POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)
URL: http://dergipark.org.tr/politeknik

Artificial intelligence in healthcare: fall risk assessment in older adults by using machine learning techniques

Sağlık hizmetinde yapay zeka: makine öğrenmesi teknikleri kullanılarak yaşlılarda düşme riski tespiti

***Yazar(lar) (Author(s))**: Gökçe ÖZDEN GÜRCAN[1], Hakan GÖKDAŞ[2], Ebru TURAN KIZILDOĞAN[3]*

ORCID[1]: 0000-0002-5386-1985

ORCID[2]: 0000-0003-3399-8658

ORCID[3]: 0000-0002-2005-604X

To cite to this article: Ozden Gurcan G., Gokdas H. and Turan Kızıldogan E., "Artificial intelligence in healthcare: fall risk assessment in older adults by using machine learning techniques", ***Journal of Polytechnic***, *(*): *, (*).

Bu makaleye şu şekilde atıfta bulunabilirsiniz: Ozden Gurcan G., Gokdas H. ve Turan Kızıldogan E., "Artificial intelligence in healthcare: fall risk assessment in older adults by using machine learning techniques", ***Politeknik Dergisi***, *(*): *, (*).

Erişim linki (To link to this article): http://dergipark.org.tr/politeknik/archive

DOI: 10.2339/politeknik.1446723

Artificial Intelligence in Healthcare: Fall Risk Assessment in Older Adults by Using Machine Learning Techniques

Highlights

- *Fall risk prediction in older adults is performed using machine learning techniques such as Random Forest, Decision Tree, and Adaptive Boosting.*
- *The prediction of falling in elderly individuals with less input will enable health professionals working in this field to gain an advantage and save time.*
- *This study facilitates data prediction as it does not require a professional employee or clinical test in the measurement and determination of selected inputs.*

Graphical Abstract

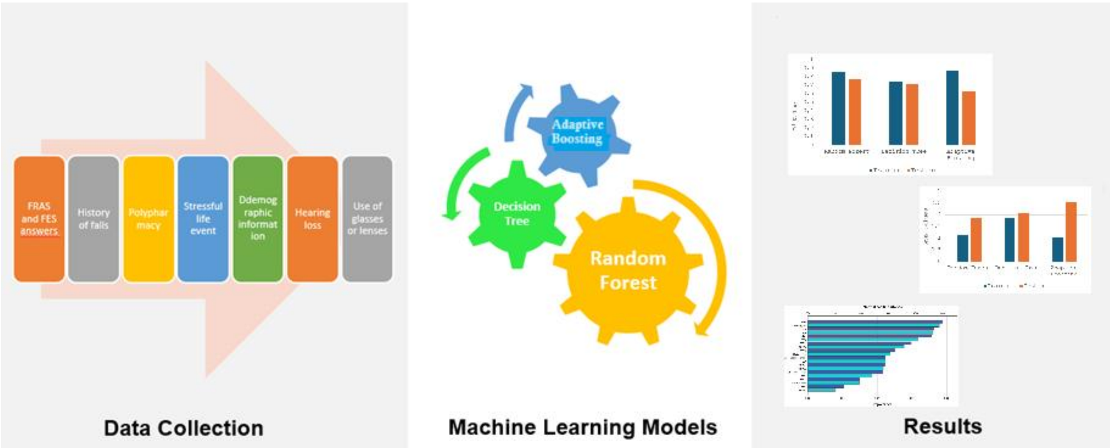

Figure. Graphical Abstract

Aim

This study aims to facilitate the fall risk assessment process for health professionals to determine the fall risk factors in elderly individuals and to make predictions.

Design & Methodology

In order to predict the risk of falling in the elderly, the Random Forest (RF), Adaptive Boosting (AB), and Decision Tree (DT) methods are used. The experimental and predicted fall risk values are compared in terms of test results. Besides physical and health factors of elderly people, FRAS and FES, questionnaire answers are used as input variables to predict BBS values for fall risk.

Originality

Random Forest, Adaptive Boosting, and Decision Tree are compared in terms of prediction efficiency in detecting fall risk using different input variables. Designing an artificial intelligence system that uses demographic characteristics as well as answers to survey questions such as FES, FRAS, and BBS to determine the risk of falling in elderly individuals will be a new contribution to the literature.

Findings

The R^2 (coefficient of determination) was 0.85 for training and 0.77 for testing the fall risk prediction of the RF model. The coefficient of determination is also obtained as 0.75 and 0.87 for the training of the DT and AB models, and 0.72 and 0.63 for testing. It is observed that machine learning methods, such as RF, DT, and AB, can be used in fall risk prediction.

Conclusion

The decision support systems that experts and elderly individuals can use in the coming periods can be designed by using the machine learning methods and the input and output structure indicated in the article. The various scales used in the clinic that evaluate the risk of balance and fall, besides the BBS, can also be used for output detection in future studies.

Declaration of Ethical Standards

The author(s) of this article declares that the materials and methods used in this study do not require ethical committee permission and/or legal or special permission.

Artificial Intelligence in Healthcare: Fall Risk Assessment in Older Adults by Using Machine Learning Techniques

Araştırma Makalesi / Research Article

Gökçe ÖZDEN-GÜRCAN[1*], Hakan GÖKDAŞ[2], Ebru TURAN-KIZILDOĞAN[3]

[1]Department of Machinery and Metal Technologies, Eskisehir Osmangazi University, Turkey
[2]Department of Industrial Engineering, Eskisehir Technical University, Turkey
[3]Vocational School of Health Services, Eskişehir Osmangazi University, Turkey

(Geliş/Received : 04.03.2024; Kabul/Accepted : 19.02.2025; Early View : 07.03.2025)

ABSTRACT

There are many attempts to provide the elderly with a more independent life. One of the main problems facing people in this age group is falls. Falls are one of the most common accidents among the elderly and may result in extended hospitalization and increased medical costs. The requirement for care services, such as fall detection, is increasing because of the growing population of elderly people. In this study, machine learning techniques: Adaptive Boosting, Random Forest, and Decision Trees, are used to predict fall risk of elderly people. Fall risk assessment methods are used to obtain inputs and outputs in addition to the physical and clinical features of people in the dataset.

