Article

Frailty Insights Detection System (FIDS)—A Comprehensive and Intuitive Dashboard Using Artificial Intelligence and Web Technologies

Bogdan-Iulian Ciubotaru 1*, Gabriel-Vasilică Sasu 2, Nicolae Goga 3, Andrei Vasilăteanu 3, Iuliana Marin 3, Ionel-Bujorel Păvăloiu 3 and Claudiu Teodor Ion Gligore 4

1 Military Equipment and Technologies Research Agency (METRA), Ministry of National Defence, 077025 Clinceni, Romania; ciubotarubogdaniulian@gmail.com
2 Faculty of Automatic Control and Computers, National University of Science and Technology Politehnica Bucharest, 060042 Bucharest, Romania; sasugabriel94@gmail.com
3 The Faculty of Engineering in Foreign Languages, National University of Science and Technology Politehnica Bucharest, 060042 Bucharest, Romania; n.goga@rug.nl (N.G.); marin.iulliana25@gmail.com (I.M.); bujor.pavaloiu@upb.ro (I.-B.P.)
4 National Institute of Medical Expertise and Recovery of Working Capacity, 077145 Pantelimon, Romania; drclaudiugligore@gmail.com

* Correspondence: andrei.vasilateanu@upb.ro

Abstract: Frailty, known as a syndrome affecting the elderly, have a direct impact on both social well-being and body’s ability to function properly. Specific to geriatric healthcare, the early detection of frailty helps the specialists to mitigate risks of severe health outcomes. This article presents the development process of a system used to determine frailty-specific parameters, focusing on easy-to-use, non-intrusive nature and reliance on objectively measured parameters. The multitude of methodologies and metrics involved in frailty assessment emphasize the multidimensional aspects of this process and the lack of a common and widely accepted methodology as being the gold standard. After the research phase, the frailty-specific parameters considered are physical activity, energy expenditure, unintentional weight loss, and exhaustion, along with additional parameters like daily sedentary time, steps history, heart rate, and body mass index. The system architecture, artificial intelligence models, feature selection, and final prototype results are presented. The last section addresses the challenges, limitations, and future work related to the Frailty Insights Detection System (FIDS).

Keywords: frailty detection system; internet of things; artificial intelligence

1. Introduction

The availability of easy-to-use devices and smart platforms enables individuals to monitor their activities and act proactively to prevent health problems, which can mitigate the risk of severe health complications. Integrating technology in health management may reduce the costs of treatment, detect different conditions in early stages, and improve life quality [1].

Life expectancy has doubled from 1900 to the present [2], leading to an exponential aging of the population, making frailty one of the most concerning subjects related to geriatric medicine [3]. As Fried et al. [4] described, this condition can be synonymous with disability, comorbidity, and is highly prevalent in old age, conferring a high risk of disability, hospitalization, and mortality. Frailty is classified as a syndrome that can be prevented or treated if detected in its early stages [5], but it is also an important component of decision making and effectively prevents and treats perioperative complications in elderly patients [6].
The scientific literature emphasizes that frailty is a geriatric syndrome in which there is a notable decrease in the body’s ability to function across various organ systems, leading to a higher chance of health problems. Studies demonstrate that frail older people have a significantly increased risk of death compared to non-frail ones. The risk of death is even higher for frail older people who have multiple chronic conditions. Frail older people should be identified and targeted with personalized interventions that may reduce their risk of death [7]. Chronic diseases, such as diabetes, hypertension, and heart disease, can contribute to frailty, as can medication side effects, sleep disturbances, and social isolation [8]. Frailty management may benefit from software-driven detection approaches, enhancing early detection, targeting a wider audience by modern day devices and technology usage. Kouroubali et al. [9] discussed the need for better coordination among healthcare professionals, patients, and caregivers and introduced a digital platform leveraging innovative tools and artificial intelligence technologies to facilitate the coordination. Sauzéon et al. [10] analyzed the potential of ambient assisted living (AAL) technologies to support aging in place for frail older adults, to prevent autonomy loss and institutionalization. Their research shows that this technology can improve daily life, reduce hospital visits, and support mental health by meeting the specific needs of older adults. Bian et al. [11] developed and tested a sensor-based system to monitor frailty at home. Their found that the early detection of frailty is crucial, and many of the sensors used were reliable, except for the smart weight scale.

This article describes a functional system prototype, Frailty Insights Detection System (FIDS), using the following parameters: physical activity, energy expenditure, unintentional weight loss, exhaustion, sedentary behavior, steps history, heart rate variation, and body mass index. The main motivation for this new approach is to develop a system that is easy to use, non-intrusive, and based on objective parameters that can be determined by using a smart device’s sensors, to avoid as much as possible subjective input, but without affecting scientific reliability.

The functional system prototype integrates various key parameters that have been linked with frailty in academic studies. Through advanced sensor technology, data analysis algorithms, and user-friendly interfaces, the system is addressed to both healthcare professionals and individuals to assess and monitor frailty-related factors comprehensively, in an objective manner.

The entire process follows a meticulous and systematic approach, starting with a user requirements questionnaire, followed by the system design and system development phases. The paper has the following structure: Section 2 describes the most common frailty detection methodologies and presents a comparison in terms of the used parameters. It also presents the system architecture, emphasizing the main components with their functions and the manual and automatic feature selection processes. The results are presented in terms of mechanisms and software/hardware subsystems with illustrations of each frailty-related parameter. The last two sections present the discussion, with achievements and limitations, and conclusion, emphasizing future research directions.

A substantial part of the research presented in this paper was performed in the context of a Eureka research and development project, called cINnAMON, ‘A Non-Intrusive Home Surveillance System for Assisting Elderly Persons with Frailty Risk’. The cINnAMON project aims to develop a non-intrusive, affordable home surveillance system to be used primarily in assisting elderly persons at risk of frailty.

