

Research Article

Elderly People Activity Recognition in Smart Grid Monitoring Environment

Anusha Ganesan ¹, Anand Paul ¹ and HyunCheol Seo ²

¹School of Computer Science and Engineering Kyungpook National University, Daegu 41566, Republic of Korea

²School of Architectural, Civil, Environmental and Energy Engineering, Kyungpook National University, Daegu, Republic of Korea

Correspondence should be addressed to HyunCheol Seo; charles@knu.ac.kr

Received 3 February 2022; Accepted 7 March 2022; Published 22 March 2022

Academic Editor: Ravi Samikannu

Copyright © 2022 Anusha Ganesan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Elderly people activity recognition has become a vital necessity in many countries, because most of the elderly people live alone and are vulnerable. Thus, more research to advance in the monitoring systems used to recognize the activities of elderly people is required. Many researchers have proposed different monitoring systems for activity recognition using wired and wireless wearable sensing devices. However, the activity classification accuracy achieved so far should be improved to meet the challenges of more precise activity monitoring. Our study proposes a smart Human Activity Recognition system architecture utilizing an open source dataset generated by wireless, batteryless sensors used by 14 healthy aged persons and unsupervised and supervised machine learning algorithms. In this paper, we also propose using a smart grid for checking regularly the wearable sensing device operational status to address the well-known reliability challenges of these devices, such as wireless charging and data trustworthiness. As the data from the sensing device is very noisy, we employ the K-means++ clustering to identify outliers and use advanced ensemble classification techniques, such as the stacking classifier for which a meta model built using the random forest algorithm gave better results than all other base models considered. We also employ a bagging classifier, which is an ensemble meta-estimator fitting the prediction outputs of the base classifiers and aggregating them to produce the ensemble output. The best classification accuracy of 99.81 was achieved by the stacking classifier in training and 99.78% in testing, respectively. Comparisons for finding the best model were conducted using the recall, F1 score, and precision values.

1. Introduction

Several countries in the world currently have a vast elderly population. This situation entails additional challenges in providing quality healthcare services and facilities for this demographic. Elderly people require more physical and psychological support and assistance. Most of the elderly people currently live on their own as their children work in the different geographic locations. This leads to a lesser likelihood of children taking care of their parents. This makes the elderly parents very vulnerable to risks with no immediate assistance available. However, this problem can be mitigated with the technological advancements in activity monitoring systems for elderly people. Data required by these systems are acquired by wearable sensing devices and then analyzed to

understand and forecast the individual's health condition and required support. These monitoring systems are referred to as Human Activity Recognition (HAR) systems. HAR systems play important roles in several domains such as healthcare, security, and smart environment deployment [1]. The operation of these systems involves five major steps, which are illustrated in Figure 1 for four basic activities such as sleeping, walking, standing, and sitting.

However, to successfully implement these steps, appropriate devices and sensors are required to ensure the efficiency of the entire system. Hence, the development of such system relies on wireless networks, machine learning (ML), data processing, and classification methods. A HAR system can detect and monitor the activities as well as the hazards that can affect the elderly people. Since the devices