This study aims to facilitate the fall risk assessment process of health professionals to determine the fall risk factors of elderly individuals, and to make predictions. Based on the results of fall prediction, individualized fall prevention interventions can be developed to reduce the fall rates of elderly individuals.

Keywords: Fall risk, elderly adults, machine learning, Decision Tree, Random Forest.

Sağlık Hizmetinde Yapay Zeka: Makine Öğrenmesi Teknikleri Kullanılarak Yaşlılarda Düşme Riski Tespiti

ÖZ

Yaşlılara daha bağımsız bir yaşam sağlamak için birçok girişimde bulunulmaktadır. Bu yaş grubundaki insanların karşılaştığı temel sorunlardan biri düşme olaylarıdır. Düşme, yaşlılar arasında en sık görülen kazalardan biridir ve hastanede kalış süresinin uzamasına, tıbbi maliyetlerin artmasına neden olabilmektedir. Yaşlı nüfusun artması nedeniyle, düşmenin tespitine yönelik bakım hizmetlerine olan ihtiyaç da artmaktadır. Bu çalışmada yaşlıların düşme riskinin tahmin edilmesi amacıyla makine öğrenmesi teknikleri (Adaptive Boosting, Random Forest, Decision Tree) kullanılmıştır. Veri setinde yer alan kişilerin fiziksel ve klinik özelliklerinin yanı sıra, girdi ve çıktıların elde edilmesi için düşme riski değerlendirme yöntemleri kullanılmıştır.

Bu çalışma, yaşlı bireylerin düşme riski faktörlerini belirlemek ve tahminlerde bulunmak amacıyla, sağlık profesyonellerinin düşme riski değerlendirme sürecini kolaylaştırmak amacıyla yapılmıştır. Bu sayede, düşme tahmininin sonuçlarına dayanarak yaşlı bireylerin düşme oranlarını azaltmak için bireyselleştirilmiş düşme önleme müdahaleleri geliştirilebilecektir.

Anahtar Kelimeler: Düşme riski, yaşlı yetişkinler, makine öğrenmesi, Karar Ağacı, Rastgele Orman Algoritması

1. INTRODUCTION

The population of people aged 65 and over is increasing worldwide. The World Health Organization (WHO) states that falls are among the most common health issues in old age. Falls are one of the important problems related to aging and are among the major causes of mortality and injuries in the elderly [1]. It is stated that approximately 30% of people aged 65 and over experience a fall at least once a year.

Our country is among the places where aging is rapid, similar to developing countries. According to Turkish Statistical Institute data, the total elderly population aged 65 and over was 6,895,385 (8.5%) [2]. Falls negatively affect individuals both physically and psychologically. Falls that reduce the quality of life not only create fear and anxiety in elderly individuals but also cause loss of independence [3].

Some comorbid conditions (e.g., postural hypotension, stroke, orthopedic diseases, visual impairment, anemia), female sex, surgery, older age, a history of falls, muscle weakness, impaired mobility, and polypharmacy are risk factors for falls [4]. WHO classifies risk factors into four categories: biological, behavioral, environmental, and socioeconomic [5]. The following are recognized behavioral risk factors: polypharmacy, fearful behavior, and lack of physical activity. Decreases in physical ability, balance issues, as well as problems with vision, hearing, and cognitive loss, are examples of biological

*Sorumlu Yazar (Corresponding Author)
e-mail : gozden@ogu.edu.tr

risk factors [6]. Environmental-external factors are caused by unfavorable physical conditions around the individual. Environmental-external factors include inadequate lighting, lack of handrails on the stairs inside and outside the house, high stair steps, lack of grab bars in the toilet and bathroom, slippery bathtubs, low toilets, unstable carpets, and slippery floors.

It is known that two-thirds of falls in the elderly can be prevented. It has been stated that identifying and predicting risk factors is important to prevent falls. Developing an extremely precise fall prediction model could help lower the rate of patient falls, which would reduce patients' injury and unnecessary medical expenses [7].

Rafiq et al. [8] found that the presence of caregivers, dizziness, anti-inflammatory drug use, diabetes, low foot sensation, heart failure, excessive alcohol consumption, coronary artery disease, and low BMI were not particularly related to the risk of falls in older people.

Although balance assessment systems are an important tool to differentiate balance disorders, they are tiring and time-consuming for clinicians. Machine learning can be a useful and practical tool for clinicians to predict the risk of falls in older adults [9]. It is important to determine the risk factors that cause falls in elderly individuals and to develop protective strategies to prevent falls. Early detection of elderly people at high risk of falls is necessary for the development of fall prevention programs. In clinical routine, healthcare practitioners must identify older people who are at higher risk of falling by using a simple and efficient clinical method [10].

In recent years, clinicians have preferred shorter, cost-efficient, and more practical applications. Recently, machine learning-based techniques have become popular in applications for prediction and diagnosis [11]. These strategies are also applied in fall risk prediction for the system to learn from past events during the prediction phase. The studies carried out in the field of fall risk prediction were recently examined and discussed in terms of the machine learning method used and input and output variables (Table 1).