2. Materials and Methods

This section presents the main frailty detection methodologies, followed by an analysis of different parameters along with the scoring mechanisms for each methodology. It introduces an architecture that integrates Internet of Things (IoT) components with artificial intelligence (AI) classification models and web-specific technologies. This combination aims to promote self-awareness and facilitate remote health monitoring, with a special focus
on elderly frailty syndrome. Additionally, the section covers the feature selection process for artificial intelligence models, presenting both manual and automatic selection processes.

2.1. Analysis of Frailty Detection Methodologies

Frailty is a complex and multifactorial condition that affects many older adults, characterized by a decline in physiological and cognitive function, decreased mobility, and increased vulnerability to stressors and disability. Several methods for detecting signs specific to frailty were developed and tested. Multiple reviews related to frailty measurements were conducted to better understand the concept of frailty [12,13]. The most widely known and tested methodologies are the following:

- Fried Frailty Phenotype;
- Rockwood Frailty Index;
- Groningen Frailty Indicator;
- Tilburg Frailty Indicator;
- Edmonton Frail Scale;
- FRAIL Scale.

Most methodologies have common parameters that are used for frailty detection. Table 1 illustrates the most common parameters and their occurrence in the frailty detection methodologies mentioned. Although most parameters can be determined through predefined, subjective questionnaires dependent on the patient’s response, their outcome can also be determined objectively using IoT systems.

Table 1. Comparison of frailty detection methodologies in terms of parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fried Frailty</th>
<th>Rockwood Index</th>
<th>Groningen Indicator</th>
<th>Tilburg Indicator</th>
<th>Edmonton Scale</th>
<th>FRAIL Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unintentional Weight Loss</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Physical Activity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Exhaustion or Fatigue</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Cognitive Function</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Psychological Health</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Social Aspects</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Multiple Health Issues/Chronic Diseases</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Functional Independence</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Detecting frailty is a complex task that requires measuring different metrics of physical, cognitive, social, and medical information. To develop a system capable of detecting signs that may indicate a person’s frailty, it is necessary to determine the parameters that can be objectively measured. An analysis of how to determine the state of frailty was carried out based on the selected methodologies.

The scoring mechanism for the selected methodologies is presented in Table 2. As can be observed, most methodologies use a specific scale by assigning a unit for the presence of a parameter (for example, if the patient unintentionally loses weight, the value is 1), with the final score being the sum of these values. In the case of the Tilburg indicator, 1 point is awarded for a ‘Sometimes’ response and 2 points for ‘Yes’.

In the case of the Rockwood Index, which can also be called the cumulative deficit index, frailty results from the accumulation of various health deficits over time, dependent on age and gender. To determine the index, a series of symptoms, signs, abnormal laboratory test values, functional impairments, and chronic diseases are defined, tailored to a target group under analysis. The frailty score is defined as a fraction between the
sum of the present deficits (1—deficit present/0—deficit absent) and the total number of defined deficits.

Table 2. Scoring mechanisms of frailty methodologies.

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Number of Items</th>
<th>Scoring Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fried Frailty</td>
<td>5</td>
<td>0: Not frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1–2: Pre-frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 or higher: Frail</td>
</tr>
<tr>
<td>Rockwood Index</td>
<td>N/A</td>
<td>0&lt;0.1: Not frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1–0.2: Mildly frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2–0.3: Moderate frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;0.3: Severely frail</td>
</tr>
<tr>
<td>Groningen Indicator</td>
<td>15</td>
<td>0–4: Not frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5–6: Mildly frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7–8: Moderately frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9 or higher: Severely frail</td>
</tr>
<tr>
<td>Tilburg Indicator</td>
<td>15</td>
<td>0–5: Not frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6–11: Mildly frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12–17: Moderately frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18 or higher: Severely frail</td>
</tr>
<tr>
<td>Edmonton Scale</td>
<td>17</td>
<td>0–5: Not frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6–7: Vulnerable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8–9: Mildly frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10–11: Moderately frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 or higher: Severely frail</td>
</tr>
<tr>
<td>FRAIL Scale</td>
<td>5</td>
<td>0: Not frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1–2: Pre-frail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3–5: Frail</td>
</tr>
</tbody>
</table>

The methodologies for detecting vulnerability are methodologies that depend on the responses of the surveyed individuals. In this regard, the degree of precision of the result has a subjective character, requiring the surveyed person to remember relevant data.

2.2. System Architecture

Multiple frailty detection methodologies were presented, each one being based on objectively and subjectively measurable parameters. The system presented in the current paper uses the following parameters, determined through objective methods:

1. Physical activity;
2. Energy expenditure;
3. Unintentional weight loss;
4. Exhaustion or fatigue;
5. Daily sedentary time and steps history;
6. Heart rate values daily variation;
7. High body mass index.

The parameters were chosen, considering their scientific relevance in frailty detection, objective measurability that ensures precision and reduces subjectivity, and the feasibility of continuous monitoring using current technologies. In what follows, the system architecture is described from a software and hardware perspective, with an emphasis on the functional components and modules.

Figure 1 presents the high-level architecture of FIDS, the Frailty Insights Detection System. This architecture has a correspondence to previously published architectures, at a higher level, with a different use case [14].
were developed using the Fitbit environment. 

2.2.2. Frailty Insight Detection System (FIDS)

This component obtains raw format data from the user through a smartwatch. The component is an application that operates in the Fitbit environment and consists of two modules: the application that runs at the smartwatch level and the companion application in the native Fitbit app (Figure 2), which runs at the smartphone level. Both applications were developed using the Fitbit environment.

![Figure 2. Workflow of the smartwatch application.](image)

**Figure 2.** Workflow of the smartwatch application.

Fitbit Device Application

This module has the role of reading sensor data and forwarding these data. Initially, sensors are initiated and the reading frequency is set for the accelerometer, gyroscope, orientation sensor, and optical sensor. After initialization, the application enters an infinite loop, reading data from the sensors and sending them forward to the companion through peerSocket connection.

