative (TN) in the confusion matrix are the successful and unsuccessful cases correctly classified by the decision models. The false negative (FN) and false positive (FP) are the successful and unsuccessful adults incorrectly categorized by the algorithms.[70]

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{2}$$

$$\mathrm{NPV} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}} \tag{3}$$

$$\mathrm{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{4}$$

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \tag{5}$$

$$\mathrm{Accuracy} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FN}+\mathrm{FP}+\mathrm{TN}} \tag{6}$$

$$\text{F-measure} = 2*\left(\frac{\mathrm{Sensitivity}*\mathrm{Specificity}}{\mathrm{Sensitivity}+\mathrm{Specificity}}\right) \tag{7}$$

K-fold cross-validation. This technique uses the train data splitting process while learning the ML algorithms. In this subsampling way, the training data is divided into equal sections, and the accuracy of all split train data is measured separately. The final performance of the ML models is the average of all sections' accuracy. This method is beneficial for preserving the generalizability of the model's performance due to the stratified random sampling.[71] In this study, the 10-fold cross-validation measures the ML models' capability.

External validation assessment. In this study, to validate and measure the generalizability of our model for predicting the SA and non-SA cases, we used external data irrespective of the data on those three centers. In this regard, we used the 147 instances in Mehravaran Shamal and Panhkala-Jonubi of elderly centers in Sari city of Mazandaran province. Among these cases, 96 and 51 cases were associated with non-SA and SA cases, respectively. We compared the performance of our model for predicting the SA in this study and the capability of predicting the these externaldata cases. In this respect, we used the confusion matrix and ROC curve.

Results

Preprocessing of the data set

After removing the records with more than 5% missing values in the output class, seven and eight cases associated with the successful and unsuccessful cases were excluded from the study. So 222 and 743 cases pertained to successful and unsuccessful cases remained in the study. The 127 and 95 cases belonged to the women and men SA cases, and 386 and 357 cases were associated with the women and men non-SA cases, respectively. Most of the cases used for this study were related to the elderly aged 60 to 70. In this respect, 45% and 42% frequency of SA and non-SA elderly cases were associated with this age range group. Through imputing the missing cases using the KNN algorithm with a specific amount of K = 1, 3, 5, and more, the data of 37 cases with less than 5% missing data were replaced by predicted values obtained by this algorithm. The principle component analysis of some features affecting is shown in Figure 2. The points in the upper and lower portions of the figure represent SA and non-SA cases, respectively.

In Figure 2, the correlation between data points associated with the SA and non-SA cases with transformed input data is shown in two dimensions in which two class types of data points are clustered into two groups using the transformed input data.

The results of determining the correlation of predictors with the output class using the binary logistics regression

Figure 2. Principal component analysis for clustering SA data point. SA: successful aging.

(BLR) as univariate regression analysis are shown in Table 2.

In Table 2, β implies the correlation of each factor affecting the SA in BLR. The more amount of β is equivalent to more correlation of factors in predicting SA and therefore is considered the most critical factor for predicting SA among the elderly. In this study, we considered the $\beta = 0.5$ as a threshold for accepting the factors on SA among the elderly. So the factors with $\beta > 0.5$ at $P < 0.05$ were considered acceptable for predicting SA among the elderly. The training of all algorithms was performed by using these variables in the data set. Based on the information provided in Table 2, the variables including age ($P = 0.01$) ($\beta = 0.5$), income level ($P = 0.01$) ($\beta = 0.55$), arterial blood pressure and cardiac diseases ($P < 0.01$) ($\beta = 0.63$), the habit of smoking ($P < 0.01$) ($\beta = 0.54$), family emotional support ($P < 0.01$) ($\beta = 0.58$), the ability of emotional role ($P < 0.01$) ($\beta = 0.5$), depression ($P = 0.01$) ($\beta = 0.54$), diabetes mellitus ($P = 0.03$) ($\beta = 0.52$), stress control capability ($P = 0.03$) ($\beta = 0.58$), individual independence ($P < 0.01$) ($\beta = 0.59$), mortality of one of friends or families ($P < 0.01$) ($\beta = 0.53$), perceived health condition ($P < 0.01$) ($\beta = 0.5$), official social relationships ($P < 0.01$) ($\beta = 0.54$), non-official social relationships ($P < 0.01$) ($\beta = 0.59$), life satisfaction ($P < 0.01$) ($\beta = 0.56$), explanation of lifestyle ($P < 0.01$) ($\beta = 0.64$), mental dysfunction ($P < 0.01$) ($\beta = 0.56$), social dysfunction ($P < 0.01$) ($\beta = 0.52$), and QoL among the elderly ($P < 0.01$) ($\beta = 0.5$) were considered as the most important factors influencing the SA among the elderly as so used for model construction. Generally, assessing all variables indicated that they obtained a confidence of more than 80% ($P = 0.15$) in predicting the SA.

Model's performance assessment

The results of measuring and comparing data mining algorithms' performance by criteria including PPV, NPV, sensitivity, specificity, accuracy, and F-score obtained for classifying the SA, and non-SA cases among the elderly with the best-adjusted hypermeters pertained to them, are shown in Table 3. The 10-fold cross-validation method was used to split the training data set and report the performance assessment results.

Similar to the previous study, the accuracy and the AUC are the best performance criteria to assess the ML-trained algorithms.[72,73] As shown in Table 3, by considering the 10-fold cross-validation for splitting the training data set, the RF model as the ensemble algorithm used the decision stump tree with a maximum depth of eight with a maximum of 20 iterations for training the algorithm. This algorithm was recognized as the best performer in predicting SA among the elderly, and the results of its performance criteria calculation are bold in Table 3. We considered the average performance criteria in classifying the SA and non-SA cases in this respect. So the RF-trained algorithm with a mean accuracy of 97.05% for SA and non-SA cases gained the best performance than other ML algorithms. Also, other performance critera calculation results obtained by this algorithm were better than others. Based on the information in Table 3 the XG-Boost algorithm with a maximum depth of tree = 6, using REP-Tree and Gb-tree for training and boosting, and minimum child weight of one, on average, with an accuracy of 94.47%, gained a pleasant performance in this regard. In contrast, the NB algorithm, with an accuracy of 72.79%, achieved the lowest performance among others.