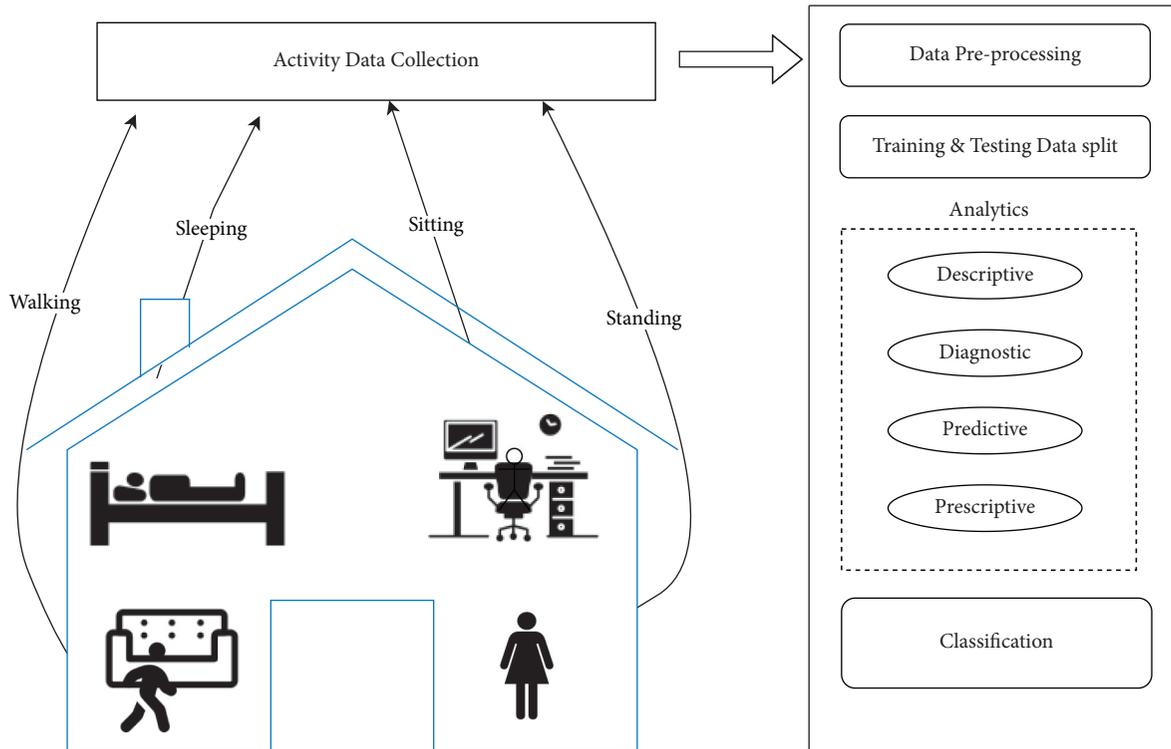


FIGURE 1: Overview of HAR system operation.

connected to users can generate huge amounts of data while monitoring their activities, ML algorithms can help to discover the patterns in the activity logs and make necessary predictions of future trends to assist in adequately supporting elderly persons.

The HAR systems can be considered as a type of Cyber-Physical Systems (CPSs). A CPS is an integrated ensemble of hardware and software components that can run a given process effectively and safely. Nowadays in Smart Homes (SHs), CPSs effectively perform the following tasks:

- (1) Monitoring human activities
- (2) Learning inhabitant preferences or needs
- (3) Providing the required assistance for activity determination, localization, and scheduling

Thus, the HAR systems adopting machine or deep learning offer very efficient health and safety monitoring method, especially for the elderly people. In these systems, data are collected by sensors, wearable devices, and cameras, and transmitted using wireless networks, to be later analyzed to detect the activity changes, behavioral patterns, social interactions, and sleep patterns [2–4]. These types of systems can even help in assessing several health risk factors such as depression, loneliness, and fall risk, as well as unexpected changes in behavior of healthy people. Many researchers have implemented various CPSs for these purposes. Mohsen et al. [5] proposed a TensorFlow-based model for classifying human activities. However, deep learning models require substantial computation power. Hence, other approaches using ML algorithms with good classification and prediction

accuracy have been proposed. Staab et al. [6] examined the performance of several supervised learning algorithms for classifying daily activities of dementia patients using data collected by smartwatches. In another article, Ullrich et al. [7] demonstrated the detection of Parkinson’s disease from inertial sensor data using unsupervised standardized gait tests. Although developing more accurate monitoring models to recognize human activities and detect problems using HAR systems or CPSs is challenging, these challenges must be overcome to enhance the effectiveness of the HAR systems. The existing literature works indicates that HAR systems can be generally based on either sensors or computer vision. In the latter, the system gathers data in form of images or videos with cameras deployed in the monitored environment. Alrashdi et al. [8] proposed a Markov recognition model using maximum entropy utilizing depth camera data. However, their method may infringe on user’s privacy as cameras capture their every motion in videos or images. Hence, the methods which use wearable sensors to gather data are efficient and are strongly preferred. Therefore, in this study we used a dataset gathered using sensors worn by healthy elderly people.

This study proposes a smart HAR system that also tracks the wearable sensing device operational status in addition to the activity data log gathered by the wearable sensors. Because the monitoring process relies exclusively on the data collected by the sensors, checking the sensing device operational state is essential to achieve accurate activity monitoring. Therefore, we propose using a smart grid (SG) in a smart environmental setup to monitor the wearable sensor operational status and report either correctly working