Fitbit Companion Application

This module acts as an interface between the smartwatch application component and the Frailty Insight Detection System (FIDS). Data are collected from the smartwatch via peerSocket, formatted into JSON, added with the phone’s location as longitude and latitude, and sent to FIDS. To optimize the data transmission process, data are sent in batches, where a fixed number of records are stored in a buffer, and when the buffer is full, the data are forwarded to be added to a queue. From this queue, data are transmitted through a secure WebSocket to FIDS. Before transmission, the WebSocket status is checked (it can be OPEN/CLOSED). When the WebSocket is successfully opened, the data are transmitted.

2.2.2. Frailty Insight Detection System (FIDS)

FIDS is composed of two distinct modules that are hosted on a Virtual Private Server (VPS). A VPS is a virtualized server that acts as a dedicated server within a larger server. FIDS operates in an isolated setting, meaning that the operations do not interfere with other operations. This setup was achieved by using the operating system Ubuntu 20.04, running on a 2 GB RAM, 1vCPU, and 50 GB storage system. The two modules and their roles are as follows:
NodeJS CoreHub with Prediction Engine (NCHPE)

This module acts as the core module of the FIDS. It is responsible for secured communication with the smartwatch application (SA) and comprises two back-end technologies: NodeJS and Flask. Both technologies are web frameworks used to build web applications and API services. The architecture and data workflow are presented in Figure 3.

Upon initializing the NodeJS module, secure WebSockets over the Internet are created to receive real-time data from the smartwatch. After defining the WebSocket, this module creates a secure web server (by using an SSL certificate), defining two routes: a GET route named /getData, which can be called by other components, transmitting real-time received data, and a POST route named /insertData, which establishes a connection with an external MongoDB database. The last route takes as input the data to be inserted and the collection name to which these data belong (for example, data received directly from the bracelet are inserted into the FIDSData collection).

Another important function of this module is to spawn a child process. The newly created process receives data from the main processes and formats them so that they can be used for prediction. Based on existent data, new data are computed (for example, vector magnitude for acceleration or the three Euler angles roll, pitch, and yaw) so that a new packet is formed for prediction purposes.

The Flask module is responsible for loading three artificial intelligence models that are used to determine the type of movement into memory. These models were trained locally and exported using the Python joblib library. The POST route named /predict is defined such that, when receiving a dataset as input, it is processed to match the specific header for the three models. A movement-type prediction is made using the models, and the predictions are subsequently transmitted in the JSON format.

Communication between the NodeJS module and the Flask module is bidirectional. Once the prediction is completed, it is transmitted to the ActivityCompute subprocess, which is responsible for creating a new packet that incorporates the prediction timestamp and inserts using /insertData into the specific collection the type of activity predicted (in this case, ActivityPrediction). Additionally, ActivityCompute adds to the new packet, based on each prediction, the number of kcal specific to the type of activity.

Figure 3. NCHPE architecture.
User Interface Dashboard for Frailty Insights (UIDFI)

The User Interface Dashboard for Frailty Insights module of the system has a Word-Press instance as its main component, which interacts through the API with the NCHPE module. This component of the system has as its primary function the display of the data extracted using NodeJS CoreHub with Prediction Engine in a user-friendly way, in interactive dashboards, as follows:

- Displaying data received in real time by calling the /getData function, as well as extracting relevant insights.
- Displaying self-report forms, which aim to query the user about their physical condition, mood, and body weight.
- Establish a secure connection to the Fitbit cloud and obtain stored data related to heart rate values.
- Composing charts to display the history of various parameters (weight, daily variation of heart rate, and determining sedentary times).

This module is composed of JavaScript routines that use open source libraries to perform various functions. For connecting to the Fitbit cloud, the module authenticates the user using an access token obtained using OAuth 2.0 mechanisms to authorize it to access the data. The authentication is performed using a unique ClientSecret and a ClientID obtained and used also for the smartwatch application.

Each function will be described in the following sub-sections to establish the connection between the method of determining a particular parameter and its association with fragility.

2.2.3. Non-Relational Database

The non-relational database module (Figure 4) is represented by a MongoDB database, a NoSQL document-oriented model. This database can store data in the JSON format; each dataset has a unique structure, organized in collections. The module is composed of collection and pipeline algorithms, which are used for data aggregation based on different rules. The aggregation pipeline is a MongoDB framework that is designed for data transformation based on custom filters, grouping rules, and sorting. The data are retrieved using the Fitbit Web API, using functions described in the official SDK documentation. The data retrieved are in the standard JSON format and can be used in front-end routines.

![Figure 4. MongoDB database structure.](image)

The purpose of this module is to store the FIDS data, aggregate them according to specific rules, and make these data available for further processing and access. The structure of the database collection is represented in Figure 5. Each collection and pipeline have the following roles:
- FIDSData—stores raw data received from the smartwatch.
- ActivityPrediction—stores activity-type predictions (5 s granularity).
- EnergyExpenditure—stores energy expenditure for each activity (minute granularity).
- HourExpenditure—stores energy expenditure for each activity (hour granularity)—this collection is created by the MinuteToHour pipeline and has as the main source the EnergyExpenditure collection.
- SedentaryPrediction—stores how many minutes each activity is performed (hour granularity)—this collection is created by the SedentaryPrediction pipeline and has as its main source the EnergyExpenditure collection.
- UserInformation—stores information provided by the user through self-report forms (weight and mood).

![Non-relational database structure](image)

**Figure 5.** Non-relational database structure.

Each collection is later accessed by using the GET endpoint /getData from NodeJS modules and used for both processing and displaying purposes.

