Accidental Fall Prediction and Detection in Elderly Persons using Ensemble Techniques

Michael Osiako Afote
Federal University Lokoja
Lokoja, Nigeria

Fati Oiza Ochepa
Federal University Lokoja
Lokoja, Nigeria

ABSTRACT
Accidental falls in the elderly have gradually become a major health concern requiring reliable prediction and timely detection. This study has adopted the use of artificial intelligence ensemble learning techniques to assist in tackling this critical issue. Specifically, bagging techniques were deployed with Random Forest (RF), Logistic Regression (LR), Support Vector Classifiers (SVC), and Decision Tree (DT). The dataset used in the study comprised both physiological and environmental-related data that serve as indicators for falls. Results from the study yielded the best performance with the bagging technique applied on the Random Forest, Logistic Regression, and Support Vector Classifiers, which yielded an accuracy of 96%. The bagged Decision Tree model also performed significantly with an accuracy of 93%. The model was deployed using Flask, with the integration of SMS alerts and a dashboard notification feature. The deployed system demonstrates potential as a valuable tool in ensuring early fall detection in the elderly by reducing the risks of sustaining injuries, enhancing safety, and improving the overall well-being and quality of life of elderly individuals.

Keywords
Accidental Fall Prediction, Ensemble Learning, Elderly Safety, Machine Learning.

1. INTRODUCTION
Globally, falls account for the second leading cause of unintentional injury-related deaths [1]. With falls being the reason for millions of emergency hospital visits annually, particularly in elderly persons ranging from 65 years and above, this has made the elderly more vulnerable and susceptible to fall-related health challenges. These incidents often result in severe physical injuries such as sprains, fractures, and head trauma. Other effects also include reduced mobility, loss of independence, and psychological anguish. Furthermore, falls can hurt an individual's quality of life, independence, and overall health. Given the significant impact of falls on the elderly population, there has been growing interest in developing methods for predicting and detecting falls to help prevent them from occurring. Advances in sensor technology, artificial intelligence, and machine learning have enabled the development and implementation of innovative fall prediction and detection systems.

Ensemble methods are a machine learning technique that combines several base machine learning models to produce one optimal predictive model. With advancement leading to the exponential progression in computational power that allows large training ensemble learning in a relatively small time frame, its applications have grown increasingly. Some of the applications of ensemble classifiers include remote sensing, computer security, emotion recognition, fraud recognition, financial decision-making, and more. Ensemble methods act as a regularizing parameter on overfitting models, increasing the stability of the final model and drastically increasing the model's accuracy. Developing an accidental fall prediction and detection system based on an ensemble approach using health related parameters like blood pressure, heart rate, and sugar level has the potential to enhance the safety and well-being of elderly individuals and reduce the burden of falls on healthcare systems [2].

Falls may be associated with syncope; syncope occurs when there is not enough blood flowing to the brain. There are many causes for this, depending on the type of syncope. Various contiguous environmental hazards, such as carpets and rugs, may also contribute to falls. Research has shown that 72.8% of theses falls occur at home, especially in women, this represents 80.2% of all fall injury victims [3].

The high incidence of falls among the elderly population necessitates the development of robust systems to predict and detect falls in real-time. Existing methods often focus solely on prediction or detection, lacking a unified approach, sensor inaccuracies, and a lack of user-friendly interfaces. Therefore, the development of effective fall prediction and detection systems is crucial to prevent falls, reduce the risk of injury, and improve the safety and well-being of elderly individuals. In this study, an ensemble-based machine learning framework combining bagging with Random Forests, Linear Regression, Support Vector Classifiers and Decision Tree models was proposed to provide on-time interventions for elderly victims who suffer falls. This is done with an overall aim of improving well-being and optimization of healthcare resource usage.

2. RELATED WORK

2.1 Review of Relevant Approach to Accidental Fall Prediction and Detection
A variety of approaches have been deployed in the prediction and detection of accident-related falls in elderly individuals. Sensor-based approaches rank as the most frequently used approach, followed by machine-learning-based techniques and mobile health technologies.

Sensor-based techniques include the utilization of ambient and wearable sensors that collect relevant data that will be used to predict and detect falls. These wearable devices range from gyroscopes and accelerometers to pressure and proximity sensors.