or faults such as faulty, short circuited, good working, or device dead. So, to create a SG monitoring system in a smart environment, we can deploy phasor measurement units (PMUs). These are high speed sensors which can measure voltage and current phasors of the power system with an accuracy in the order of one microsecond [9, 10]. A PMU can help to understand the power status of the device to which it is connected. Using the data collected from PMUs, we can also identify the condition of wearable sensors in a wireless network. However, this study proposes the including such SG elements for sensor state monitoring with a future long-term goal of implementing them. Thus, this study aims to design a smart environment for the proposed system and focuses on developing better computational model for the HAR system. As mentioned earlier, a monitoring system for elderly people must be accurate to predict the required assistance and make prompt decisions to implement necessary actions. In our study, we used the data gathered from 14 healthy elderly people using a batteryless wearable sensing device placed at the sternum level and transmitted via a RFID reader antenna [11–13, 14]. The approach we propose starts with preprocessing of the noisy sensor data. So we use the K-means ++ algorithm to form clusters of data points corresponding to predefined activities, which leaves remaining spurious data in separate, distinct clusters. As the clustering algorithm is used prior to building the classifier, the recognition accuracy is significantly improved and identification of abnormal activity data becomes easier and quicker. The classification model takes the clusters obtained in the previous step as its input. We adopted the ensemble ML algorithms such as the stacking and bagging classifiers at this stage. By analyzing the classification results, we obtain information about the person's activity or the monitoring device condition. Based on the detection outcome, an appropriate alert will be triggered to notify the care taker.

In summary, the key contributions of this paper are as follows:

- (i) Design of a smart HAR system for elderly people physical activity recognition
- (ii) Proposal for using a SG monitoring system to track the wearable device operational status to enhance the reliability of the entire system
- (iii) Evaluation of the efficiencies of the numerical models using ensemble ML algorithms utilizing statistical metrics such as recall, F1 score, and precision values of training and testing accuracies

The rest of the paper is organized as follows. We first provide a short overview of the literature on existing activity recognition systems. In Materials and Methods, we present the materials and methods used to design our proposed system. This section also describes the dataset used in our study. The next section is Proposed Smart HAR System where we explain the design of our system. The experimental setup and results are described and discussed in Results and Discussion. Finally, we conclude the paper with an overview of the future research directions for the development of the proposed system.

2. Literature Survey

This section contains a brief survey of literature works about the HAR system and wearable sensing systems. The review covers the earlier and more recent studies in recognition, classification, and monitoring the human activity patterns. During the current COVID-19 pandemic, many individuals are quarantined in designated places or isolated at home; thus there is an increased need to monitor their physical activities. In this context, Tan et al. [15] used smartphone sensor data and proposed an ensemble learning algorithm comprising a convolutional neural network (CNN) stacked on a gated recurrent unit and deep neural network. Similarly, using smartphone-based accelerometer sensor data, Prasad et al. [16] used a CNN to recognize the six basic human activities of jogging, sitting, standing, and walking up and down the stairs. In fact, in recent years, there have been many studies using the smartphone-based sensor data for activity recognition. In 2020, Sarwar et al. [17] proposed a double-layer approach called physical activity recognition having complex interclass variations (PARCIV). The authors used the smartphone semantic data in their study. Chen et al. [18] analyzed smartphone-based sensor data for HAR using the maximum full a posteriori algorithm. On the other hand, some researchers used the Global Positioning System (GPS) tracking feature available in almost every smartphone along with the sensor data for monitoring the physical activities within indoor and outdoor environments. Selected efforts related to using GPS data for activity recognition are mentioned herein. Wannenburg and Malekian [19] made a prediction model using the kNN and kStar algorithms to recognize the physical activities. Wan and Lin [20] used GPS data collected by a smartphone to classify the human activity patterns utilizing fuzzy logic and aggregation approach. Wu et al. [21] used smartphone motion sensors and kNN classifier to classify physical activities. Kelishomi et al. [22] presented a model for the recognition of the environment type, such as indoor or outdoor, based on the activity data collected using a smartphone. Nan et al. [23] presented a deep learning-based model for elderly people activity recognition using a pocket-worn smartphone. The model was built using a CNN with Long Short-Term Memory algorithm.

Although there have been several studies on recognizing the physical activities using smartphone sensor data, the drawback of such data is that they are likely to have more variation due to dynamic movement of the devices. This can affect the accuracy while monitoring the physical activities of elderly people. However, the CPSs can assist in monitoring of the human activities with wearable sensors and in improving the accuracy of physical activity recognition. The studies reviewed below give an indication of the potential applications and benefits of CPSs for developing efficient monitoring systems for activity recognition. Amarasinghe et al. [24] proposed a framework for data-driven status monitoring of CPS which utilized an additional CPS testbed apart from the key components of the status monitoring system, such as data acquisition, state identification and estimation, cyber-physical status monitoring, and warning generation. The CPS testbed

included a microgrid simulator and cyber network that connected the grid to its controller. The experimental results demonstrated that such a framework can be useful for extracting the relationships between the data features and for highly accurate monitoring of patients or elderly people. Pereda et al. [25] adopted a CPS to perform daily activities in a SH with the help of robots. The paper argued that CPSs can also be used to perform a wide range of activities that makes life easier. Bocicor et al. [26] presented a CPS-based system for indoor ambient monitoring for assisted living, equipped with recognition, monitoring, and alerting features. Bergweiler [27] designed a CPS model for an assistance system based on virtual digit object memories. These types of systems are typically adopted for commercial purposes such as manufacturing of spare parts for automobile and quality assurance for consumer products.