To depict the predictive performance capability of each of ensemble and non-ensemble ML algorithms by considering the 10-fold cross-validation, we used the ROC of all of them. Figure 3 shows the ROC curve pertained to all ML algorithms (the vertical and horizontal vertices indicate true positive rate (TPR) and false positive rate (FPR), respectively.)

The ROC curve indicates the algorithms' capability in classifying the cases. If this curve is closer to TPR vertices, the area under the ROC curve (AUC) will be increased. This increase indicates the higher capability of the model in classifying the cases. As shown in Figure 3, the RF model with AUC = 0.975 obtained the highest AUC than other algorithms. Also, the XG-Boost algorithm with AUC = 0.932 got the pleasant capability in classifying the SA and non-SA cases. The study results showed that the ROC curves of the RF and XG-Boost models were closer to the TPR vertices than other ML models. Therefore, these models have better prediction capability than others in SA with the highest AUC. In contrast, The NB algorithm with AUC = 0.655 had the lowest performance with an increasing rate of random classification capability than other models. Also, the Ada-boost mode, with AUC = 0.858, had a pleasant performance for predicting the SA with AUC > 0.8. We investigated the ROC curve of all ML algorithms. We concluded that the ensemble algorithms gained better discriminant capability in classifying the SA and non-SA cases than the non-ensemble algorithms. Generally, the study results showed that the RF as ensemble ML algorithm with PPV = 90.96%, NPV = 99.21%, sensitivity = 97.48%, specificity = 97.14%, accuracy = 97.05%, F-score = 97.31%, and AUC = 0.975 was recognized as the best performing model for predicting the SA among the elderly.

Predictor's assessment

In this study, we ranked our SA predictor features through the RF classifier algorithm as the best algorithm for predicting SA among the elderly. The feature importance is calculated by node impurity reduced weighted by the probability of achieving to that node. The node probability is computed by dividing the number of cases for achieving the node by all cases. The higher value indicates the more important

Table 2. The univariate regression technique to select favorable attributes.

Variables	β	Odd ratio	95% of the confidence interval	P-value
Age	**0.50**	**1.55**	**[1.49–1.62]**	**0.01***
Sex	0.21	1.23	[0.986–1.45]	0.06
BMI	0.13	1.14	[1.03–1.25]	0.06
Educational level	0.21	1.23	[1.18–1.33]	0.09
Marital situation	0.17	1.53	[1.37–1.62]	0.09
Past occupation type	0.19	1.53	[1.31–1.75]	0.07
Income level	**0.55**	**1.00**	**[0.68–1.35]**	**0.01***
Number of children	0.22	1.14	[0.98–1.29]	0.15
Family payment support	0.31	1.57	[1.32–1.78]	0.07
Past family structure	0.16	0.86	[0.55–1.25]	0.11
Insurance situation	0.02	1.74	[1.69–1.8]	0.09
Arterial blood pressure and cardiac diseases	**0.63**	**1.35**	**[1.33–1.37]**	**<0.01***
Habit of smoking	**0.54**	**1.08**	**[1.06–1.08]**	**<0.01***
Habit of alcohol	0.13	1.63	[1.41–1.85]	0.06
Medical treatment	0.33	0.86	[0.55–1.212]	0.12
Governmental subsidies	0.17	1.73	[1.56–1.91]	0.07
Family visits	0.13	0.89	[0.85–0.92]	0.07
Family emotional support	**0.58**	**0.83**	**[0.78–0.89]**	**<0.01***
The ability of emotional role	**0.50**	**1.08**	**[1.06–1.11]**	**<0.01***
Liver diseases	0.16	1.79	[1.57–2.05]	0.1
Arthritis	0.47	1.08	[1.03–1.15]	0.09
Respiratory diseases	0.06	0.88	[0.82–0.96]	0.07
Digestive diseases	0.10	1.23	[1.15–1.34]	0.1
Musculoskeletal diseases	0.16	0.89	[0.621–1.08]	0.13
Depression	**0.54**	**1.67**	**[1.58–1.75]**	**0.01***
Allergy	0.08	1.12	[1.08–1.17]	0.14
Stroke	0.10	1.68	[1.47–1.89]	0.15

(continued)

Table 2. Continued.

Variables	β	Odd ratio	95% of the confidence interval	P-value
Convalescences	0.15	1.1	[0.98–1.25]	0.09
Eye disease	0.15	1.16	[1.06–1.25]	0.13
Skin diseases	0.42	1.05	[0.996–1.07]	0.10
Hearing diseases	0.15	1.12	[1.06–1.18]	0.12
Diabetes	**0.52**	**1.89**	**[1.86–1.92]**	**0.03***
Cancer	0.12	1.49	[1.27–1.66]	0.07
Activity daily living	0.05	1.04	[1.01–1.07]	0.08
Sporting exercises	0.13	1.58	[1.48–1.65]	0.12
Exercise duration	0.42	0.91	[0.89–0.95]	0.05
Exercise type	0.18	1.51	[1.38–1.65]	0.14
Sexual condition	0.12	0.86	[0.75–0.97]	0.11
Sexual relationships	0.19	1.23	[1.17–1.29]	0.07
Stress control capability	**0.58**	**1.72**	**[1.68–1.76]**	**0.03***
Food habitat	0.16	1.51	[1.47–1.55]	0.07
Recreational activities	0.05	1.05	[1.01–1.11]	0.13
Healthcare utilization	0.12	0.91	[0.75–1.18]	0.10
Mortality of one of friends or families	**0.53**	**1.28**	**[1.26–1.31]**	**<0.01***
Instrumental activities of daily living	0.11	1.01	[0.985–1.04]	0.13
Individual independence	**0.59**	**1.21**	**[1.15–1.28]**	**<0.01***
Mobility	0.56	1.13	[1.05–1.17]	0.06
Perceived health condition	**0.50**	**1.12**	**[1.08–1.16]**	**<0.01***
Official social relationships	**0.54**	**1.04**	**[1.01–1.06]**	**<0.01***
Non-official social relationships	**0.59**	**1.7**	**[1.68–1.72]**	**<0.01***
Life satisfaction	**0.56**	**0.95**	**[0.93–0.97]**	**<0.01***
The general explanation of lifestyle	**0.64**	**1.83**	**[1.78–1.87]**	**<0.01***
Assessment of body pain	0.19	1.56	[1.42–1.72]	0.13
Physical dysfunction	0.08	1.74	[1.71–1.78]	0.13
Fatigue	0.07	0.98	[0.93–1.04]	0.10

(continued)

Table 2. Continued.