2.3. Feature Selection for Physical Activity Detection

Previous work [15] demonstrated that it is feasible to use artificial intelligence classification models to classify different activity types. In a frailty context, the aim is to assess how much time a person performs a certain physical activity. This information is used to determine different frailty insights, such as daily energy expenditure, physical activity levels, and daily sedentary times.

Data were collected for four types of activities: Fast walk (the subject is walking at a fast pace), Slow walk (the subject is walking at slow pace), Resting (the subject is standing still), and Climbing upstairs (the subject is going up a staircase).

*By using data acquired from a smartwatch’s sensor, the aim is to determine four main activities performed by an elderly person during a day: resting, walking, running, and climbing stairs. For these activities, five one-minute recording sessions were carried out. A five-minute pause was taken between sessions to avoid data compromise. The sensors (Fitbit Versa built-in sensors: accelerometer, gyroscope, orientation sensor, and optical sensors) were set to record data ten times per second (except the optical sensor, which only recorded one datum per session) [15]. Data recording sessions were*
performed by one healthy individual, with an average weight and height and no health conditions, classified as non-frail.

Based on our previous analysis detailed in Ciubotaru et al. (2023) [15], the decision tree classifier (DTC), random forest classifier (RFC), and gradient boosting classifier (GBC) models were identified as the most reliable in terms of performance. This selection was made after comparing these models against other algorithms, such as logistic regression (LR), k-nearest neighbors (KNN), support vector (SVC) and Gaussian naive Bayes (GNB) classifiers. The chosen models demonstrated remarkable consistency in their performance metrics across diversified datasets [15], making them the optimal choice for developing a high performance classifier for activities.

A total of 5515 data samples was recorded using smartwatch sensors, annotated using the methodology presented by Ciubotaru et al. 2023 [15]. The set has the following attributes:

- Accelerometer_x, Accelerometer_y, and Accelerometer_z—accelerations (m/s^2), obtained from the accelerometer sensor; and AccMagnitude—vector magnitude derived from accelerations (m/s^2);
- Gyroscope_x, Gyroscope_y, and Gyroscope_z—angular velocities (rad/s), obtained from the gyroscope sensor;
- Roll, pitch, and yaw—rotation angles in three angles, derived from quaternions (rad), obtained from the orientation sensor;
- Activity: annotated type of activity for models to distinguish between (resting, walking, running, and stair climbing).

To identify the optimal set of attributes, considering the three chosen algorithms, a routine was developed that takes the default dataset as the input and calculates all possible feature combinations, ranking them based on their performance. To assess the performance of a combination, accuracy, a specific machine learning metric, was used. Accuracy is the ratio between the correct predictions and the total number of predictions, in this specific case, the ratio between the correct classifications and total classifications.

Considering that the default attribute set has 10 features, the total number of feature combinations is 2^{10} − 1, meaning 1023 possibilities. To check each possibility’s performance, the Python class concurrent.futures library was used. This library provides the capability of asynchronous execution of code, using a pool of processes that are running in parallel. The number of processes is defined by the max_workers variable, allowing multiple CPU cores to execute instructions concurrently. After running the simulation with 10 max_workers (allowing Python to create 10 concurrent subprocesses), the performance metrics for each feature combination were saved.

In Figure 6, the distributions for each model are illustrated, with the upper limit being the maximum accuracy for the specific number of features, while the lower limit represents the minimum accuracy. As observed, starting with attribute sets containing at least six features, the accuracy values exceed 85%, while all combinations of at least nine features have a minimum accuracy of 90%. It is noteworthy that relevant information can be extracted with fewer features.

In Table 3, the first 5 feature combinations are sorted based on accuracy, in descending order, according to the average of the chosen 3 models. As observed, in terms of percentages, the differences are insignificant. Also, considering the randomness of training data, the actual combinations of features presented in table can be similar, and further analysis is required. It should be mentioned that, after analyzing the top 100 combinations by average accuracy, an average value of 7.18 features per attribute set was determined to achieve optimal results.

The best 100 feature combinations were determined in terms of accuracy, with a median value of accuracy among the three classifiers of 0.9524, ranging from 0.9607 to 0.9456. It is plausible to state that a combination of 7 collected features leads to an average 95% accuracy for the models considered. During the feature combination search, the confusion matrix was computed. A confusion matrix is a metric used to evaluate the performance. It consists of a table that summarizes the number of correct and incorrect predictions, categorized by...
each class. Figure 7 shows the distribution of the sums of diagonal values of the confusion matrix for the same 100 best combinations of characteristics. The average diagonal sum for all records is 1050, which is 95% from the 1100 maximum sum (20% of total attributes, 5515). The random forest diagonal elements sum ranges around 1070–1073, while the gradient boosting range is 1064–1070. Correlated with Figure 6, the decision tree has the lowest sum.

Figure 6. Accuracy distribution for the number of features.

Table 3. Top 5 combination of features by median accuracy.

<table>
<thead>
<tr>
<th>Features</th>
<th>Random Forest</th>
<th>Gradient Boosting</th>
<th>Decision Tree</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accelerometer_x, Accelerometer_y, Gyroscope_x, Gyroscope_y, Gyroscope_z, AccMagnitude, Roll, Pitch, Yaw</td>
<td>0.9728</td>
<td>0.9701</td>
<td>0.9393</td>
<td>0.9607</td>
</tr>
<tr>
<td>Accelerometer_x, Accelerometer_y, Accelerometer_z, Gyroscope_y, Gyroscope_z, Roll, Yaw</td>
<td>0.9692</td>
<td>0.9646</td>
<td>0.9474</td>
<td>0.9604</td>
</tr>
<tr>
<td>Accelerometer_x, Accelerometer_y, Accelerometer_z, Gyroscope_y, Gyroscope_z, Roll, Pitch, Yaw</td>
<td>0.9710</td>
<td>0.9655</td>
<td>0.9447</td>
<td>0.9604</td>
</tr>
<tr>
<td>Accelerometer_x, Accelerometer_y, Gyroscope_y, Gyroscope_z, AccMagnitude, Roll, Pitch, Yaw</td>
<td>0.9683</td>
<td>0.9674</td>
<td>0.9456</td>
<td>0.9604</td>
</tr>
<tr>
<td>Accelerometer_x, Accelerometer_y, Accelerometer_z, Gyroscope_y, Gyroscope_z, AccMagnitude, Roll, Pitch, Yaw</td>
<td>0.9710</td>
<td>0.9655</td>
<td>0.9438</td>
<td>0.9601</td>
</tr>
</tbody>
</table>