Table 1. The studies using machine learning techniques for fall risk prediction

Authors	Method	Inputs	Outputs
Silva et al., 2024 [12]	Multilayer Perceptron (MLP), Multiple Linear Regression (MLR), Random Forest (RF), Random Tree (RT), K-nearest Neighbors (KNN), and Least-Squares Support Sector Regression (LS-SVR)	EMG signals and dynamometer data	BBS, fall risk
Wang et al., 2023 [13]	Ensemble Classification Model (ECM), Linear, Naive Bayes Classifier (NBC), Binary Decision Classification Tree, Discriminant Analysis Classifier (DAC), KNN, Support Vector Machine classifier (SVM)	States wing time, and step time, gait speed, trunk angle, step length, gait duration, Center of Mass (COM), toe clearance	Fall risk
Chen, Lingxiao et al., 2023 [14]	Light Gradient Boosting Machine (Light GBM), Adaptive Boosting (AB), LR, SVM, RF	Demographic factors, health status factors, lifestyle factors, medication factors, psychological factors, home environment factors, physical functions, blood indices	Occurrence of falls, occurrence of fall-related injuries
Chen et al., 2023 [15]	Logistic Regression	Demographic, health status, medication, lifestyle, psychological factors, socio economic factors	Fall risk
Yongjian et al., 2023 [16]	Gradient Boosting and Ridge Regression	Demographic characteristics, socioeconomic status, and self-reported physical mental health, health behaviors, social capital, and community environment	Functional disabilities
Sharma et al., 2023 [17]	Tree-based and linear ML algorithms (eg, XGBoost, CatBoost, logistic regression)	Pharmaceutical information network, population, and vitality statistics data and hospitalizations emergency department visits, physician visits/claims	Risk of fall
Langsetmo et al., 2023 [18]	RF and the fine-gray model	History of fracture after 50 years of age, low physical activity, shrinking, age, self-reported race/ethnicity, height, weight, health status, smoking status, walking speed, weakness, poor energy, recalled height and weight at 25 years of age, and medication, dual-energy X-ray absorptiometry	5-year risk of competing mortality, 5-year risk of hip fracture
Ikeda et al., 2022 [19]	RF-based Boruta algorithm and the eXtreme Gradient Boosting algorithm	History of falls during the past year, self-rated health, age, fear of falling, ability to stand up from chairs, depressive symptoms, choking, dry mouth, arthritis, difficulty in eating tough foods, ability to climb stairs, sense of coherence, incontinence, and number of remaining teeth	Index of Competence, Japanese Geriatric Depression Scale, sense of coherence scale- fall risk

Table 1. (Cont.) The studies using machine learning techniques for fall risk prediction

Authors	Method	Inputs	Outputs
Lathouwers et al., 2022 [20]	Random Forest Classifier	Age, physical activity, gender, home ownership, housing issues, physical vulnerability, social vulnerability, loneliness, physical exertion, mental activity, environmental vulnerability, home type income, level of education, mental activity, insecurity, psychological vulnerability, civil status, surrounding density, feeling unsafe	Fall risk
Mishra et al., 2022 [21]	Logistic Regression, Decision Tree (DT), linear SVM, and RF, Shapley Additive Explanations, SVM, KNN	ADL, IADL, MMSE, GDS, SF12, fall history, age, gender, gait speed, FAP	Fall outcome (next 6 months)
Gökler et al., 2022 [22]	Adaptive Neuro-Fuzzy Inference System (ANFIS)	Risk values and risk classes	Evaluation of spatial risks in nursing homes
Makino et al., 2021 [23]	Decision-Tree Algorithm	Age, sex, prescribed medication, lower limb pain, gait speed, and fall history, knee osteoarthritis	Fall detection
Yoo & Oh, 2018 [24]	Artificial Neural Networks (ANNs)	Acceleration sensor data	Fall detection
Aicha et al., 2018 [25]	Long Short-Term Memory (LSTM), and a hybrid of the two methods (ConvLSTM), Convolutional Neural Network (CNN),	Accelerometer data	Fall risk
Razmara et al., 2018 [9]	ANNs	Psychological factors and public factors	Fall risk
Deschamps et al., 2016 [26]	Decision-Tree Algorithm	Gender, taking medications, functional autonomy, impaired cognition, postural sway, physical lifestyle, anthropometric measures, and various systemic domains.	Risk of a first fall (next year)
Vidigal et al., 2015 [27]	ANNs	Acceleration signals	Elderly falls detection

Silva et al. [12] used Random Tree (RT), Multilayer Perceptron (MLP), Random Forest (RF), K-Nearest Neighbors (KNN), Least-Squares Support Vector Regression (LS-SVR), Multiple Linear Regression (MLR), to predict the Berg Balance Scale Score (BBS) in elderly people. The pair of MLP and RF-based prediction models that had 10 features extracted from the EMG signals was the best. Chen et al. [14] developed prediction models for fall-related injuries and falls for older Chinese individuals. The fall and fall-related injury risk models for 3 years were developed using five machine learning algorithms. The best performance among the fall-related and fall injury prediction models was achieved by the Logistic Regression model. House temperature, flush toilets and sex were significant variables only related to the fall model, while smoking, lung function, and Internet access were only connected to the fall-related injury model. Yongjian et al. [16] predicted the functional disability of older people using machine learning techniques. The models that predicted functional disability the best were Gradient Boosting and Ridge Regression. Age, self-rated health, fall and posture stabilization factors, and Parkinson's and dementia diagnoses were significant factors in both models. Makino et al. [23] developed a Decision-Tree (DT) algorithm for fall prediction. Age, fall history, prescribed medication, fear of falling, sex, knee osteoarthritis, gait speed, lower limb pain, and timed up and go tests were used as input variables for fall prediction. Their findings offered helpful information for early fall risk screening and promotion of prevention methods. Razmara et al. [9] used Artificial Neural Networks to predict the fall risk in the elderly based on their physiological profile. The proposed model achieved effective results on the basis of people's physiological profiles, according to the experimental outputs. Nait Aicha et al. [25] presented models to predict fall risk in elderly people. Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and a hybrid of the two methods (ConvLSTM) were compared in terms of their performance. They found that wearable sensor data was a useful tool for assessing fall risk through deep learning models, especially multi-task learning. Deschamps et al. [26] developed a model based on DT algorithm to predict the risk of first fall onset in one year. Their study presented a prototype tool that gerontologists might readily use to improve their assessment of the risk of first fall onset and rank the most successful preventive techniques.