Among the demographics targeted for monitoring their physical activities, the elderly people are a very important group because of the support they may require. There are some studies in which models have been proposed for monitoring their daily activities and selected medical problems, and issuing alerts to the caregivers or hospitals, or healthcare specialists [28]. Lim et al. [29] designed an interactive CPS for assisting elderly people and individuals with disabilities by reminding them about the sequence of their activities and in supporting their daily routine using a smart agent. In the early 2000's, Ohta et al. [30] developed a health monitoring system for elderly people living alone. The monitoring system adopted infrared sensors placed in every room. This early publication demonstrates the need to monitor elderly people health. Chernbumroong et al. [31] proposed an activity recognition and classification model for elderly people. They collected data using wrist worn sensors. However, their classification was at most only 90% accurate. To alleviate the problem, the authors suggested to combine temperature data with acceleration. In a later review paper, Wang et al. [32] presented a detailed account of the wearable sensing technologies for elderly persons to accurately recognize and track their indoor activities and detect their health condition promptly. This review paper is useful for determining which wearable sensors may be suitable for designing a highly accurate HAR system.

It is also necessary to overview the existing HAR algorithms to select an appropriate one for building our model. Sriwichian and Muangprathub [33] adopted different deep learning and ML algorithms for elderly people activity recognition using the R program and compared their results to determine the best performing model (an artificial neural network). Papagiannaki et al. [34] proposed an activity recognition model for aged people using ML algorithms SVM, CNN3 classifiers, etc. While the authors admitted that the accuracy of their model might be lower than that of other approaches, they claimed that low accuracy was because of a single sensor placed at the sternum, and the age of the population from whom the data were collected. However, our own experimental results indicate that better classification results can be achieved even for a single sensor and elder people. We succeeded in developing a more accurate recognition model by applying clustering and ensemble

models to the data. The following section explains the benefits of the methods used in the proposed model along with the materials required.

3. Materials and Methods

This section provides a description of the materials and methods used to design the proposed model.

3.1. Computational Environment. Simulations were executed in the Python environment and Python environment integrated with sklearn libraries for ensemble model libraries to create the classification models and evaluate them using statistical metrics.

3.2. Dataset Description. The dataset used in this study was generated by monitoring 14 healthy elderly people, aged within 66 to 86 in two clinical rooms, while they performed the broadly defined activities using a batteryless, wearable sensor attached to their clothing at the sternum level. Room 1 and Room 2 had four and three RFID reader antennas, respectively, for collecting of activity data. The dataset consisted of 9 features listed in Table 1. There were 75,128 observations in the dataset and 4 classes of the activity labels: (1) sitting on the bed (16,406 instances), (2) sitting on the chair (4,911 instances), (3) lying on the bed (51,520 instances), and (4) ambulating (walking or standing within the room) (2,291 instances). The entire dataset was randomly divided in an 80:20 ratio for the training and testing sets. Hence, the training set had 60,102 samples and the testing set had 15,026 samples, respectively.

3.3. Methods. In this section, firstly, we present a description of the HAR system working and then explain the methods used for designing the proposed system.

3.4. HAR System. HAR systems are typical representatives of pattern recognition systems [35]. They work in two stages: (1) training and (2) recognition. In fact, both stages consist of almost identical steps. The training stage involves gathering prior knowledge of the activities which must be recognized. The recognition stage uses the information collected in the training stage to accurately recognize those activities. In other words, the second stage strongly depends on the success of the first stage. The training stage includes steps such as data collection, feature selection, and learning from the selected features. Data can be collected from different types of sensors employed by the system. However, data are often noisy and must be preprocessed to clean them before further analyses. This preprocessing occurs during the feature selection process. In the feature selection process, the collected data can be classified into structural and statistical features. Structural features are those which indicate correlations between data points. Statistical features are those resulting from the application of statistical methods directly to the data or after they have been transformed, and include the mean, variance, and others. The most widely used

TABLE 1: Dataset features.