After manual analysis, an automated analysis using the sklearn.feature_selection module, used for the automatic selection of features, was performed. The methods used were as follows:

- Recursive Feature Elimination (RFE), a feature selection technique that recursively removes the least important feature and evaluates the performance for the new set of features;
- SelectKBest, a method that helps to select the K best features, with a predefined K, with chi-squared test that is specific for classification.
In case of RFE, the code module was adapted to run subsequently for 5, 6, 7, 8, and 9 features out of 10. The results illustrated in Table 4, where gray means that the feature was included in the results, show that the best features, considering the occurrence, are those related to acceleration. The features that are highlighted in black in Tables 4 and 5 are the ones that were not selected from the full list of initial features. Other important parameters are the angular velocities on the Y- and Z-axes, while the rotation angle pitch can be eliminated.

![Distribution of average diagonal sums across models](image)

**Figure 7.** Distribution of average diagonal sums for confusion matrix.

**Table 4.** Recursive feature elimination (RFE) analysis.
Table 5. SelectKBest chi-squared test analysis.

<table>
<thead>
<tr>
<th>Number of Features</th>
<th>acc_X</th>
<th>acc_Y</th>
<th>acc_Mg</th>
<th>gyro_X</th>
<th>gyro_Y</th>
<th>gyro_Z</th>
<th>Pitch</th>
<th>Roll</th>
<th>Yaw</th>
<th>RFC Accuracy</th>
<th>GBC Accuracy</th>
<th>DTC Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9680</td>
<td>0.9695</td>
<td>0.9289</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9630</td>
<td>0.9644</td>
<td>0.9340</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9630</td>
<td>0.9608</td>
<td>0.9419</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9615</td>
<td>0.9601</td>
<td>0.9274</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9572</td>
<td>0.9586</td>
<td>0.9282</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9441</td>
<td>0.9427</td>
<td>0.9245</td>
</tr>
</tbody>
</table>

SelectKBest was used with the chi-squared test, which helps to find a significant difference between the expected outcomes and predicted outcomes in different categories, which were activities in this specific case. Like in the RFE, the analysis was performed for various values of K, from 10 to 5. As it can be observed in Table 5, in that specific feature selection method, the accelerometer values keep their importance, while the rotation angles are more important than the gyroscope. Considering that this analysis is independent from any model, after the best feature set was selected, models with those features were trained, and the accuracy was computed.

It is more than clear that not all data are required for a good classification of an activity type. The new dataset has a reduced header, presented in Table 6: acc_X, acc_Y (accelerations), acc_Mg (acceleration vector magnitude), gyro_X, gyro_Y, gyro_Z (angular velocities), roll, yaw (rotation angles), and activity type. While there might be better combinations of features, the differences between the accuracies for eight features were insignificant.

Table 6. Data header.

<table>
<thead>
<tr>
<th>acc_X</th>
<th>acc_Y</th>
<th>acc_Mg</th>
<th>gyro_X</th>
<th>gyro_Y</th>
<th>gyro_Z</th>
<th>roll</th>
<th>yaw</th>
<th>activity</th>
</tr>
</thead>
</table>

After the models were trained, the confusion matrix was calculated using a built-in function from sklearn for each model. The confusion matrix is shown in Table 7, and it represents in rows the actual class (running, walking, resting, and stair climbing) and, in the same order, the predicted class in columns. The correspondence between activity types and rows/columns is: R1/C1 means running, R2/C2 means walking, R3/C3 means resting, and R4/C4 stair climbing.

Table 7. Confusion matrix for the models.

<table>
<thead>
<tr>
<th>Gradient Boosting Classifier</th>
<th>Decision Tree Classifier</th>
<th>Random Forest Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 C2 C3 C4</td>
<td>C1 C2 C3 C4</td>
<td>C1 C2 C3 C4</td>
</tr>
<tr>
<td>R1 119 2 2 3 113 3 4 6 123 1 2 0</td>
<td>R2 1 312 1 0 2 295 1 16 0 309 1 4</td>
<td>R3 1 2 578 1 3 2 575 2 2 2 578 0</td>
</tr>
</tbody>
</table>

The Accuracy, F1 score, Precision, and Recall were calculated for each model, and the overall results have an accuracy of over 95%:

- RFC: Accuracy: 0.972, F1 Score: 0.972, Precision: 0.972, Recall: 0.972.
- GBS: Accuracy: 0.97, F1 Score: 0.969, Precision: 0.969, Recall: 0.977.
- DTC: Accuracy: 0.939, F1 Score: 0.938, Precision: 0.938, Recall: 0.939.

The learning curve for artificial intelligence models is an instrument used to analyze the performance of a model and to visualize the evolution of the model performance on both
the training and validation sets. The scope of analysis is to check if models are overfitted or underfitted. Overfitting occurs when a model captures noise from input data, resulting in a model with good performance on training data, but cannot distinguish between activities on new, unseen data. With underfitting, the model cannot capture patterns in the data, resulting in a low performance on both the training and unseen data.

Fitting is an important stage in defining a model and is used to ensure that the model predicts unseen data correctly. Learning curves use only data used for training (as previously mentioned, it represents 80% of full input data) in a process called cross-validation. This process depends on a variable called folds of cross-validation (cv), which defines the number of how many subsets are created from a set. For example, if a set has X values, the data are divided into the number of folds defined subsets used in the cross-validation process. The cross-validation process involves training and iterating the model on different subsets of data for performance evaluation.