In the literature, it is observed that demographic factors, health status, psychological factors, home environment factors, and physical functions are commonly used as input variables in the prediction models. In this study, besides the physical and health factors of elderly people, Falls Risk Assessment Score (FRAS), and Falls Efficacy Scale (FES) questionnaire answers were used as input variables to predict Berg Balance Scale (BBS) values for fall risk. The aim of this study is to compare three machine learning methods (Adaptive Boosting, Random Forest, and Decision Tree) in terms of prediction efficiency for fall risk detection using different input variables. The design of an artificial intelligence system that uses demographic characteristics as well as answers to survey questions such as FES, FRAS, and BBS to determine the risk of falling in elderly individuals will be a new contribution to the literature.

The following sections present the remainder of this study. The data analysis is given in the second section. The methodology and machine learning techniques are presented in the third section. The fourth section consists of the results and discussion. The final section concludes with a summary and suggestions for further research.

2. DATA ANALYSIS FOR FALL RISK ASSESMENT

In this study, the experimental data of Menezes et al. [28] were used. The BBS values and fall history (https://data.mendeley.com/datasets/3d4vr4dwjs/3) were evaluated through clinical trials. Even though FRAS, FES, and FRAT-up (Fall Risk Assessment Tool) presented similar accuracy to that reported in the literature [28-30], BBS, stands out from the others. Therefore, BBS values are used as outputs of the prediction model for this study.

2.1. Fall Risk Assessment in Older Adults

Fall risk assessment is a routine practice in healthcare for the elderly, and special tests are used to determine the fall risk. The Berg Balance Scale (BSS) and the Timed Up and Go Test (TUGT) are two of the tools that are most frequently used in clinical practice [12]. Although there are clinical tests that are scientifically accepted and implemented, these analysis tools are subjective and often take a long time to implement; therefore, automation of this process can benefit healthcare professionals.

BBS values are used in the prediction model with values ranging from 44 to 56 and a mean value of 54.29. The inputs for the prediction models were selected based on the effect of the input variables on the output variable. The most effective inputs were used for predicting BBS values (Table 2).

Table 2. Input variables and their levels

Input Variables	Levels
Age	68 - 80
Height (cm)	149 - 180
Mass (kg)	47 - 93
Hearing loss	0 1
Use of glasses/ lenses	0 1
History falls	0 - 5
Polypharmacy	0 1
FRAS 1	0 1
FRAS 2	0 1
FRAS 3	0 1
FRAS 4	0 1
FES 4	1 2 3
FES 6	1 2 3

Table 2. (Cont.) Input variables and their levels

FES 7	1 2 3 4
FES 8	1 2 3 4
FES 11	1 2 3 4
FES 13	1 2 3 4
FES 14	1 2 3 4
FES 15	1 2 3 4
Stressful life event	0 1 2 3 4

Table 2 shows some physical and health factors along with the FES and FRAS questionnaire answers. In this study, fifty-two individuals (85% female) were chosen to participate, and all of them completed the 6-month follow-up. The participants' ages were between 69 and 80 and the mean was 74. Urinary incontinence (n = 21, 40%), hypertension (n = 29, 56%), and hearing loss (n = 20, 38%) were the most prevalent disorders. They also commonly reported having experienced at least one stressful situation in the previous 12 months (52%, n = 27) and having poor eyesight (90%, n = 47) requiring the use of glasses (69%, n = 36). The participants at the baseline investigation reported having fallen at least once in their lives, with 48 participants (92%), whereas 12 (23%) and 15 (29%) said they had fallen in the previous 6 or 12 months, respectively [28].

2.2. Selection of Input Variables

In this study, demographic information (age, body height, and mass), the stressful life event, history of falls, polypharmacy, hearing loss, use of glasses or lenses, FRAS (1, 2, 3, 4), and FES (4, 6, 7, 8, 11, 13, 14, 15) answers were selected as inputs. First, all demographic, illness-related information, and some FRAS, FES questions (totally 25 features) were considered. Their importance levels were evaluated by using the SPSS software. Finally, the features are selected whose importance score is over 20%. As a result, 20 features are used for the prediction models as inputs. The most effective inputs are used to predict outputs (BBS values). For this purpose, independent variable importance analysis was performed using SPSS software. Figure 1 shows the most valuable features ranked according to their importance score.

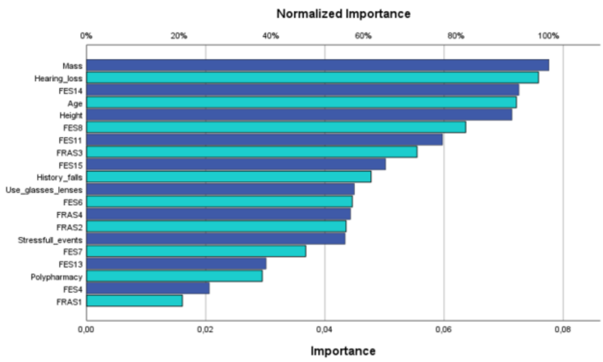

Figure 1. Input variables and importance ranking

In Figure 1, the importance scores for the output variables are shown: namely body mass (which emerged as the most effective parameter), hearing loss, FES 14, age, height, FES 8, FES 11, FRAS 3, FES 15, history of falls, use glasses or lenses, FES 6, FRAS 4, FRAS 2, stressful life events, FES 7, FES 13, polypharmacy, FES 4, and FRAS 1 were 100.0%, 97.7%, 93.5%, 93.0%, 92.0%, 82.0%, 77.0%, 71.5%, 64.7%, 61.5%, 57.9%, 57.5%, 57.1%, 56.1%, 55.9%, 47.4%, 38.8%, 38.0%, 26.5%, and 20.7%, respectively.