Variables	β	Odd ratio	95% of the confidence interval	P-value
Mental dysfunction	**0.56**	**1.02**	**[0.98–1.08]**	**<0.01***
Social dysfunction	**0.52**	**1.35**	**[1.27–1.41]**	**<0.01***
Quality of life	**0.57**	**1.34**	**[1.29–1.405]**	**<0.01***
Physical activity	0.18	1.01	[0.977–1.06]	0.06
Debarment activities when occurring disease	0.11	0.89	[0.83–0.96]	0.07

BMI: body mass index. The correlation (β), odd ratio, confidence interval, and P-value for critical factors are bold.

feature. In all decision trees, the node importance is calculated using Gini importance as follows:

$$nx_i = w_i c_i - w_{left\ (i)} c_{left\ (i)} - w_{right\ (i)} C_{right\ (i)} \tag{8}$$

In Equation 8, the (nx_i) is the importance of the node (i), W_i is the weighted number of cases to achieve the node of (i), Ci is equaled to the impurity value associated with the node (i), and left (i) and right (i) are the left and right child nodes of the node (i), respectively. Therefore, the feature importance of the feature (x) can be calculated as the following formula:

$$FI(x) = \frac{\sum_{i:\ \text{node i splits on the factor (x)}} nx_i}{\sum_{K \in \text{all nodes}} nx_k} \tag{9}$$

In Formula 9, The FI and nx_i indicate the importance of the predictor and node (i). Also, in the RF algorithm, the feature importance value can be calculated by averaging the (FI) from all trees built.

$$FI(x)\text{in RF} = \frac{\sum_{i \in \text{all trees}} FI_{xi}}{T} \tag{10}$$

FI(x) in RF is the feature importance of the feature (x) calculated from all trees, FI (xi) is the feature importance of the feature (x) in the tree (i), and T is the total number of trees used for the ensemble. We reported relative importance (RI) in this study.

The results of ranking the SA predictors using the RI by RF model as the best performer for predicting SA among the elderly are indicated in Table 4.

Table 4 shows that the QoL among the elderly (RI = 0.47) is the most crucial factor influencing SA among the elderly. Also, other factors, including life satisfaction (RI = 0.44), official social relationships (RI = 0.38), lifestyle (RI = 0.4), non-official social relationships (RI = 0.35), and social dysfunction (RI = 0.31), with RI > 0.3, were considered as the essential factors in this respect. The familial and governmental support factors with an average RI = 13.6 gained the lowest score in this respect. Generally, based on the factors gained from the RF model, we concluded that social factors are crucial in predicting SA among the elderly.

External validation appraisal

The results of assessing the RF-trained model by confusion matrix in classifying external cases pertained to the SA and non-SA cases obtained from two different centers in the Mazandaran province as the external validation cohort are shown in Table 5.

Based on Table 5, The RF with TP = 45, FN = 6, FP = 8, and TN = 88 gained an accuracy of 0.90. Comparing the RF when the model was built with the accuracy of 97.05% and the model tested by external data demonstrated that the model's accuracy was not reduced significantly. For internal validation, as obtained by the current study, we got the AUC = 0.975 for RF-trained algorithm. By classifying the new external cases from two elderly centers belonging to Mazandaran province by the best-trained algorithm, the RF model gained $AUC_{-external} = 0.896$ as the result. This little difference in the model's degree of accuracy in the two states shows our prediction model's pleasant generalizability in different test settings. The ROC curve of RF in two states (Figure 4) confirms this generalizability.

Discussion

In this study, we aimed to build a predictive model for SA by using ML models. We first investigated the relationship of each factor with SA and got the essential factors statistically; for this purpose, we used the univariable analysis. Next, after preprocessing the data and gaining the best factors affecting the SA, we selected the appropriate ML models to implement the model predicting the SA. To this end, we chose the AB, XG-Boost, J-48, RF, ANN, SVM, and NB algorithms. Finally, the best predictive model for SA is obtained based on comparing and evaluating all the ML models' performance using various criteria.

Table 3. The results of the performance criteria of selected models.

Algorithm	Important hyperparameters used	Class	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	F-score (%)
AB	Number of iterations= "15"; Classifier= "Rep-tree"; Use resampling= "true";	SA	68.12	92.26	84.75	88.15	86.15	86.42
		Non-SA	78.17	96.15	76.38	94.35	92.2	84.42
		Average	73.14	94.21	80.56	91.25	89.17	85.57
XG-Boost	Maximum depth ="6"; Classifier= "REP-Tree"; Booster type= "gb-tree"; Min of child weight= "1";	SA	79.27	98.16	95.06	96.12	95.52	95.59
		Non-SA	83.12	97.85	92.27	92.16	93.43	92.21
		Average	81.19	98	93.66	94.14	94.47	93.90
SVM (linear)	Kernel type= "Linear"; C= "10"; Gamma= "1.0";	SA	48.68	88.25	63.12	76.15	75.55	69.03
		Non-SA	57.15	92.77	79.25	85.95	84.35	82.46
		Average	52.91	90.01	71.17	81.5	79.95	75.99
ANN	Learning rate= "0.3"; Hidden layer= "15"; Normalize attributes and numeric class= "true"; Training time= "500";	SA	42.91	85.14	56.12	74.75	74.41	64.11
		Non-SA	53.07	91.36	68.25	85.14	78.15	75.77
		Average	47.99	88.25	62.18	79.94	76.28	69.95
SVM (RBF)	Kernel type= "RBF"; C= "15"; Gamma= "1.0"; RBF gamma= '0.1';	SA	52.15	86.15	73.42	80.88	78.52	76.97
		Non-SA	63.33	97.71	79.51	87.14	85.41	83.15
		Average	57.74	91.93	76.46	84.01	81.96	80.06
J-48	Confidence factor = "0.2"; Number of folds= "3"; Minimum number of objects per leafs= "2"; Binary splitting =" true";	SA	51.27	88.26	68.75	79.54	76.98	73.75
		Non-SA	58.67	93.92	75.5	84.63	82.8	79.80
		Average	54.97	91.09	72.13	82.08	79.89	76.78
RF	Ensemble decision tree type= "Decision Stump"; Maximum depth= "8"; Number of iterations= "20";	SA	**87.52**	**98.43**	**96.12**	**95.57**	**96.65**	**95.84**
		Non-SA	**94.41**	**100**	**98.85**	**98.71**	**97.45**	**98.78**
		Average	**90.96**	**99.21**	**97.48**	**97.14**	**97.05**	**97.31**
NB	Estimator type= "simple"; Search algorithm= "K2 algorithm";	SA	39.75	85.12	57.73	74.45	69.92	65.03
		Non-SA	48.12	89.9	62.15	80.37	75.67	70.10
		Average	43.93	87.51	59.94	77.41	72.79	67.56