Machine learning algorithms have the ability to make meaning out of patterns that can indicate falls by analyzing the data from these sensors [4]. Data from non-sensor sources such as electronic health records and social media have been analyzed using machine learning techniques.

These approaches serve to offer additional information that may be of significance to fall prediction, in ways where it can guide towards decisions such as changes in medication regimens or social support networks [5]. Fall can also be detected via mobile health technologies like smartphones and smartwatches integrated sensors. Explorations have also been made into the viability of virtual and augmented reality technologies in preventing falls and offering rehab ideas. In essence, the field of fall prediction and detection in elderly people is evolving rapidly, presenting significant potentials towards safety improvement and quality of life [6].

2.2 Review of Related Literature

Krishnan [7] presented a fall detection system using the YOLO v8 object detection model which achieved an impressive accuracy of 90%. The study addressed fall in the elderly by proposing a non-intrusive approach thereby eliminating user compliance issues. However, future works suggested in the study includes privacy concerns and validation. [8] developed the FallCNN, a fall detection system for elderly individuals using deep CNN architecture to analyze data from accelerometer and gyroscope sensor and transform it into signal-based images. The model recorded a classification accuracy of 98% highlighting the potential of deep learning adoption for efficient fall detection. The research undertaken by [9] resulted in the deployment of a wearable device that uses an accelerometer and gyroscope for fall detection in elderly individuals. The study recorded an accuracy of 93.6%, after testing on a group.

A new fall detection system based on accelerometer for elderly people, was the output of the study undertaken by [10]. The system was built to respects privacy and comfort by examining 7,700 features from three datasets and using more sophisticated approaches to extract important information. [11] developed an edge device that uses IOMT to predict or detect falls based on image orientation and physiological sensors. The Good-Eye system, which includes a remote offline camera, achieved 95% accuracy in fall prediction and detection. An IoT-based system for fall detection using accelerometers and gyroscope sensors was the focus of the study by [12]. The system achieved a high accuracy of 98.5% in both fall and non-fall activities from remotely monitoring and analyzing data.

Despite advancements recorded in previous studies, there are still a number of limitations that surrounds existing fall detection systems. In most studies, the focus has been on either predicting falls or detecting falls, rather than integrating both functionalities in the systems. Other significant issues include small sample sizes and limited diversity in the data which has affected the generalization of results to larger populations. Many systems require high computational resources for real-world deployment while false positives and negative results impact the reliability of the systems hence, affecting user trust and system adoption [13]. The present study seeks to address these limitations by introducing an integrated framework that handles detection and prediction using ensemble learning techniques, making use of a comprehensive dataset to improve model training and evaluation, utilizing advanced ensemble methods yielding a real-time system designed that is suitable for deployment in practical environments.

3. METHODOLOGY

This study adopts the Cross Industry Standard Process for Data Mining (CRISP-DM), an open standard process model that describes common approaches used in building the fall prediction and detection models.

3.1 Architectural Design

The architectural design offers a high-level view of the structure and interconnections among the various components involved in building the model. Figure 1 depicts the architectural design of the proposed model, highlighting the data processing, features, modeling, and output components.

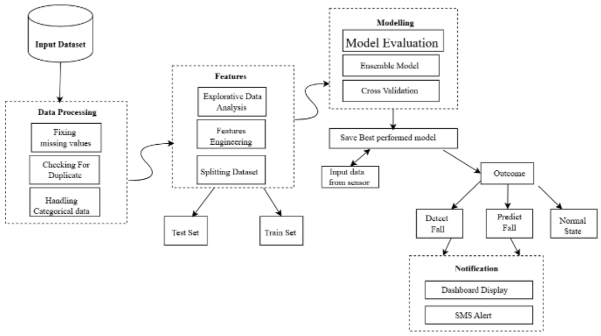


Fig 1: Architectural design for fall prediction and detection

3.2 Logical Design

The logical design in the fall detection system depicted in Figure 2 depicts the conceptual representation of interactions between the input data (Distance, Heart Rate, Blood Pressure, Accelerometer), data processing algorithms (machine learning models and ensemble models), and output predictions (all detect and predict).