3. MATERIAL AND METHOD

In this study, the experimental data of Menezes et al. [28] were used. The BBS values and history of falls (https://data.mendeley.com/datasets/3d4vr4dwjs/3) were evaluated for the machine learning algorithms. The steps followed from data acquisition to the evaluation of performance criteria are given in Figure 2.

Figure 2. Steps of the modelling process

Figure 2 shows the modelling processes such as data acquisition, feature selection, data normalization, regression analysis with machine learning algorithms, and evaluation of performance criteria.

3.1. Methodology of Fall Detection

Medical practitioners and physiotherapists usually use several types of standardized and approved balance tests to determine the patient's fall risk. The Berg Balance Scale (BBS) is one such assessment. Medical practitioners frequently utilize the BBS, a conventional and validated assessment, to evaluate the risk of falls. The extensive BBS consists of 14 motor tasks with different levels of difficulty, such as turning, stepping, sitting, and standing up from a chair. The total of the 14 task scores was summed to determine the final BBS score. A five-level scale that ranges from 0 (unable) to 4 (independent), is used to score each job. Fall risk is categorized as high for scores between 0 and 20, medium for scores between 21 and 40, and low for scores between 41 and 56 [31]. In other words, between 0 and 20 points: that the person is dependent on a wheelchair and has a 100% fall risk. Between 21-40 points: that the person can walk with help because there is a fall risk. Between 41-56 points: it states that he can walk independently with a lower risk of falling. The fall risk score (BBS) between 0–41 is generally considered a high risk of falling. This score is considered for elderly patients who cannot walk without any help. The cut off value of BBS score were obtained as “45” for elderly who can live and walk independently in their home in the literature [32].

The data set [28] that we used in our study deals with the elderly individuals who can live independently in their homes, so the scores they received were observed between 45–56. Chiu et al. [33] revised the cut off value as “47” in their study. It means that for these people, over 47 is low risk and under 47 is high risk. As a result, our data set includes high and low risk BBS values for elderly people who can live and walk independently in their homes.

The Falls Efficacy Scale (FES) evaluates falling when people are going about their daily lives both indoors and in public. The questionnaire [34] assesses concern about the potential of falling while participating in 16 different activities on a scale from 1 to 4. The cutoff point to determine whether participants were at a low or high risk of falling was 23 points or more (sensitivity = 47%, specificity = 66%) [29].

FRAS is a five-question questionnaire [35] addressing clinical variables. Higher scores on the FRAS, which has a score range of 0 to 6.5, indicate a higher risk of falls. Each of the following categories received a score: more than one fall within the last 12 months (yes = 2); age (0.02 per a year increase from 60-year-old); loss of balance (yes = 1); poor vision (yes = 1); weak hand grip (yes = 1); slow walking speed in gait (yes = 1.5). Considering polypharmacy, the participants' risk of falling was categorized as either low (< 5 medications) or high (≥ 5 medications) based on the number of medications they were taking (specificity = 67, sensitivity = 49%).

The Stressful Life Events questionnaire was used to evaluate stressful life events that occurred within the year before the study's completion [36]. In elderly individuals, demographic changes such as age and body weight are selected as inputs to estimate the risk of falling because they are associated with loss of balance.

3.2. Machine Learning Techniques for the Prediction of Fall Risk

In this study, machine learning techniques such as Random Forest, Decision Tree, and Adaptive Boosting are used for the prediction model.

Random Forest

The Random Forest (RF) method is based on the values of random vectors sampled independently of each tree it contains, the pseudocode of the RF algorithm [37]. It is a combination of tree predictors with the same distribution for trees in the forest. Training and testing stages are used for RF design, similar to other supervised machine learning techniques. This algorithm performs the process by extracting predictions from the labeled training data to predict the label of new unlabeled input data. During the operations, a generalization error is obtained for forests. The generalization error of a tree classifier forest depends on the strength of each tree in the forest and the correlation between trees. The general result gets closer based on the generalization error as the number of trees in the forest increases [38, 39] . RF provides a random approach to the tree model when expanding trees. Using this method, when segmenting a tree node, the algorithm searches for the best feature within a random subset of features instead of looking for the most important feature in the tree. One of the most important advantages of RF is that it can provide solutions to both regression and classification problems, which form the basis of other machine learning algorithms [40].

Decision Tree Algorithm

The Decision Tree (DT) consists of three main parts called nodes, branches, and leaves. The first node that has no input is defined as the root node. Nodes whose outputs are inputs to another node are called internal nodes; nodes whose outputs are not inputs to another node are called leaf nodes. In the decision tree, each internal node is split into two or more parts. Decision tree algorithms generate a tree structure with a minimum error rate [41]. Determining the branch-splitting criteria in decision tree structures is of great importance to increase the success rate of the algorithm. Some approaches such as information gain, chi-square statistic, and GINI index are the preferred approaches in determining splitting criteria [42]. One of the approaches used to enhance the performance of the decision tree is the pruning method. The pruning method simplifies the tree structure and reduces complexity by eliminating sub-trees that have low statistical validity. Many DT algorithms such as C4.5, C5.0, ID3, and classification and regression trees (CART) have been developed since the automatic interaction detector (AID) algorithm [43]. The CART algorithm uses the GINI index approach as a splitting criterion and grows the decision tree by splitting without any stopping rule. After the completion of splitting, pruning from the leaf to the root is performed (e.g., CART algorithm [44]).

Adaptive Boosting

Adaptive Boosting (AB) is the boosting technique introduced by Freund and Schapire [45]. The most frequently used kind of boosting algorithm that improves several poor learners , into a single robust learner is the adaptive boosting technique. Adaptive Boosting (AdaBoost) can be used for both classification and regression problems.