PPV: positive predictive value; NPV: negative predictive value; SA: successful aging. The performance criteria for best ML-trained algorithm are bold.

So far, few studies have been conducted concerning SA prediction models, especially in the ML domain. Cai et al. built a prediction model for SA based on the more physical factors using machine learning approaches. Their studies showed that the deep learning model with AUC = 90%, specificity = 93.1%, and accuracy = 83.9% gained the best performance to predict the SA. Also, the variables of age, sitting and standing for 30 seconds, arm curl, and reaction

Table 4. The importance of each factor affecting SA by RF algorithm.

Variable type	Variable name	Number of nodes used in the RF model	RI	Average RI in each group
Demographic factors	Age	967	0.3	0.176
	Sex	684	0.17	
	BMI	714	0.18	
	Educational level	1517	0.29	
	Marital situation	324	0.11	
	Past occupation type	401	0.15	
	Income level	303	0.12	
	Number of children	323	0.11	
	Insurance situation	382	0.16	
Familial and governmental support factors	Past family structure	341	0.10	0.136
	Family payment support	419	0.12	
	Family visits	840	0.18	
	Family emotional support	646	0.15	
	Governmental subsidies	754	0.13	
Medical condition factors	Medical treatment	515	0.16	0.167
	Arterial blood pressure and cardiac diseases	684	0.12	
	The ability of emotional role	722	0.28	
	Liver diseases	583	0.18	
	Arthritis	424	0.14	
	Respiratory diseases	744	0.16	
	Digestive diseases	606	0.12	
	Musculoskeletal diseases	1707	0.19	
	Depression	722	0.15	
	Allergy	1098	0.16	
	Stroke	1616	0.10	
	Convalescences	1201	0.13	

(continued)

Table 4. Continued.

Variable type	Variable name	Number of nodes used in the RF model	RI	Average RI in each group
	Eye disease	1783	0.22	
	Skin diseases	362	0.13	
	Hearing diseases	369	0.15	
	Diabetes mellitus	733	0.31	
	Cancer	653	0.15	
Physical factors	Activity daily living	1479	0.24	0.19
	Sporting exercises	845	0.11	
	Exercise duration	846	0.13	
	Exercise type	632	0.16	
	Instrumental activities of daily living	671	0.24	
	Individual independence	1452	0.12	
	Mobility	1130	0.19	
	Healthcare utilization	831	0.26	
	Physical activity	1994	0.22	
	Debarment activities when occurring disease	1432	0.17	
	Physical dysfunction	562	0.17	
	Fatigue	1523	0.13	
	Assessment of body pain	1066	0.13	
Mental factors	Mortality of one of friends or families	985	0.16	0.237
	Stress control capability	840	0.28	
	Perceived health condition	1746	0.28	
	Mental dysfunction	651	0.23	
	Life satisfaction	1406	0.44	0.44
	The general explanation of lifestyle	1686	0.40	0.40
Epidemiological and environmental factors	Food habitat	586	0.20	0.147
	Recreational activities	1704	0.12	

(continued)

Table 4. Continued.

Variable type	Variable name	Number of nodes used in the RF model	RI	Average RI in each group
	Habit of smoking	705	0.13	
	Habit of alcohol	621	0.14	
Social factors	Non-official social relationships	1274	0.35	0.346
	Official social relationships	1228	0.38	
	Social dysfunction	1215	0.31	
	Quality of life	522	0.47	0.47
Sexual factors	Sexual condition	367	0.28	0.275
	Sexual relationships	405	0.25	

PPV: positive predictive value; NPV: negative predictive value; SA: successful aging; RF: random forest; RI: relative importance; BMI: body mass index.