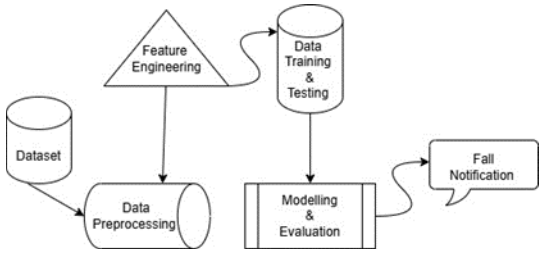


Fig 2: Logical design for fall prediction and detection

3.3 Machine Learning Algorithm

Machine learning has the ability to learn from data without being explicitly programmed to do so, such as in the dataset used in this research, which includes both medical and sensor data. Below are the machine learning algorithms used for detecting and predicting falls:

Decision Tree, a non-parametric supervised learning algorithm, serves as a tool for fall prediction and detection. Its hierarchical tree structure consists of a root node, branches, internal nodes, and leaf nodes. This helps to make it learn from data more efficiently.

$$E(S) = \sum_{i=1}^{c} - p_1 \log_2 p_1 \qquad (1)$$

where p is the i-th order probability,

$$G(S, C) = E(S) - \sum_{w2values(C)} \frac{Sw}{S} E(S_w)$$

Equation 1 calculates the information gain (G) when splitting a dataset S based on a categorical attribute, C.

Logistic Regression calculates or predicts the probability of a binary outcome (such as a fall occurring or not occurring). It is mathematically represented in Equation 2.

$$P(Y = 1/X) = \frac{1}{1+e^{-(\omega.x+b)}} \qquad (2)$$

P(Y=1|X): Probability of class 1 given data X
w: weight vector. x: input vector. b: bias term.

Support Vector Machine (SVM) creates the best line or decision boundary that can segregate n-dimensional space into classes so that the new data point can be placed easily in the correct category in the future. Equation 3 mathematically represents the SVM as follows:

$$f(x) = sign(w.x + b) \qquad (3)$$

Where w is the weight vector and b is the bias.

Random Forest is used to combine more than one decision tree algorithm for better accuracy. Where B is the number of trees and Tb(x) is the prediction of the b-th tree. It is represented mathematically in equation 4.

$$f(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \qquad (4)$$

Bagging, also called Bootstrap Aggregation, is an ensemble learning method that helps improve the performance and accuracy of the fall prediction and detection model. Boosting converts multiple weak learners of the dataset into a single strong learning model to improve machine model predictive accuracy and performance. A voting classifier uses an ensemble of models to predict an output class based on their highest probability of choosing it. It aggregates the findings of each classifier, predicting the output class based on the combined majority of voting.

3.4 Performance Evaluation

Accuracy measures the number of correct predictions made by a model from the total number of predictions made [14]. It is computationally represented in Equation 5.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \qquad (5)$$

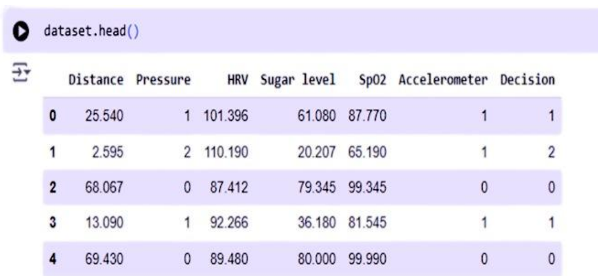
Confusion Matrix presents a table layout of the different outcomes of the prediction. Table 1 highlights the possible outcomes of a decision as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Table 1. The Confusion Matrix

	PREDICTED POSITIVE	**PREDICTED NEGATIVE**
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	True Negative (TN)	False Positive (FP)

3.5 Dataset

The dataset consists of 2040 rows and 7 columns. The features in this dataset are Distance, Blood Pressure, Heart Rate Variability (HRV), Sugar Level, Accelerometer, Oxygen Saturation (SpO_2) levels, as shown in Figure 3.



```
dataset.head()
```

	Distance	Pressure	HRV	Sugar level	SpO2	Accelerometer	Decision
0	25.540	1	101.396	61.080	87.770	1	1
1	2.595	2	110.190	20.207	65.190	1	2
2	68.067	0	87.412	79.345	99.345	0	0
3	13.090	1	92.266	36.180	81.545	1	1
4	69.430	0	89.480	80.000	99.990	0	0

Fig 3: Sample of the dataset

4. RESULTS AND DISCUSSION

This section explains the results obtained from the research carried out on fall detection and prediction in the development process. The development process includes the preprocessing steps, modeling, evaluation, and deployments.