In Figure 8, the graphs that show the learning curve for the three chosen models are illustrated. The vertical axis shows the accuracy of the model, while the horizontal axis displays training datasets (80% of the dataset). The red line, representing the training score, shows how the model performs on the training data as the training data size increases. The learning curve displays an accuracy starting from 10% of total training data with an increasing size of 10% until it reaches 100%, with a number of 10 trainings for each model. For each dataset (for example, the first 10% of data), the model is trained in an iterative way for cv (folds number) times and is validated on the same data. For a better understanding, the data are divided into cv subsets, the model is trained on cv-1 sets (the number of folds used is 5, meaning that the data are divided into 5 sets), and the accuracy is tested on the same cv-1 (the initial 10% of data, which was divided into 5 folds, is used to train and validate the model five times, each time with different 4 folds).

Using the green line, the accuracy illustrated is the median value of cross-validation scores. The process is like the one for the red line, but the validation is conducted on the remaining set (for a better understanding, for the case of 10% data, they are divided into five folds, where 4 are used for training and the remaining one for validation). This process is performed 5 times for each subset of the entire data. The lighter red and green areas illustrate the other accuracy values for the same subset, while the points represent its mean values.

Figure 8 shows the learning curves for the gradient boosting classifier, decision tree classifier, and random forest classifier in terms of the accuracy performance metric across variable attribute set sizes. All three plots highlight how both training and validation scores converge as the number of training attribute sets increases. While DTC and RFC start from 0 on the horizontal axis, the gradient boosting classifier starts from the third subset, because of its nature of training (it has an initialization and iterative fitting to the errors phase, compared to the other two models). While all three classifiers show a strong performance, the random forest classifier appears to be the most balanced in terms of learning from the training data and generalizing to unseen data. The decision tree classifier is the least robust, showing signs of overfitting. The gradient boosting classifier offers a middle ground solution that performs well, but with room for improvement in generalization. In short, the models are already performing well with the data provided, and adding more is unlikely to make any significant difference.

Another important analysis of the models’ performance is the loss curve, which is a graphical representation showing how the error evolves during iterations, with each iteration representing a new weak learner (with GBC) or a new tree to the forest (with RFC). Because of its non-iterative training process, it is not possible to compute the loss curve. In Figure 9, where the curve loss during 100 iterations is displayed, it can be observed that both models gradually learn as the number of iterations increases. Gradient boosting classifier has both training and testing loss starting high, decreasing gradually, with no signs of overfitting (both curves stay relatively close to each other). In the case of the random forest classifier, it can be observed training loss starts high and rapidly drops to a
low value, meaning that the model fits data well, while the testing loss starts high, with no smooth decrease, reaching a plateau after 40 iterations. Both classifiers are effective in reducing the loss as the number of iterations increases.

Figure 8. Learning curves for all three models, across the size of attribute sets.
A valid way of improving the models’ accuracy is to adjust their hyperparameters (each model has specific parameters). The best way to find the correct hyperparameters is to run a Grid Search for parameters, which is a simulation with random values for parameters. After the models are trained using a random pair of parameters, the performance is stored, and the process moves to the next random pair. The goal is to determine the pair of parameters for which the performance is the best. Such a process is resource-intensive in terms of hardware, requiring multiple successive trainings of the models using randomized parameter values.

3. Results

After the design phase of the system, a fully working prototype was deployed in a custom environment accessible using a web browser from a device connected to the Internet. The User Interface Dashboard for Frailty Insights was developed according to the architecture presented in Section 3 and contains the following modules, which address different parameters.

3.1. Physical Activity-Type Detection

The prediction results are used for both the determination of energy expenditure and the evaluation of sedentary time. For a better understanding, a front-end module was developed that displays real-time data (Figure 10). This interface is updated every 5 s by the addition of a new prediction in each classifier box in the section Most Frequent Prediction History. After 12 consecutive predictions, the minute prediction is computed and added to the Activity History table. This table is created each time the page is open, being a front-end instance. In the back-end, the ActivityCompute routine from NCHPE is responsible for computing data and inserting them into specific collections.
As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.

The number of calories consumed by the user is stored in the HourExpenditure collection for each hour. This calculation considers the predictions of the three models, the MET values specific to the identified activity types, and the self-reported weight of the user. As previously described, a significant part of the overall daily energy expenditure of an individual is determined by physical activity. For this purpose, the concept of Metabolic Equivalent of Task (MET) is used. MET is a universal tool for measuring energy consumption in relation to a reference level, considering body weight. It was first introduced by Gagge et al. in 1941 [16], and since then, multiple compendiums of physical activities were developed, where for each human activity, a correspondent MET value exists [17].

One MET is defined as the resting energy expenditure, which is 1 kcal/kg/h for an average person. It means that a 100 kg person who sits for one day has a daily energy consumption of 1 kcal × 100 × 24 = 2400 kcal. This means that, to determine how many calories were burnt, one has to multiply the MET value for a certain activity by body weight and duration in hours. Due to its general nature, MET does not consider individual differences, such as metabolic rate or muscle mass, but it can be used as an estimator.
Figure 11. Screen capture of the daily energy expenditure module.

3.3. Unintentional Weight Loss and Exhaustion

The most common parameter relevant for detecting frailty is unintentional weight loss. In the context of frailty, an unintentional weight variation of at least 5% in a downward trend is considered a warning sign. Weight determination can be conducted automatically using a smart scale that has the ability to transmit the weight to a database or by manual weighing using a traditional scale. Exhaustion can be determined statistically using user records related to mood.