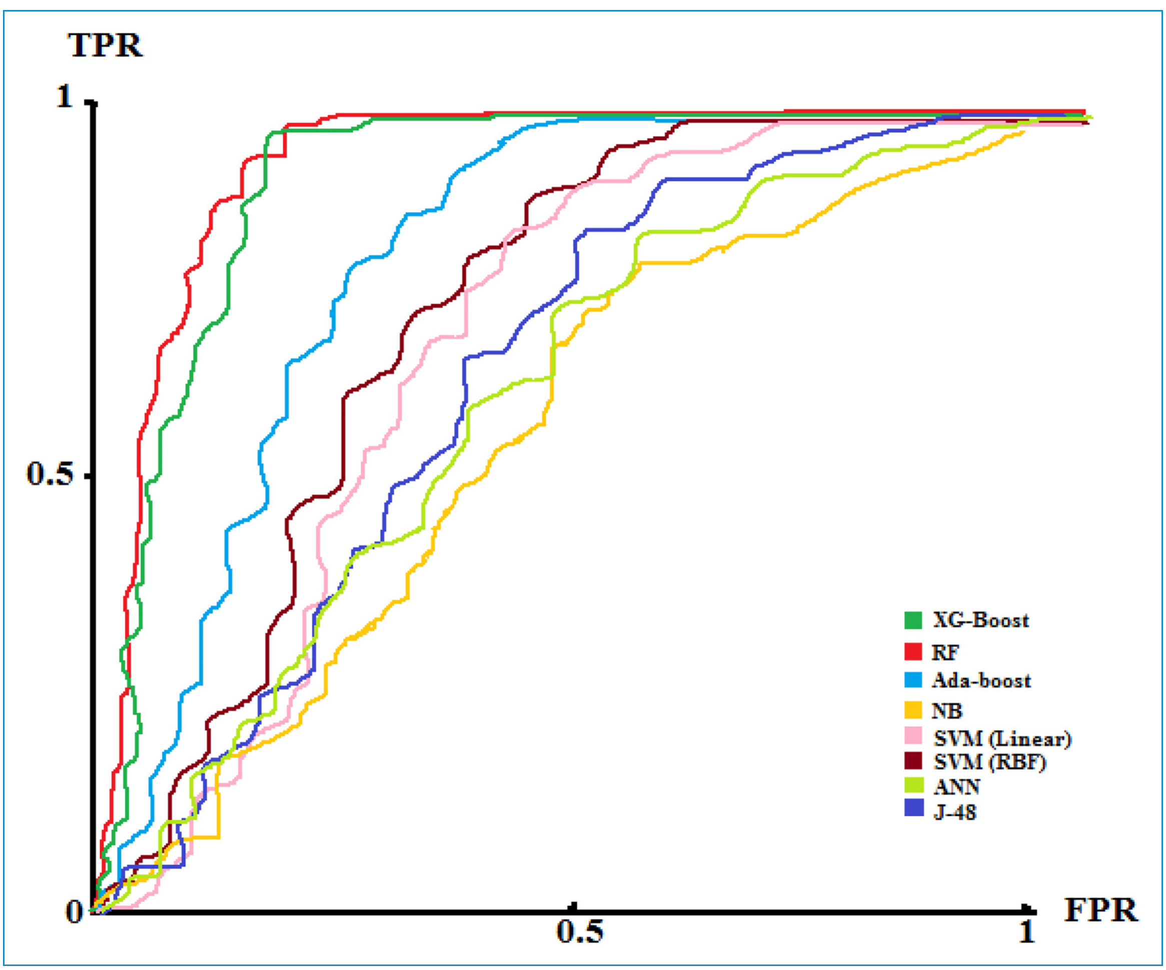

Figure 3. The ROC of the models. ROC: receiver operator characteristics.

time were recognized as essential factors affecting the SA by the model.[28] In our study, the more social factors affecting SA prediction are considered in addition to the physical aspects. Based on these factors, the RF model with PPV = 90.96%, NPV = 99.21%, sensitivity = 97.48%, specificity = 97.14%, accuracy = 97.05%, F-score = 97.31%, and AUC = 0.975 obtained the best performance predicting the SA. The RF model extracted variables of life satisfaction, QoL, and official and non-official social relationships as the best factors.

Table 5. Confusion matrix for external data cases.

	Predicted as SA	Predicted as non-SA
Real SA cases	45	6
Real non-SA cases	8	88

SA: successful aging.

Li et al. estimated the QoL using physical factors, laboratory measurements, demographic factors, and healthy behaviors based on machine learning algorithms. They found that age, walking speed, sleep duration, handgrip strength, body mass index, blood pressure, and sitting and standing duration are the most critical factors affecting QoL.[74] In the present study, the QoL is considered one of the dimensions affecting the SA, and social and mental factors are considered in addition to the introduced factors in that study. Kyoung-Sae Na was surveyed to predict cognitive impairment 2 years later in the elderly. They showed that the gradient boosting machine (GDM) algorithm with sensitivity = 0.96, specificity = 0.825, and AUC = 0.921 gained the best ability for cognitive impairment prediction, and age was considered the best factor obtained from GDM.[75] In the current study, age (RI = 0.3) is considered as an important factor for predicting SA.

Byeon constructed the predictive model for the social participation of the elderly using the ANN and Quest algorithms. The ANN with AUC = 0.718 and Quest with 0.754