In this study, AdaBoost.R2 which is one of the boosting algorithms for regression problems was used. The final prediction in AdaBoost.R2 is a weighted mean of the predictions made by each weak learner. The algorithm works by feeding the information from the previous weak learner to the next, improving the previous learner's error, thereby making a particular weak learner's performance dependent on the previous one [46]. Equal weighting factors are used to train the initial weak learner; these weighting coefficients will be adjusted in subsequent boosting rounds. The weights of the cases with poor predictions increase while those with good predictions decrease [47].

The prediction models developed in this study were performed using RF, DT, and AB algorithms with the Python programming language and the Scikit-Learn 1.3.1 library [44]. The maximum depth parameter is defined as 3 for both the Decision Tree and Random Forest, with the minimum sample leaf parameter defined as 4 for the Decision Tree and 2 for the Random Forest. The Adaptive Boosting algorithm uses the same parameters as the Decision Tree algorithm.

4. RESULTS AND DISCUSSION

Adaptive Boosting, Random Forest, and Decision Tree algorithms were examined to develop a fall risk model. The input values were normalized into a predetermined range. The input data in the current study were scaled using the max–min method to fit into the range [0,1]. The analysis was conducted by using the 5-fold cross validation technique for each algorithm. 5-fold cross-validation uses the complete dataset for part-to-part training and validation, instead of dividing the dataset at random. The data set is depicted for cross-validation in Figure 3.

Table 3. Performance criteria for Random Forest, Decision Tree and Adaptive Boosting models

Performance criteria	Random Forest		Decision Tree		Adaptive Boosting	
	Training	Testing	Training	Testing	Training	Testing
MAE	0.72	0.94	0.89	1.01	0.84	1.13
MAPE	1.38	1.80	1.69	1.91	1.56	2.14
MSE	1.16	1.89	1.88	2.09	1.01	2.55
RMSE	1.07	1.27	1.37	1.37	1.00	1.52
R^2	0.85	0.77	0.75	0.72	0.87	0.63

Figure 3. Using 5-fold cross-validation and data set splitting

There were fifty-two elderly people in the original sample. There are five equal sets to these fifty-two data points. This indicates that every part has ten data points, and the folding operation is carried out five times. 20% of the data from elderly people is utilized for validation in the first fold, while the remaining data is used for training. Likewise, the second subset is used for validation in the second fold. The remainder of the folding is done in this manner, as seen in Fig. 3. Thus, overfitting of the model is avoided by using cross-validation by folding the dataset. Since the training procedure is carried out using various training sets each time, the training results are more generalized and robust [48].

The performance criteria were evaluated, including mean absolute error (MAE), mean absolute percentage error (MAPE), mean square error (MSE), root mean square erros (RMSE) and coefficient of determination (R^2) (Table 3).

The Random Forest approach achieved a high coefficient of determination for training (0.85) and testing (0.77). MAPE and MSE values were found to be 1.38 and 1.16 for training and 1.80 and 1.89 for testing, respectively. The Decision Tree approach also achieved a high coefficient of determination for training (0.75) and testing (0.72). MAPE and MSE values were found to be 1.69 and 1.88 for training and 1.91 and 2.09 for testing, respectively. These values are also effective for precise prediction models. The Adaptive Boosting algorithm achieved a high coefficient of determination (0.87) for training; however, a low value for testing (0.63). MAPE and MSE values were 1.56 and 1.01 for training and 2.14 and 2.55 for testing, respectively. Coefficient of determination and MSE values are also depicted in Figure 4 and Figure 5.

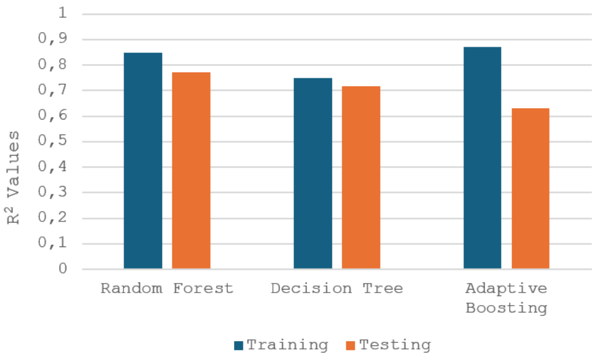

Figure 4. Coefficient of determination values for RF, DT and AB training and testing results

Figure 5. MSE values for RF, DT and AB training and testing results

Figure 4 shows that the RF and DT models' predicted fall risk values closely match the experimental results for testing. The R^2 values are obtained as 0.77 and 0.72 for testing the fall risk prediction of the RF and DT models, respectively. These values are higher than the results of the AB algorithms. The R^2 values are obtained 0.85 and 0.75 for training of RF and DT models. Similarly, the MSE values are obtained as 1.89 and 2.09 for testing the RF and DT models. These values are smaller than the results obtained using the AB algorithm (2.55) for testing (Figure 5). Although the results of the three algorithms are generally reasonable, RF and DT perform better based on the test results. BBS was the most reliable method for screening the risk of falling, due to the highest predictive accuracy among the FRA approaches, whereas the other methods presented limited screening capacity [30]. According to the results, it is observed that machine learning methods, primarily RF, DT, and then the AB method, can be used in fall risk prediction.

5. CONCLUSION

Determining risk of falling is of great importance in the health care of the elderly and the strategies necessary for the prevention of falls and the secondary problems related to them. In this study, alternatives for using machine learning techniques to evaluate the risk of falling and predict the score of the BBS were obtained, and the results were analyzed.

This study provides convenience for the data prediction, as it does not require a professional employee and clinical test in the measurement and determination of selected inputs. The prediction of falling among elderly individuals with less input will enable health professionals working in this field to gain a time advantage. The prediction of falls in the elderly and the implementation of measures to prevent them will serve to decrease the health issues that result from falls.