4.1 Data Preprocessing and Exploratory Data Analysis

The dataset was visualized, and it was observed that the three expected outcomes from the dataset (0: no fall predicted or detected, 1: fall predicted, and 2: fall detected) were equally distributed, which could lead to overfitting. To address this, some rows were removed, resulting in a cleaned and pre-processed dataset with 1,409 rows and 7 columns.

The correlation of the selected features was visualized using a heatmap in order to identify the most critical features that can impact the model's performance. Figure 4 shows a heatmap that indicates a strong connection between Pressure, HRV, Accelerometer, and Distance with the decision variable. This means they play an important role in predicting and detecting falls. Notably, Pressure and HRV (0.92 and 0.83, respectively) demonstrate a high positive correlation with the outcome, indicating their potential as key predictive features. Similarly, Accelerometer (0.83) and Distance (-0.87) also show strong associations with the fall outcome, reinforcing their importance in fall detection models. In contrast, sugar level exhibits weaker correlations with other variables, suggesting a limited direct impact on fall prediction. These results informed the decision to prioritize the highly correlated features in the proposed model to enhance predictive accuracy.

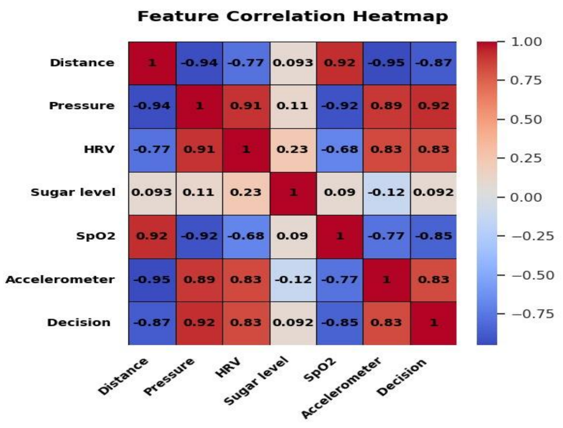


Fig 4: Heatmap visualization of features

4.2 Presentation of Results

This section discusses the results obtained by the algorithms that were trained on the dataset.

4.2.1 Model Performance Evaluation

Decision Tree, Logistic Regression, Random Forest and Support Vector Machine models were trained and evaluated. The accuracy of the four models is depicted in Figure 5 with

SVM and Linear Regression having the highest accuracy score of 96%.

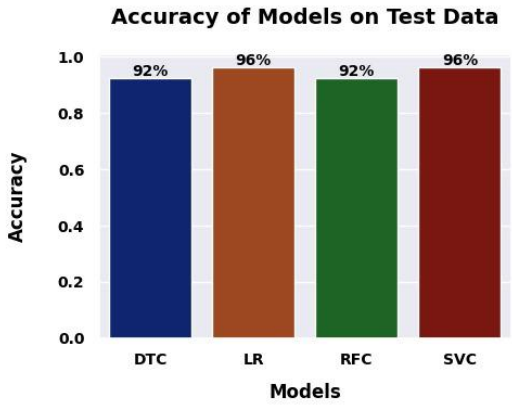


Fig 5: Accuracy score for the machine learning models

4.2.2 Bagging

Bagging is an ensemble technique that reduces variance in models like decision tree, random forest, support vector machine, and logistic regression, improving accuracy and preventing overfitting. Figure 6 highlights the improved accuracy recorded by bagging the developed models.

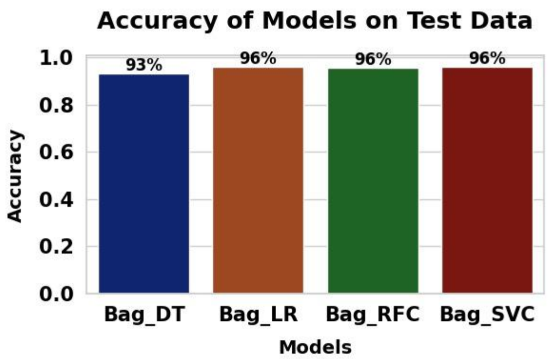


Fig 6: Accuracy score of the bagged models

4.3 Discussion of Results

The results of this study provide insights into the performance and efficacy of the models deployed for predicting and detecting falls in elderly persons. To ensure safety, the data used were curated and pre-processed, including the removal of irrelevant attributes and addressing any missing values. The initial dataset comprised 2040 rows and 7 columns. The dataset was visualized, and it was discovered three possible outcomes (0-no fall, 1-fall prediction, 2-fall detection).