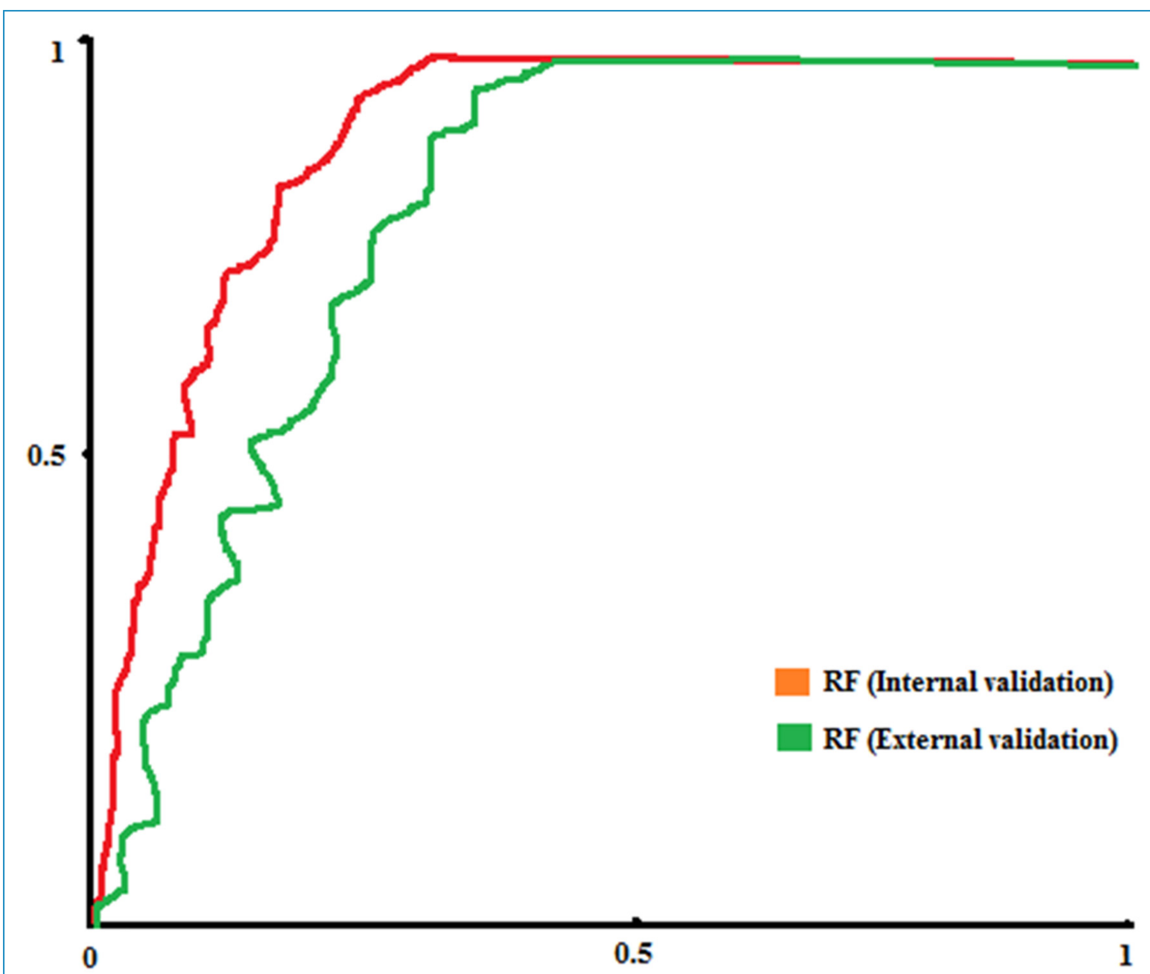

Figure 4. The ROC of RF in internal and external validation states. ROC: receiver operator characteristics; RF: random forest.

with 10-fold cross-validation were relatively efficient in classifying the elderly as a social activity. The subjective health situation with Normalized Importance (NI) = 100% and age with NI = 89% were reported as the essential factors in this regard.[76] In the current study, the factors affecting SA were noticed more with the inclusion of diseases that may exist in the elderly and constrained physical and social participation. Moreover, more ML models, including ensemble and non-ensemble, are utilized to predict SA among the elderly.

This study benefits from more mental and social factors to build the prediction model for SA using ML models, contrary to most previous studies investigating SA. In addition, we used the ensemble technique to construct the prediction model through various decision tree algorithms and observed the enhancement of our prediction model for SA. Our model can predict SA in the elderly earlier to enhance the lifestyle and QoL in this age group. The mental and physical disorders will be detected in the elderly and will be improved. Therefore, the probability of SA in them will be increased.

Limitation and suggestion

The limitations of our study were first using the qualitative variables in the database as the factors affecting the SA. Using the qualitative variables may somewhat decrease the performance of our ML models. Second, we were forced to fill in lost data using regression. Therefore, this may affect the performance and generalizability of our ML models. In the future, to get a more optimal SA predictive model for better generalization, we suggest training ML models with more data from more nursing centers and filling in missing data in features by acquiring the actual data instead of imputing methods.

Conclusion

This study used AB, XG-Boost, J-48, RF, ANN, SVM, and NB algorithms to develop prediction models for SA. The research shows that life satisfaction, QoL, and official social relationships are the best factors affecting SA. We conclude that the RF model can assist the gerontologist in increasing the speed of assessing the situation of SA in elderlies and introduce the best solution for improving it by considering the physical, mental, and specially social factors using the best knowledge extracted from the data.

Acknowledgements: The authors thank the people who assisted us in all steps of this study.

Contributorship: Study concept and design: MA, and RN; analysis and interpretation of data: RN and SN; drafting of the manuscript: MA; critical revision of the manuscript for important intellectual content: MA, RN, and SN; and statistical analysis: RN.

Declaration of Conflicting Interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval: This article was extracted from the research project approved by the research committee affiliated with the Iran University of Medical Sciences (IUMS) with the ethical code of IR.IUMS.REC.1401.780.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article.

Guarantor: Raoof Nopour.

Informed Consent: Due to the retrospective nature of this study, it's waived form the informed consent.

ORCID iD: Raoof Nopour https://orcid.org/0000-0003-3770-2375

Supplemental Material: Supplemental material for this article is available online.

References

1. Machado AH, Casarin VA, De Mirinda RL, et al. O comportamento de consumo de vestuário de um grupo de terceira idade: estudo de caso em um Município do Estado do Rio Grande do Sul. *Revista Espacious* 2016; 37: 19.
2. Coto M, Lizano F, Mora S and Fuentes J (eds). Social media and elderly people: research trends. In: *Social computing and social media applications and analytics*. Cham: Springer International Publishing, 2017.
3. Pisano M, Ceresoli M, Cimbanassi S, et al. 2017 WSES and SICG guidelines on acute calcolous cholecystitis in elderly population. *World J Emerg Surg* 2019; 14: 1–16.
4. Wiggins T, Markar SR, Mackenzie H, et al. Evolution in the management of acute cholecystitis in the elderly: population-based cohort study. *Surg Endosc* 2018; 32: 4078–4086.
5. Yu Z, Kong D, Peng J, et al. Association of malnutrition with all-cause mortality in the elderly population: a 6-year cohort study. *Nutr Metab Cardiovasc Dis* 2021; 31: 52–59.
6. Prasad GL, Anmol N and Menon GR. Outcome of traumatic brain injury in the elderly population: a tertiary center experience in a developing country. *World Neurosurg* 2018; 111: e228–ee34.
7. Sotoudeh GR, Mohammadi R, Mosallanezhad Z, et al. The prevalence, circumstances and consequences of unintentional falls among elderly Iranians: a population study. *Arch Gerontol Geriatr* 2018; 79: 123–130.
8. Mentis M, Gouva M, Antoniadou E, et al. Psychological and physical problems in elderly people with problems of falls. *Eur Psychiatry* 2017; 41: S659–SS60.