For determining unintentional weight loss as well as exhaustion, a self-report survey was defined (Figure 12), which has the role of storing weight and state, together with the date at which the record was made. These data are stored using the /insertData route in the UserInformation collection. Based on the daily report, users can view history entries (Figure 13) and how they varied in the last 30 days (Figure 14).

Figure 12. Self-report survey.

Figure 13. Weight and exhaustion history.
When a survey is submitted, the information is sent to the database via a cURL request and recorded in the UserInformation collection. Independent JavaScript routines were created to extract insights and display them on the dashboard, which, using the charts.js library, plot the information in an intuitive way (Figure 15). The duration selector has three options: last month, last 2 months, and last 3 months. It automatically refreshes the display and shows data accordingly.

**Figure 14.** Frailty detection dashboard—weight and mood.

**Figure 15.** Heart rate values and evolution.
3.4. Daily Sedentary Time and Steps History

A sedentary lifestyle is one of the main issues that society needs to address. Human evolution has involved the development of tools that reduce physical activity to improve quality of life and provide access to various resources without significant physical effort. In this way, a new research direction emerged: researching and developing new tools and technologies to help people become aware of both the implications of low activity level and the long-term effects on health outcomes [18]. Sedentary behavior refers to static activities, such as resting, which consume little energy, with MET values < 1.5 [19]. Elderly people tend to become more sedentary as they age due to loss of mobility and physical abilities.

Increasing physical activity is necessary to eliminate the risk of causing fatal diseases, such as cardiovascular issues, diabetes, and mobility problems. A study conducted over an 18-year period on a group of elderly people showed that physically active people have a lower mortality rate compared to sedentary people [20]. Sedentary behavior is a causative factor of frailty, and the early detection of this lifestyle is necessary for detecting early signs of frailty.

To provide the user with information regarding the time spent conducting various activities, as well as to display a history of the daily number of steps taken, the following procedure was followed:

Daily sedentary time
Determining these times is conducted at the non-relational database level, using the SedentaryPrediction pipeline, which extracts from the EnergyExpenditure collection the times spent each hour for a certain predicted activity (for each type of activity, a number of minutes is expressed). By running this pipeline, a new collection is created, called SedentaryPrediction, where the date, hour, and number of minutes for each type of activity are stored and grouped according to the classification model. To reorganize and analyze a dataset involving the predictions of activities, we first convert the ‘time’ field to ‘dateMinute’ format (YYYY-MM-DDTHH:MM:00.000Z) and then group the data by ‘dateMinute’, retaining the first value of each prediction type. We also transform the ‘_id’ field into ‘dateHour’ format (YYYY-MM-DDTHH:00:00.000Z) and subsequently group by ‘dateHour’ to count occurrences of specific activities (Resting, Walking, Running, Stair climbing) within each prediction. Then, these grouped data are reorganized to neatly organize activity counts by prediction type. Finally, we sort the results by ‘date’ in descending order and store the organized data in the ‘SedentaryPrediction’ collection for further analysis or reporting.

Steps history
The number of steps performed is a quantitative method to assess daily activity. In determining frailty, the number of steps can be used in conjunction with the time spent on various activities (previously determined) [21]. Fitbit bracelets are recognized for their accuracy in determining the number of daily steps and are used in extensive research projects [22–24].

The daily step count is obtained by calling the Fitbit WebAPI, with this number being stored in the Fitbit cloud with day-level granularity. A PHP $_GET request is made to retrieve these data using the specific address from the official documentation. For the step count, an iterative routine is used to fetch the number of steps for each day, which are subsequently stored in the browser’s temporary memory to be displayed.

3.5. Heart Rate Values

The heart rate is an important factor that has implications both in determining frailty and in detecting specific heart diseases. In their study, Chaves et al. (2008) demonstrated that, in the case of older women who have lower heart rate variations over time, the heart rate fluctuates in a more regular and repetitive way, proving a reduced cardiovascular control adaptability, which was linked to frailty [25].

In the FIDS context, the system retrieves information using Fitbit WebAPI at a one-minute granularity level. The data are fetched and processed to display the tables and plot
from Figure 15, where different parameters like daily mean value, a five-day period mean value, and deviation from mean value are computed.

3.6. High Body Mass Index

The body mass index (BMI) is a simplified method for categorizing an individual into a risk category based on their weight and height. The BMI is calculated as follows: “the body weight, expressed in kilograms, is divided by the square of the height, expressed in meters”. The World Health Organization published a metric classification system based on the BMI in 2000, as follows [26]:

- “Underweight: <18.5 kg/m²”;
- “Normal weight: 18.5–24.9 kg/m²”;
- “Overweight: 25–29.9 kg/m²”;
- “Obesity class I: 30–34.9 kg/m²”;
- “Obesity class II: 35–39.9 kg/m²”;
- “Obesity class III: ≥40 kg/m²”.

The system calculates the BMI value and displays it, highlighting the category into which the user falls, as presented in Figure 16.

![Body Mass Index Example](image)

**Figure 16.** Screen capture of the BMI index with classification categories.

4. Discussion

The final outcome of the prototype is visually and successfully computes and displays data related to the following parameters:

- Physical activity;
- Energy expenditure;
- Unintentional weight loss;
- Exhaustion or fatigue;
- Daily sedentary time and steps history;
- Heart rate daily values variation;
- High body mass index.

Otherwise, while the chosen parameters were those that were the most mentioned in the reviewed literature, other parameters like cognitive, psychological, and social aspects cannot be assessed in a non-intrusive manner and they might require specialized assistance. Integrating these additional parameters could enhance the accuracy of the system, and more research is necessary to develop accurate and easy-to-use methods of assessing these parameters, exploring new ways of capturing the metrics of these parameters by using additional sensor data. Taking into account that physical activity, energy expenditure, and daily sedentary times are based on models trained with data from one person, the system may have a better accuracy if more diverse data are used for training, from both non-frail and pre-frail individuals, young and elderly.