Feature number	Description
1	Time in seconds
2	Frontal axis acceleration
3	Vertical axis acceleration
4	Lateral axis acceleration
5	ID of sensor reading antenna
6	Received signal strength indicator (RSSI)
7	Phase
8	Frequency
9	Activity label

transformations are the Fourier and wavelet transforms. Following feature selection, the learning process is executed as the final step in the training phase of a HAR system. In this step, a recognition model is trained from the data gathered by sensors. Several algorithms can be used for creating a best-fit model for recognition. All these three steps are again performed in the recognition stage. At the conclusion of the recognition stage, the most accurate, fastest, and resource-efficient recognition model is chosen. We propose a smart HAR system with the most accurate recognition model for elderly people activity monitoring and prompt assistance. To enhance the effectiveness and reliability of the HAR system in smart environments, we use a SG monitoring system in the proposed system.

3.5. Smart Grid Monitoring System. SGs are electrical power grids, which can be employed for advanced networking, instantaneous monitoring, and controlling to save energy, reducing costs, and enhancing security, interoperability, and reliability [36]. SGs can automatically and spontaneously react to changes in the grid environment. However, they require specific smart sensors to provide real-time monitoring information and status. These sensors are used to measure several physical parameters of the environment and the devices connected to that environment, including energy storage, transmission connectivity, and power generation. Examples of such sensors are PMUs, voltage transformers, current transformers, smart meters, temperature sensors, gas sensors, accelerometers, humidity sensors, and network cameras. However, the sensors must meet certain requirements to be able to perform their intended tasks with efficiency. For instance, the intelligent capabilities required for these sensors are self-identification, self-calibration, self-diagnostics, self-awareness, and localization. Using these sensors, a SG monitoring system can be designed. Such systems monitor the operations of the network and its devices in real time [37]. These systems also aim at lowering the energy consumption and enhancing life quality by saving time and reducing failure rates and stress. The other advantages of SG monitoring systems include monitoring the operational status of the physical devices in the network, providing greater reliability, and increasing the sustainability of the data transmission through clean and energy efficient resources. The core ability of the SG is the ability to communicate the status of the system from the sensors used. Figure 2 shows a schematic diagram of a wireless SG monitoring system. The

SG monitoring system is active when the sensor module present in each SG node collects the voltage and current data from the device connected to it. These data are then sent over the Internet to a server for storage, from where they can be accessed remotely on smart phones. However, data must be interrogated to extract the relevant information about the state of devices connected to the SG system.

Therefore, we propose including this SG monitoring system in the smart environment equipped with specific sensors, such as PMUs, that can support the active monitoring of the system status and obtain the operational data of devices connected to the system. However, only when the quality of the data from a smart environment is enhanced can the end users experience the full benefit of the wearable device. Thus, we employ a clustering technique, the K-means++ algorithm, for data cleaning.

3.6. Clustering. Clustering is an unsupervised technique in which unlabeled input data are divided into groups called clusters based on a certain user-defined criterion. It is a data mining technique, as it partitions the data into clusters based on similarities or patterns common within each cluster. Therefore, data cleaning and outlier removal can be performed by generating a prototype of the data from a clustering algorithm. Many researchers [38, 39] used clustering algorithms for data cleaning to enhance data quality. In our study, we adopted the K-means++ clustering algorithm for data prototyping after removing duplicates from the dataset, which enhanced clustering performance as discussed in the results section. K-means++ algorithm is much similar to the K-means algorithm; however, it works smarter because of the steps followed during the centroid initialization. In the result section, we explain about the algorithmic steps followed by K-means++ to select the centroids. So, in this section, we describe the mathematics behind the K-means++ algorithm for clustering. Like K-means clustering, K-means++ algorithm also uses the centroids to create the clusters. A centroid of m datapoints on an X-Y plane is another point having its own x and y coordinates such that it is a geometric center of the m points. So we have to create K clusters for the given data points based on their Euclidean distance from their centroid. For m data points and $K =$ number of clusters, the centroid initialization is the first step. In K-means++ algorithm, only one centroid is randomly selected from one of the datapoints. Then, for every datapoint x in a set of datapoints given by S , it calculates the distance between every datapoint x and the centroid chosen c using Euclidean distance formula:

$$d(x, c) = \sqrt{\sum_{i=1}^n (c - x_i)^2}. \quad (1)$$

The next cluster centroid is chosen with this distance calculated by selecting the datapoint which is at maximum distance from the first centroid. However, the subsequent centroids are identified with the distance calculated between the datapoint x and the nearest selected centroids. And the weighted probability distribution is used to select the cluster