9. Chalise H. Provincial situation of elderly population in Nepal. *Aging Sci Res* 2020; 1: 9–11.
10. Afshar PF, Asgari P, Shiri M, et al. A review of the Iran's elderly status according to the census records. *Galen Med J* 2016; 5: –6.
11. Pilleron S, Soto-Perez-de-Celis E, Vignat J, et al. Estimated global cancer incidence in the oldest adults in 2018 and projections to 2050. *Int J Cancer* 2021; 148: 601–608.
12. Purnomo E, Komalasari R and Manurung EI. Elderly Patients Quality of Life at Outpatient Department Hospital X, Banten. In: The 1ST ICGH international conference on global health, Jakarta. Universitas Indonesia: Faculty of Public Health Universitas Indonesia; July, 2017.
13. Estebsari F, Dastoorpoor M, Khalifehkandi ZR, et al. The concept of successful aging: a review article. *Curr Aging Sci* 2020; 13: 4–10.
14. Zanjari N, Sharifian Sani M, Chavoshi MH, et al. Successful aging as a multidimensional concept: an integrative review. *Med J Islam Repub Iran* 2017; 31: 00.
15. Asiamah N. Social engagement and physical activity: commentary on why the activity and disengagement theories of ageing may both be valid. *Cogent Med* 2017; 4: 1289664.
16. Kooij DTAM, Zacher H, Wang M, et al. Successful aging at work: a process model to guide future research and practice. *Ind Organ Psychol* 2020; 13: 345–365.
17. Emlet CA, Harris L, Furlotte C, et al. 'I'm happy in my life now, I'm a positive person': approaches to successful ageing in older adults living with HIV in Ontario, Canada. *Ageing Soc* 2017; 37: 2128–2151.
18. Rowe JW and Kahn RL. Successful aging. *Gerontologist* 1997; 37: 433–440.
19. Rowe JW and Kahn RL. Successful aging and disease prevention. *Adv Ren Replace Ther* 2000; 7: 70–77.
20. Crowther MR, Parker MW, Achenbaum WA, et al. Rowe and Kahn's model of successful aging revisited: positive spirituality—the forgotten factor. *Gerontologist* 2002; 42: 613–620.
21. Vaillant GE and Mukamal K. Successful aging. *Am J Psychiatry* 2001; 158: 839–847.
22. Depp CA and Jeste DV. Definitions and predictors of successful aging: a comprehensive review of larger quantitative studies. *Am J Geriatr Psychiatry* 2006; 14: 6–20.
23. Torregrosa-Ruiz M, Gutiérrez M, Alberola S, et al. A successful aging model based on personal resources, self-care, and life satisfaction. *J Psychol* 2021; 155: 606–623.
24. Pedras S, Carvalho R and Pereira MG. A predictive model of anxiety and depression symptoms after a lower limb amputation. *Disability Health J* 2018; 11: 79–85.
25. Zeng F, Deng G, Cui Y, et al. A predictive model for the severity of COVID-19 in elderly patients. *Aging Sci Res* 2020; 12: 20982.
26. DuMontier C, Clough-Gorr KM, Silliman RA, et al. Health-related quality of life in a predictive model for mortality in older breast cancer survivors. *J Am Geriatr Soc* 2018; 66: 1115–1122.
27. Byeon H. Development of a physical impairment prediction model for Korean elderly people using synthetic minority over-sampling technique and XGBoost. *Int J Adv Comput Sci Appl* 2021; 12: 36–41.
28. Cai T, Long J, Kuang J, et al. Applying machine learning methods to develop a successful aging maintenance prediction model based on physical fitness tests. *Geriatr Gerontol Int* 2020; 20: 637–642.
29. Romo RD, Wallhagen MI, Yourman L, et al. Perceptions of successful aging among diverse elders with late-life disability. *Gerontologist* 2013; 53: 939–949.
30. Al-Hashem MA, Alqudah AM and Qananwah Q. Performance evaluation of different machine learning classification algorithms for disease diagnosis. *Int J E-Health Med Commun* 2021; 12: 1–28.
31. Alqudah AM, Al-Hashem M and Alqudah A. Reduced number of parameters for predicting post-stroke activities of daily living using machine learning algorithms on initiating rehabilitation. *Informatica* 2021; 45.
32. Alqudah AM. Ovarian cancer classification using serum proteomic profiling and wavelet features a comparison of machine learning and features selection algorithms. *J Clin Eng* 2019; 44: 165–173.
33. Alquran H, Alqudah AM, Abu-Qasmieh I, et al. ECG classification using higher order spectral estimation and deep learning techniques. *Neural Netw World* 2019; 29: 207–219.
34. Alqudah AM, Qazan S and Obeidat YM. Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds. *Soft COmput* 2022; 26: 13405–13429.
35. Alqudah AM, Alquraan H, Qasmieh IA, et al. Employing image processing techniques and artificial intelligence for automated eye diagnosis using digital eye fundus images. *J Biomimetics, Biomater Biomed Eng* 2018; 39. 10.4028/www.scientific.net/JBBBE.39.40
36. Alsharif R, Al-Issa Y, Alqudah AM, et al. Pneumonianet: automated detection and classification of pediatric pneumonia using chest X-ray images and CNN approach. *Electronics (Basel)* 2021; 10: 2949.
37. Qasmieh IA, Alquran H and Alqudah AM. Occluded iris classification and segmentation using self-customized artificial intelligence models and iterative randomized Hough transform. *Int J Electr Comput Eng* 2021; 11: 4037.
38. Obeidat Y and Alqudah AM. A hybrid lightweight 1D CNN-LSTM architecture for automated ECG beat-wise classification. *Trait Signal* 2021; 38. doi:10.18280/ts.380503
39. Alqudah AM, Alquran H and Qasmieh IA. Classification of heart sound short records using bispectrum analysis approach images and deep learning. *Network Model Anal Health Inf Bioinf* 2020; 9: 1–16.
40. Masad IS, Alqudah A, Alqudah AM, et al. A hybrid deep learning approach towards building an intelligent system for pneumonia detection in chest X-ray images. *Int J Electr Comput Eng* 2021; 11: 5530–5540.
41. Al-Issa Y, Alqudah AM, Alquran H, et al. Pulmonary diseases decision support system using deep learning approach. *Comput Mater Contin* 2022; 73: 311–326.
42. Alqudah A, Alqudah AM, Alquran H, et al. Recognition of handwritten Arabic and Hindi numerals using convolutional neural networks. *Appl Sci* 2021; 11: 1573.
43. Alqudah AM, Qazan S, Al-Ebbini L, et al. ECG Heartbeat arrhythmias classification: a comparison study between different types of spectrum representation and convolutional neural networks architectures. *J Ambient Intell Humanized Comput* 2021; 13: 1–31.