The dashboard serves as a tool for specialized personnel to detect and assess insights related to frailty. Based on the information provided, the frailty syndrome can be attributed to a person linked with other specific clinical tests. Even though the initial purpose of the dashboard is frailty detection, it can also be used as a tool for monitoring a person’s overall health. In the next phase, the entire system is planned for an extended clinical study.
aimed at collecting data from elderly individuals in a coordinated, ethical, and scientifically approved manner.

Integration with wearable health devices, such as smartwatches and fitness trackers, offers real-time data to enhance the system’s accuracy. Furthermore, telehealth applications for the remote monitoring of older adults, especially those in remote or underserved areas, have significant potential. The development of personalized interventions, predictive analytics to forecast frailty risk, and customized user interfaces for older individuals are all options worth exploring. Additionally, ethical considerations, validation, and certification, user education, population-specific models, and cost-effectiveness studies are components of future work, ensuring the system’s effectiveness, ethical compliance, and accessibility while advancing the well-being of older adults.

Another limitation is that most IoT devices require an intermediary to connect to the Internet. In this case, a mobile phone is needed, which increases the cost of such a system. Addressing a limited number of detected activity types narrows the system’s accuracy in confidently predicting the exact level of sedentariness and energy consumption. At the same time, considering that the system is designed to analyze data in a distributed context over extended periods, the accuracy is offset by the nature of the process.

The subjective nature of some parameters can introduce data interpretation errors. Improving such a system is achievable by integrating smart scales, as well as algorithms that can determine gait based on the number of steps taken over a period. Another technology that can be used for a more accurate determination is indoor localization, which can also provide information about the exact distance covered and the precise step size of a person.

The development of such systems, where relevant health data about vulnerable patients are used, must take into account the ethical dimension of data collection. The security and confidentiality of these data are essential to prevent unauthorized use for purposes other than those for which they are intended. In this regard, future work involves the development of proprietary hardware solutions that transmit data in a single direction, towards an isolated on-premise system. Other aspects that require extensive research are represented by the particularities of the analyzed population, considering that the cultural and social differences that make the level of acceptance for the integration of such systems vary.

By comparing the FIDS with other similar projects, the following advantages can be determined: the system is easy to use and non-intrusive (the main component is smartwatch that does not disrupt or inconvenience daily routine of the users and a web-accessible dashboard over Internet) and cost-effective (by using a commercial off-the-shelf product that can be simply configured to be integrated in the system). Disadvantages may include the need for an Internet connection, dependence on the Fitbit ecosystem, and lack of cognitive, psychological, and social parameters.

5. Conclusions

This paper presents a functional system aimed at determining and displaying specific parameters related to frailty. The system is based on the results of a questionnaire designed to identify functional requirements and the study of the specialized scientific literature. This system is based on the architecture proposed in [14] and developed in [15]. The main specific parameters, namely the level of daily energy consumption and physical activity, are based on artificial intelligence classification models, which were determined and refined based on the results presented in [15]. The overall accuracy of the trained models is over 96%, with values between 94% and 97%, depending on the classified activity type.

Looking ahead to future developments, several opportunities are present for the system designed to determine and visualize parameters related to fragility in older adults. One option involves refining the artificial intelligence classification models to enhance their precision and dependability in assessing fragility-related parameters. This encompasses the incorporation of additional data sources and the exploration of cutting-edge deep learning techniques. Longitudinal studies provide valuable information on the progression
of fragility and the effectiveness of potential interventions. Throughout the development of the system, several challenges and research opportunities were identified that could improve such systems by addressing them. The high costs needed for a proprietary device determined the usage of market-available commercial off-the-shelf devices. In this sense, the developed system is partially dependent on the data provided by Fitbit WebAPI but independent in terms of interfacing between a smartwatch and the system. In an institutionalized framework, where the main focus point is placed on the data storage location, such a project approach is not feasible, given that some of the data are stored in an external cloud.

Future work should address a clinical trial, where pre-frail individuals could be monitored to check if the usage of such system is feasible, so it can be later be refined and transformed into a Technical Readiness Level (TRL) 9 product.

Frailty detection is an area of increasing interest within the scientific community. The progress made in recent years sets the premises for the emergence of user-friendly commercial solutions, adaptable to specific needs, and at reduced costs.


Funding: A part of the research presented in the current paper was conducted in the context of the Romanian Ministry of Education and Research, CCC DI-UEFISCDI, project number PN-III-P3-3.5-EUK-2019-0202, cINnAMON project, nr. 189/01.10.2020, within PNCDI III.

Data Availability Statement: The data presented in this study are openly available in Figshare at https://doi.org/10.6084/m9.figshare.26629675.v1 [27].

Conflicts of Interest: The authors declare no conflicts of interest.

References
3. Strandberg, T.E.; Nieminen, T. Future Perspectives on the Role of Frailty in Cardiovascular Diseases; Springer: Cham, Switzerland, 2020; Volume 1216. [CrossRef]
7. Clegg, A.; Young, J.; Iliffe, S.; Rikkert, M.O.; Rockwood, K. Frailty in elderly people. Lancet 2013, 381, 752–762. [CrossRef]
11. Bian, C.; Ye, B.; Mihailidis, A. The Development and Concurrent Validity of a Multi-Sensor-Based Frailty Toolkit for In-Home Frailty Assessment. Sensors 2022, 22, 3552. [CrossRef]
14. Ciubotaru, B.I.; Sasu, G.V.; Goga, N.; Vasilateanu, A.; Popovici, A.F. Architecture of a Non-Intrusive IoT System for Frailty Detection in Older People. Electronics 2023, 12, 2043. [CrossRef]


27. Bogdan-Iulian, C. Dataset with data from a Fitbit Versa, recorded by a healthy individual, 29 years old, performing different activities: Walking, running, climbing stairs and resting. *Figshare J. Contrib.* 2024, 11, 51. [CrossRef]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.