aggregates the individual predictions to form the final prediction. This meta-classifier can be used to reduce the variance of a black-box estimation methods, such as neural networks or gradient boost methods, by using randomly selected data for establishing separate prediction models and forming their ensemble output. The bagging classifiers have been found to be very effective in various applications [47, 48]. Therefore, we have created stacking and bagging classifier models using various meta-level classifiers and evaluated their performance using several statistical classification metrics. We found that the stacking classifier with the random forest (RF) as the meta-model performed the best as the activity recognition model in our proposed system. An illustration of a stacking ensemble algorithm is given in Figure 4. As shown in the figure, the original dataset is fed as input to the first level learners, which are the base classifier models. Using the outputs of these base models, a new dataset is generated for the second-level learner, which is the meta-classifier model. The aim is to train the base models using the original data and generate new data for training the meta-model so that the risk of overfitting can be avoided. When the same data is used to train the base and meta-models, there can be a high risk of overfitting. Hence, the instances used for generating the new dataset should be excluded from the training data used for the base models, and cross-validation is also recommended.

Therefore, in our experiment, we performed the K-fold cross-validation to train the models. In order to explain this in detail, we consider that given dataset $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$ are randomly split into k equal parts approximately, D_1, D_2, \dots, D_k . D_j and $D_{(-j)} = \{D \setminus D_j\}$ are defined as testing and training sets for the j th fold. Then, suppose M learning algorithms are given; the output of the first-level learning algorithm $h_m^{(-j)}$ is obtained by invoking the m^{th} learning algorithm on $D_{(-j)}$. Then, for every element x_i in D_j , a variable z_{im} is defined to store the output $h_m^{(-j)}$ for every x_i . So, at the end of this cross-validation process, the new dataset for the meta level learning algorithm is generated from the M individual learners as

$$D' = \{(z_{i1}, \dots, z_{iM}, y_i)\}_{i=1}^n. \quad (2)$$

And the resulting output function from the meta level learner is given as h' .

4. Statistical Evaluation Metrics

4.1. Accuracy. Accuracy shows how often the classifier makes correct predictions. It is defined as the ratio of the number of correct predictions to the total number of predictions:

$$\text{accuracy} = \frac{\text{true positive} + \text{true negative}}{\text{true positive} + \text{true negative} + \text{false positive} + \text{false negative}} \times 100\%. \quad (3)$$

However, a very high model accuracy does not always identify the best model; thus metrics such as precision and recall need to be used to judge the model quality in more detail.

4.2. Precision. Precision is a useful indicator when the cost of false positives is high. It is a measure of how accurately the model detects the actual positives compared to the total predicted positives:

$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}. \quad (4)$$

4.3. Recall. A false negative is an occurrence classified as a negative rather than as an actual positive. Recall is the ratio of actual positives to the sum of actual positives and false negatives:

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}. \quad (5)$$

5. Proposed Smart HAR System

The detailed structure of proposed smart HAR system is shown in Figure 5. The system performs the following five major processes:

- (1) Data collection

- (2) Setting up a smart environment
- (3) Data management
- (4) Computational modelling
- (5) Activity recognition or alert notification

5.1. Data Collection. This process is concerned with the generation of the data from the users of the smart HAR system. The data is collected by the wearable sensors. In our study, we used the data gathered from elderly people using wearable batteryless sensor devices because our aim was to propose this smart system for elderly people who need immediate assistance in emergency situations. However, setting up a smarter environment ensures the efficiency of this process and the overall quality of this proposed system which is dependent entirely on the data collected.

5.2. Setting Up a Smart Environment. In this process, a smart environment for data collection from the wearable sensing devices, monitoring and issuing alerts to care takers, is set up. Here, we deployed a SG monitoring system with appropriate sensors, such as PMUs, to monitor the device operational status and the system data transmission status. An activity data log is collected using RFID reader antennas, which is used for recognizing the physical activities. However, use of appropriate sensors and wireless networks assists