44. Ahmadi M and Nopour R. Clinical decision support system for quality of life among the elderly: an approach using artificial neural network. *BMC Med Inf Decis Making* 2022; 22: 293.
45. Tagharrobi Z, Sharifi K and Sooky Z. Psychometric evaluation of Shah version of modified Barthel index in elderly people residing in Kashan Golabchi nursing home. *KAUMS J* 2011; 15: 213–224. [In Persian].
46. Bayani AA, Koocheky AM and Goodarzi H. The reliability and validity of the satisfaction with life scale. *Dev Psychol* 2007; 3: 259–265.
47. Bandari R, Mohammadi Shahboulaghi F and Montazeri A. Development and psychometric evaluation of the healthy lifestyle questionnaire for elderly (heal). *Health Qual Life Outcomes* 2020; 18: –9.
48. Chandrashekar G and Sahin F. A survey on feature selection methods. *Comput Electr Eng* 2014; 40: 16–28.
49. Miao J and Niu L. A survey on feature selection. *Procedia Comput Sci* 2016; 91: 919–926.
50. Li J, Cheng K, Wang S, et al. Feature selection: a data perspective. *ACM Comput Surv* 2017; 50: 1–45.
51. Subasi A, Alickovic E and Kevric J. Diagnosis of chronic kidney disease by using random forest. CMBEBIH 2017: Springer; 2017. p.589-94.
52. Speiser JL, Miller ME, Tooze J, et al. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst Appl* 2019; 134: 93–101.
53. Belgiu M and Drăguţ L. Random forest in remote sensing: a review of applications and future directions. *ISPRS J Photogramm Remote Sens* 2016; 114: 24–31.
54. Tang F and Ishwaran H. Random forest missing data algorithms. *Stat Anal Data Mining: ASA Data Sci J* 2017; 10: 363–377.
55. Chengsheng T, Huacheng L and Bing X. AdaBoost typical Algorithm and its application research. In: MATEC Web of Conferences, Poland. Cracow University of Technology: EDP Sciences 2017.
56. Schapire RE. Explaining AdaBoost. In: Schölkopf B, Luo Z and Vovk V (eds) *Empirical Inference: Festschrift in Honor of Vladimir N Vapnik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp.37–52.
57. Priya MB, Juliet PL and Tamilselvi P. Performance analysis of liver disease prediction using machine learning algorithms. *Int Res J Eng Technol* 2018; 5: 206–211.
58. Iliyas MMK, Shaikh MIS and Student M. Prediction of heart disease using decision tree. *Allana Inst Manage Sci, Pune* 2019; 9: 1–5.
59. Abdar M, Kalhori SRN, Sutikno T, et al. Data mining on the heart disease with the use of different algorithms. *Int J Electr Comput Eng* 2015; 5.
60. Shanmuganathan S. Artificial neural network modelling: an introduction. In: Shanmuganathan S and Samarasinghe S (eds) *Artificial neural network modelling*. Cham: Springer International Publishing, 2016, pp.1–14.
61. Wu Y-c and Feng J-w. Development and application of artificial neural network. *Wirel Pers Commun* 2018; 102: 1645–1656.
62. Kukreja H, Bharath N, Siddesh C, et al. An introduction to artificial neural network. *Int J Adv Res Innov Ideas Educ* 2016; 1: 27–30.
63. El_Jerjawi NS and Abu-Naser SS. Diabetes prediction using artificial neural network. *Int J Adv Sci Technol* 2018; 121: 54–64.
64. Al-Issa Y and Alqudah AM. A lightweight hybrid deep learning system for cardiac valvular disease classification. *Sci Rep* 2022; 12: 14297.
65. Pioroński S and Górecki T. Using gradient boosting trees to predict the costs of forwarding contracts. In: 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS). IEEE. 2022.
66. Afrash MR, Kazemi-Arpanahi H, Nopour R, et al. Proposing an intelligent monitoring system for early prediction of need for intubation among COVID-19 hospitalized patients. *J Environ Health Sustainable Dev* 2022; 7: 1698–1707.
67. Suthaharan S. Support vector machine. In: Suthaharan S (eds) *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*. Boston, MA: Springer US, 2016, pp.207–235.
68. Pisner DA and Schnyer DM. Chapter 6—support vector machine. In: Mechelli A and Vieira S (eds) *Machine learning*. United States: Academic Press, 2020, pp.101–121.
69. Zhang Z. Naïve Bayes classification in R. *Ann Transl Med* 2016; 4: –5.
70. Alqudah A and Alqudah AM. Artificial intelligence hybrid system for enhancing retinal diseases classification using automated deep features extracted from OCT images. *Int J Intell Syst Appl Eng* 2021; 9: 91–100.
71. Berrar D. Cross-validation. In: *Encyclopedia of bioinformatics and computational biology*. Amsterdam, The Netherlands: Elsevier, 2019.
72. Zhang W, Ye B, Liang W, et al. Preoperative prognostic nutritional index is a powerful predictor of prognosis in patients with stage III ovarian cancer. *Sci Rep* 2017; 7: 1–8.
73. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem* 2008; 54: 17–23.
74. Lee S-H, Choi I, Ahn W-Y, et al. Estimating quality of life with biomarkers among older Korean adults: a machine-learning approach. *Arch Gerontol Geriatr* 2020; 87: 103966.
75. Adnan K and Akbar R. An analytical study of information extraction from unstructured and multidimensional big data. *J Big Data* 2019; 6: 91.
76. Byeon H. Developing a model to predict the social activity participation of the senior citizens living in South Korea by combining artificial neural network and quest algorithm. *Int J Eng Technol* 2019; 8: 214–221.