Using the machine learning methods and the input and output structure indicated in the article, decision support systems can be designed for use by experts and elderly individuals in the coming periods. In future studies, the various scales used in the clinic that evaluate the risk of falling, besides the BBS, can also be used for output detection . The other machine learning techniques that have good results in the literature can be used and compared with each other.

DECLARATION OF ETHICAL STANDARDS

The authors of this article declare that the materials and methods used in this study do not require ethical committee permission and/or special legal permission.

AUTHORS' CONTRIBUTIONS

Gökçe ÖZDEN-GÜRCAN: The development of a research idea, analysis of experimental data, implementation of ML algorithms, interpretation of findings, manuscript writing and general revision.

Hakan GÖKDAŞ: The development of a research idea, analysis of experimental data, implementation of ML algorithms, interpretation of findings, manuscript writing.

Ebru TURAN-KIZILDOĞAN: The development of a research idea**,** analysis of experimental data, interpretation of findings, manuscript writing.

CONFLICT OF INTEREST

There is no conflict of interest in this study.

REFERENCES

[1] Scheckel B., Stock S. and Müller D., "Cost-effectiveness of group-based exercise to prevent falls in elderly community-dwelling people", ***BMC geriatrics***, 21, 1-9, (2021).

[2] "İstatistiklerle Yaşlılar 2017", Türkiye İstatistik Kurumu (TUİK) 27587, (2018).

[3] Gale C.R., Cooper C. and Aihie Sayer A., "Prevalence and risk factors for falls in older men and women: The English Longitudinal Study of Ageing", ***Age and ageing***, 45(6), 789-794, (2016).

[4] Rodrigues A.R.G.d.M., Assef J.C. and Lima C.B.d., "Assessment of risk factors associated with falls among the elderly in a municipality in the state of Paraíba, Brazil. A cross-sectional study", ***Sao Paulo Medical Journal***, 137, 430-437, (2020).

[5] Organization W.H. and Unit L.C., "WHO global report on falls prevention in older age", ***World Health Organization***, (2008).

[6] Pérez-Ros P., Martínez-Arnau F.M., Orti-Lucas R.M. and Tarazona-Santabalbina F.J., "A predictive model of isolated and recurrent falls in functionally independent community-dwelling older adults", ***Brazilian journal of physical therapy***, 23(1), 19-26, (2019).

[7] Liu C.-H., Hu Y.-H. and Lin Y.-H., "A machine learning–based fall risk assessment model for inpatients", CIN: Computers, ***Informatics, Nursing***, 39(8), 450-459, (2021).

[8] Rafiq M., McGovern A., Jones S., Harris K., Tomson C., Gallagher H. and de Lusignan S., "Falls in the elderly were predicted opportunistically using a decision tree and systematically using a database-driven screening tool", ***Journal of clinical epidemiology***, 67(8), 877-886, (2014).

[9] Razmara J., Zaboli M.H. and Hassankhani H., "Elderly fall risk prediction based on a physiological profile approach using artificial neural networks", ***Health informatics journal***, 24(4), 410-418, (2018).

[10] Bongue B., Dupré C., Beauchet O., Rossat A., Fantino B. and Colvez A., "A screening tool with five risk factors was developed for fall-risk prediction in community-dwelling elderly", ***Journal of clinical epidemiology***, 64(10), 1152-1160, (2011).

[11] Millet A., Madrid A., Alonso-Webber J.M., Rodríguez-Mañas L. and Pérez-Rodríguez R., "Machine Learning techniques applied to the development of a fall risk index for older adults", ***IEEE Access***, (2023).

[12] Silva D.A., Branco N.F.L.C., de Andrade Mesquita L.S., Branco H.M.G.C. and de Alencar Barreto G., "Electromyography and dynamometry in the prediction of risk of falls in the elderly using machine learning tools", ***Biomedical Signal Processing and Control***, 88, 105635, (2024).

[13] Wang S., Nguyen T.K. and Bhatt T., "Trip-Related Fall Risk Prediction Based on Gait Pattern in Healthy Older Adults: A Machine-Learning Approach", ***Sensors***, 23(12), 5536, (2023).

[14] Chen X., He L., Shi K., Wu Y., Lin S. and Fang Y., "Interpretable machine learning for fall prediction among older adults in China", ***American journal of preventive medicine***, (2023).

[15] Chen X., Lin S., Zheng Y., He L. and Fang Y., "Long-term trajectories of depressive symptoms and machine learning techniques for fall prediction in older adults: Evidence from the China Health and Retirement Longitudinal Study (CHARLS)", ***Archives of Gerontology and Geriatrics***, 111, 105012, (2023).

[16] Yongjian L., Koryu S., Masato N., Hirokazu M., Katsunori K. and Naoki K., "Machine Learning–Based Prediction of Functional Disability: a Cohort Study of Japanese Older Adults in 2013–2019", ***Journal of General Internal Medicine***, 1-8, (2023).

[17] Sharma V., Kulkarni V., Joon T., Eurich D.T., Simpson S.H., Voaklander D., Wright B. and Samanani S., "Predicting falls-related admissions in older adults in Alberta, Canada: a machine-learning falls prevention tool developed using population administrative health data", ***BMJ open***, 13(8), e071321, (2023).

[18] Langsetmo L., Schousboe J.T., Taylor B.C., Cauley J.A., Fink H.A., Cawthon P.M., Kado D.M., Ensrud K.E. and Group O.F.i.M.R., "Advantages and disadvantages of random forest models for prediction of hip fracture risk versus mortality risk in the oldest old", ***JBMR plus***, 7(8), e10757, (2023).

[19] Ikeda T., Cooray U., Hariyama M., Aida J., Kondo K., Murakami M. and Osaka K., "An interpretable machine learning approach to predict fall risk among community-dwelling older adults: a three-year longitudinal study", ***Journal of General Internal Medicine***, 37(11), 2727-2735, (2022).