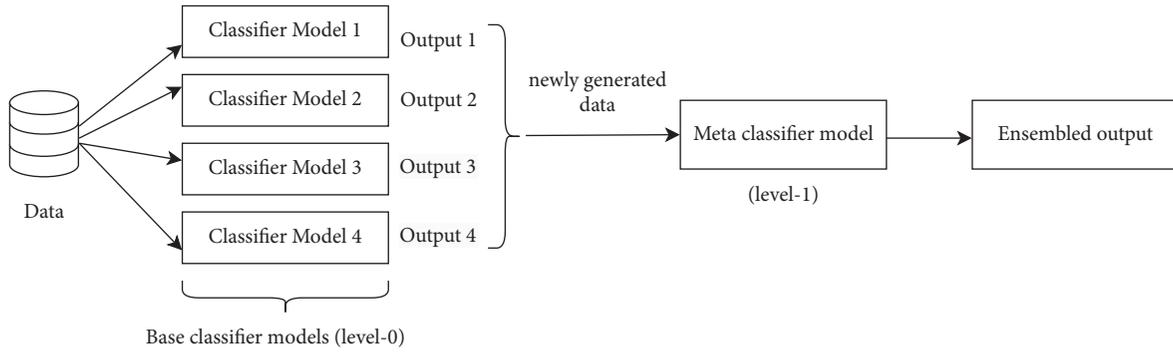


FIGURE 4: Stacking ensemble classifier model.

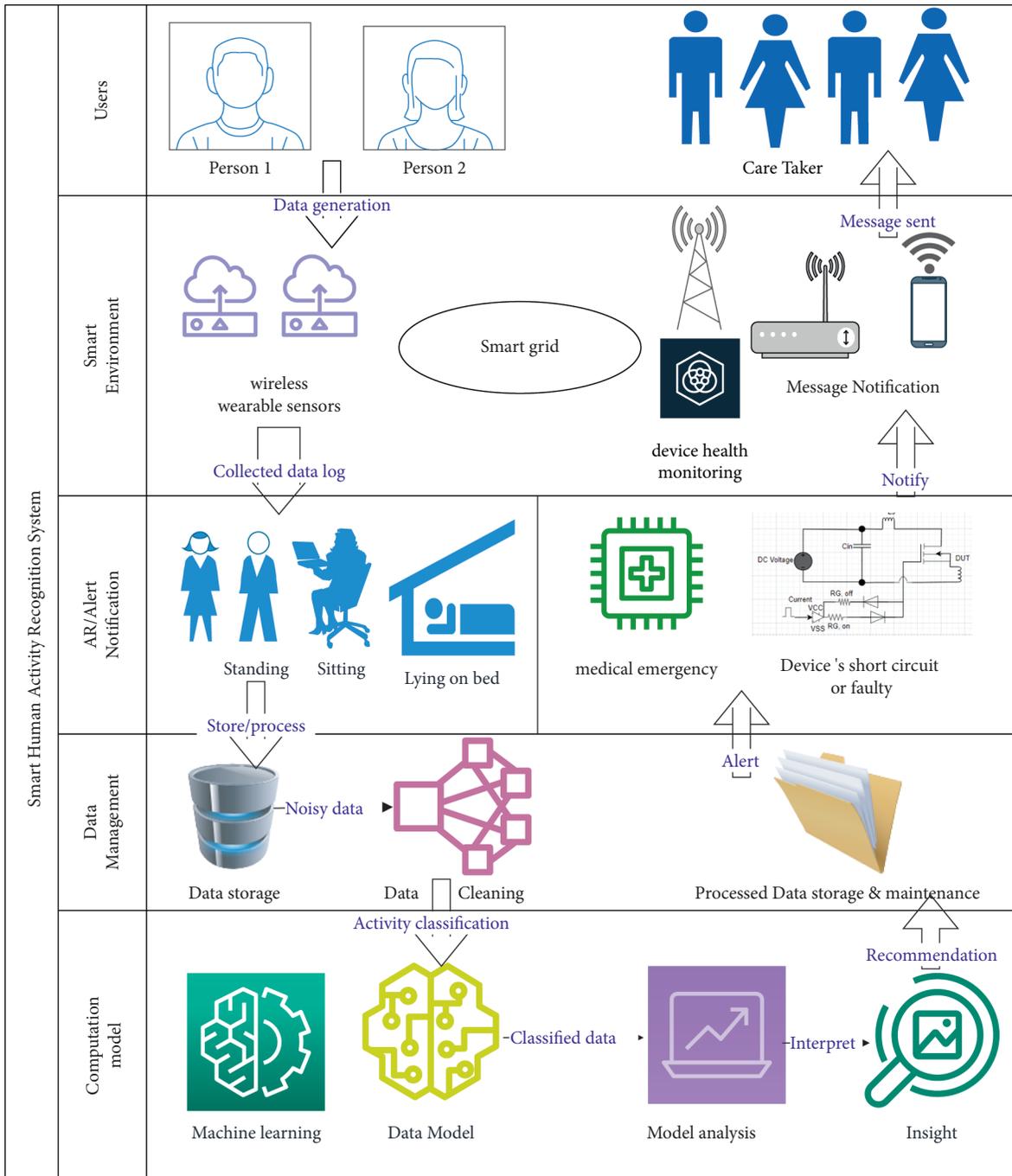


FIGURE 5: Proposed smart HAR system.

in accurate monitoring the physical activities and recording a proper log of data in the cloud or on servers. But, the raw sensor data is likely very noisy; the data management process must also include data cleaning.