The dataset used in the study was pre-processed and trimmed down to 1409 rows of 7 features: Distance, Blood Pressure, Heart Rate Value, Sugar Level, SpO_2, Accelerometer. The feature called "Decision" is the outcome of the variable; the outcome is encoded in three forms: 0: no fall predicted or detected, 1: fall predicted (slipped or tripped), or 2: fall detected an actual fall has occurred). The dataset was split in the ratio of 80% for training and 20% for testing the four selected models. The bagging and boosting techniques reduced variance and bias in some of the models, thereby enhancing their performances.

As documented in Table 2, the DT model recorded an accuracy of 92% that increased to 93% with the bagged model; likewise, RFC accuracy improved from 92% to 96% with the bagged technique. However, LR and SVC models with an accuracy of 96%, respectively, remained unchanged with the bagged versions of the LR and SVC models.

Table 2. Model Performance Evaluation

Model	Accuracy score	Bagging accuracy
Logistic Regression	96%	96.0%
Decision Tree	92%	93.2%
Random Forest	92%	96.4%
Support Vector Classifier	96%	96.0%

4.4 Deployment

Flask, a web-based framework built on Python, was utilized for deployment; the Python programming language was used for scripts, while Visual Studio Code (VSCode) was adopted for the integrated development environment of the system. The file was saved as a pickle file, which is lightweight and ideal for saving machine learning models. Render, a hosting server for web applications, hosted the system to ensure efficient model storage. This approach allows for seamless integration and deployment of the trained model.

This user interface, as seen in Figure 7, is designed to allow users to input critical parameters that are physiological and movement-related, such as heart rate variability (HRV), blood pressure, sugar levels, oxygen saturation (SpO_2), and accelerometer data for elderly or at-risk individuals. Additionally, it collects contextual information such as the individual's name, a caretaker’s phone number, and distance measurements. To check for fall risks, required values are entered in the provided fields on the interface, and once the “Predict” button is clicked, the system analyzes the inputs to see if the individual has fallen and sends alerts accordingly via SMS to the provided phone number, as shown in Figure 8.

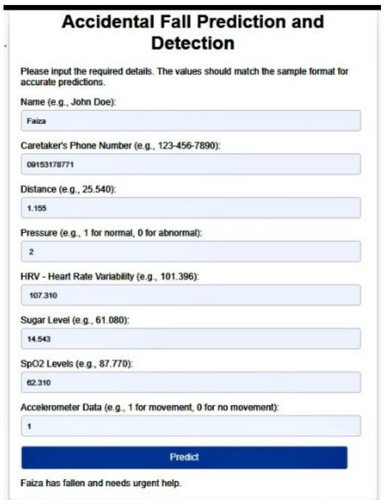


Fig 7: The user interface

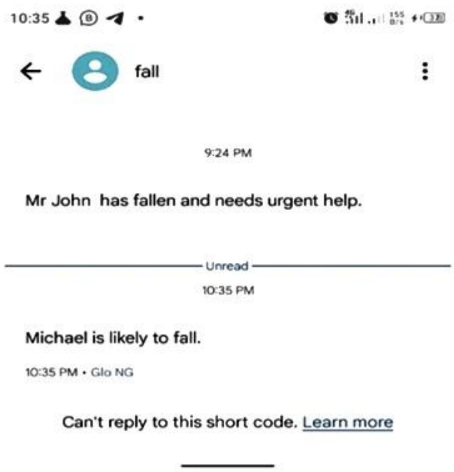


Fig 8: SMS fall status alert

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

This research developed an Accidental Fall Prediction and Detection system for elderly individuals, the study achieved this by using ensemble learning techniques, specifically bagging, with models including Random Forest, Logistic Regression, Support Vector Classification (SVC), and Decision Trees. The system was trained on a dataset that incorporates both physiological and environmental features to detect early indicators of falls. The developed system demonstrates high responsiveness at 96% accuracy for the bagging Support Vector Classifier, Random Forest, and Logistic Regression respectively while Bagging Decision Tree also performed closely with an accuracy of 93%. By integrating SMS alerts and dashboard notifications, timely alerts and potential interventions are ensured to enhance elderly well-being, hence demonstrating potential to reduce the risk of injuries, hospitalizations, and mortality occurring through accidental falls. While the results are promising, further real-world validation is necessary to confirm its impact on reducing fall-related injuries.