[20] Lathouwers E., Dillen A., Díaz M.A., Tassignon B., Verschueren J., Verté D., De Witte N. and De Pauw K., "Characterizing fall risk factors in Belgian older adults through machine learning: a data-driven approach", ***BMC public health***, 22(1), 2210, (2022).

[21] Mishra A.K., Skubic M., Despins L.A., Popescu M., Keller J., Rantz M., Abbott C., Enayati M., Shalini S. and Miller S., "Explainable fall risk prediction in older adults using gait and geriatric assessments", ***Frontiers in digital health***, 4, 869812, (2022).

[22] Gökler S.H., Yılmaz D., Ürük Z.F. and Boran S., "A new hybrid risk assessment method based on Fine-Kinney and ANFIS methods for evaluation spatial risks in nursing homes", **Heliyon**, 8(10), (2022).

[23] Makino K., Lee S., Bae S., Chiba I., Harada K., Katayama O., Tomida K., Morikawa M. and Shimada H., "Simplified decision-tree algorithm to predict falls for community-dwelling older adults", ***Journal of clinical medicine***, 10(21), 5184, (2021).

[24] Yoo S. and Oh D., "An artificial neural network–based fall detection", ***International Journal of Engineering Business Management***, 10, 1847979018787905, (2018).

[25] Nait Aicha A., Englebienne G., Van Schooten K.S., Pijnappels M. and Kröse B., "Deep learning to predict falls in older adults based on daily-life trunk accelerometry", **Sensors**, 18(5), 1654, (2018).

[26] Deschamps T., Le Goff C.G., Berrut G., Cornu C. and Mignardot J.-B., "A decision model to predict the risk of the first fall onset", ***Experimental gerontology***, 81, 51-55, (2016).

[27] Vidigal M., Lima M. and Neto A.D.A., "Elder falls detection based on artificial neural networks", ***Fourteenth Mexican International Conference on Artificial Intelligence (MICAI)***, 226-230, (2015).

[28] Menezes M., de Mello Meziat-Filho N.A., Araújo C.S., Lemos T. and Ferreira A.S., "Agreement and predictive power of six fall risk assessment methods in community-dwelling older adults", ***Archives of Gerontology and Geriatrics***, 87, 103975, (2020).

[29] Camargos F.F., Dias R.C., Dias J. and Freire M.T., "Adaptação transcultural e avaliação das propriedades psicométricas da Falls Efficacy Scale-International em idosos brasileiros (FES-I-BRASIL)", ***Brazilian journal of physical therapy***, 14, 237-243, (2010).

[30] Gates S., Smith L.A., Fisher J.D. and Lamb S.E., "Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults", ***Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews*** [Internet], (2008).

[31] Berg K., Wood-Dauphine S., Williams J. and Gayton D., "Measuring balance in the elderly: preliminary development of an instrument", ***Physiotherapy Canada***, 41(6), 304-311, (1989).

[32] Berg K., "Measuring balance in the elderly: Development and validation of an instrument", ***Canadian journal of public health***, 83, 7-11, (1992).

[33] Chiu A.Y., Au-Yeung S.S. and LO S.K., "A comparison of four functional tests in discriminating fallers from non-fallers in older people", ***Disability and rehabilitation***, 25(1), 45-50, (2003).

[34] Yardley L., Beyer N., Hauer K., Kempen G., Piot-Ziegler C. and Todd C., "Development and initial validation of the Falls Efficacy Scale-International (FES-I)", ***Age and ageing***, 34(6), 614-619, (2005).

[35] El Miedany Y., El Gaafary M., Toth M., Palmer D. and Ahmed I., "Falls risk assessment score (FRAS): time to rethink", ***Journal of clinical Gerontology and Geriatrics***, 2(1), 21-26, (2011).

[36] Lopes C.S., Faerstein E. and Chor D., "Stressful life events and common mental disorders: results of the Pro-Saude Study", ***Cadernos de saude publica***, 19, 1713-1720, (2003).

[37] Wild Ali A.B., "Prediction of employee turn over using random forest classifier with intensive optimized pca algorithm", ***Wireless Personal Communications***, 119(4), 3365-3382, (2021).

[38] Freund Y. and Schapire R.E., "Experiments with a new boosting algorithm", ***icml***, 148-156

[39] Breiman L., "Random forests", ***Machine learning***, 45, 5-32, (2001).

[40] Biau G. and Scornet E., "A random forest guided tour", ***Test***, 25, 197-227, (2016).

[41] Rokach L. and Maimon O., "Data mining and knowledge discovery handbook", ***Springer New York***, (2010).

[42] Kothari R. and Dong M., "Decision trees for classification: A review and some new results", ***Pattern recognition: from classical to modern approaches***, 169-184, (2001).

[43] Niuniu X. and Yuxun L., "Notice of Retraction: Review of decision trees", ***3rd international conference on computer science and information technology***, 105-109, (2010).

[44] Pedregosa F., "Scikit-learn: Machine learning in python Fabian", ***Journal of machine learning research***, 12, 2825, (2011).

[45] Freund Y. and Schapire R.E., "A decision-theoretic generalization of on-line learning and an application to boosting", ***Journal of computer and system sciences***, 55(1), 119-139, (1997).

[46] Drucker H., "Improving regressors using boosting techniques", ***Icml***, e115, (1997).

[47] Shrestha D.L. and Solomatine D.P., "Experiments with AdaBoost. RT, an improved boosting scheme for regression", ***Neural computation***, 18(7), 1678-1710, (2006).

[48] Sejuti Z.A., Islam M.S., “A hybrid CNN–KNN approach for identification of COVID-19 with 5-fold cross validation”, ***Sensors International***, 4 (2023).