5.2 Suggested Areas for Future Research

Real-time human activity recognition utilizing CNN models can be undertaken to facilitate the same prediction and detection of falls. Expanding the dataset for greater diversity while incorporating various image-based data to serve as input to train models to enhance and distinctively isolate falls from other activities. Furthermore, this model may be integrated with IoT to obtain sensor parameters in real time, thereby eliminating the need for human input on the web application.

6. REFERENCES

[1] World Health Organization. 2023. Falls.

[2] Seni, G., and Elder, J. 2010. Ensemble methods in data mining: improving accuracy through combining predictions. Morgan and Claypool Publishers.

[3] Rosen, T., Mack, K. A., and Noonan, R. K. 2013. Slipping and tripping: Fall injuries in adults associated with rugs and carpets. Journal of Injury and Violence Research, 5, 61-69.

[4] Pinto, A.B.A., de Assis, G.A., and Torres, L.C.B. 2022. Wearables and Detection of Falls: A Comparison of Machine Learning Methods and Sensors Positioning. Neural Process Lett 54, 2165 -179. https://doi.org/10.1007/s11063-021-10724-2

[5] Noman D., Martijn C. S., Martijn W H., Otto, M., Jonathan B., Nathalie V. V., and Ameen, A. H. 2023. Predicting future falls in older people using natural language processing of general practitioners' clinical notes, Age and Ageing, 52,(4), https://doi.org/10.1093/ageing/afad046

[6] Mohan, D., Al-Hamid, D. Z., Chong, P. H. J., Gutierrez, J., and Li, H. 2024. Fall prediction in elderly through vital signs monitoring - A fuzzy-based approach. IEEE Internet of Things Journal, 11(20), 33439 - 33449. https://doi.org/10.1109/JIOT.2024.3429516.

[7] Krishnan, V., Abhilash, B. K., and Govind, S. 2024. A robust fall detection system for elderly persons using YOLO. 2024 5th International Conference on Innovative Trends in Information Technology. https://doi.org/10.1109/ICITIIT61487.2024.10580626

[8] Sarobin, M. V. R., Jishnuraj, K., Anbarasi, L. J., Rukmani, P., Jasmine, S. G., and Narendra, M. 2022. Fall detection among elderly persons using FallCNN and transfer learning models. Preprints. https://doi.org/10.21203/rs.3.rs-1884093/v1

[9] Al-Tamimi, A., Hussain, A., and Alamri, A. 2021. Fall detection and alert system for elderly people based on accelerometer and gyroscope data. SN Computer Science, 2(2), 1-10.

[10] Nahian, M. J. A., Ghosh, T., Banna, M. H. A., Aseeri, M. A., Uddin, M. N., Ahmed, M. R., Mahmud, M., and Kaiser, M. S. 2021. Towards an accelerometer-based elderly fall detection system using cross-disciplinary time series features. IEEE Access, 9, 22355–22367. https://doi.org/10.1109/ACCESS.2021.3056441

[11] Rachankada, S., Basavaraju, R. T., and Kulkarni, P. 2020. Good-eye: a device for automatic prediction and detection of elderly falls in smart homes. International Journal of Pervasive Computing and Communications, 16(3), 255-271.

[12] Kim, Y. S., Park, K. H., and Kim, Y. T. 2020. Real-time fall detection using CNNs and an IoT-based wearable device. Sensors, 20(17), 4956.

[13] Yhdego, H., Li, J., Paolini, C., and Audette, M. 2021. Wearable sensor gait analysis of fall detection using attention network. 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 3137-3141.

[14] Zhou, M., Zhang, G., Wang, N., Zhao, T., Liu, Y., Geng, Y., Zhang, J., Wang, N., Peng, N., & Huang, L. 2024. A novel score for predicting falls in community-dwelling older people: a derivation and validation study. BMC Geriatr 24, 491. https://doi.org/10.1186/s12877-024-05064-4