5.3. Data Management. There are three main tasks involved in the data management process. Firstly, the data which is collected have to be properly stored for future retrieval. Then, the collected data are likely to require transformation into a suitable format for creating a more accurate recognition model. Later, there is an increased likelihood of noise in the sensor-based data; therefore we have to perform data cleaning. In our study, we used data clustering for data cleaning as it is very efficient. In our experiment, we considered four activities typically performed by elderly people who live alone in their homes. We selected an appropriate dataset and applied the K-means++ clustering algorithm to group the data for these four activities. The clustering algorithm produced four clean clusters, which were fed as inputs into the activity classification model as a part of the HAR system training. Once the model is adequately trained using the ensemble classification method, it can be tested to recognize all the activity data that is generated from the sensors during the HAR system recognition stage. This process also takes care of the management of processed data.

5.4. Computational Modelling. In this process, the clustered output is fed to the ensemble classification model so that it can learn and formulate a hypothesis about the activities to be recognized during the training phase and ensure that those activities are accurately identified during the recognition phase. In our study, we used the stacking and bagging classifiers as the ensemble algorithms. We designed these models using four base models, including the logistic regression (LR), support vector machine (SVM), RF, and decision tree (DT) classifiers, and modified the meta-classifier or evaluation classification model using the outputs of the base models. However, the obtained models must be compared and evaluated to find the best one.

Hence, we compared the models using several statistical metrics. So, whenever there is abnormal activity identified by the model, the corresponding classification results are visualized. And, based on the classification result, the alert notification process is initiated.

5.5. Activity Recognition or Alert Notification. The activity recognition or alert notification is carried out based on the output of the computation model. As mentioned above, we included the SG system for monitoring the sensing device status separate from the sensing system gathering the activity data. Hence, the data collected from the smart environment also include the data related to the condition status of the monitoring devices. Therefore, the classification model can also be trained to classify the condition of the sensors, but this functionality will only be implemented in the future. This device status can be classified as working correctly, faulty, short-circuited, or no power. Thus, the

alert notification is triggered not only in a medical emergency but also when the device operational status is deviating from correct working. Hence, the entire proposed HAR system for monitoring the elderly people physical activities is more reliable by monitoring the wearable sensing devices as well.

6. Results and Discussion

Using the publicly available dataset taken from the UCI repository, we performed an experiment using our proposed system with the hardware and software specifications listed in Table 2. We started the experiment with data cleaning using the K-means++ clustering algorithm for the four labeled activity classes (sitting on the bed, sitting in the chair, lying on the bed, and ambulating (walking and standing in the room)). An important feature of the K-means++ algorithm is intelligent initialization of centroids that produce better clusters. The steps to find such centroids are as follows:

- (1) Select a random centroid point from the given dataset
- (2) For each instance “i” in the dataset, find the distance x from that instance “i” to the closest, earlier chosen centroid
- (3) Select the subsequent centroid from the dataset with the termination goal that the probability of selecting a point as centroid corresponds to the distance from the closest, recently selected centroid
- (4) Repeat Steps 2 and 3 until we find the K centroid points

The output of the K-means ++ clustering algorithm is shown in Figure 6. It produced four distinct clusters of activity labels based on the features such as acceleration, phase, and frequency selected from the given dataset. We used the elbow method to find the optimal number of clusters. The optimal k value for clustering all data samples was determined to be 4.

Next, a classification model was created using the clustered output, which supported improving the accuracy of the recognition results. We used ensemble classification techniques, including stacking and bagging classifiers. The models were created with base models including LR, SVM, RF, and DT classifiers. In this way, we created the following ensemble classification models: stacking-LR (S-LR), stacking-RF (S-RF), stacking-DC (S-DC), bagging-LR (BAG-LR), bagging-RF (BAG-RF), and bagging-DC (BAG-DC). These models were evaluated using several statistical metrics to identify the best classification model for the proposed system. Figures 7 and 8 show comparisons of training and testing accuracies from all the constructed models, from which we determined that the stacking model with RF performed much better than the remaining models.

However, in concluding which classification model is the best, we cannot restrict ourselves exclusively to the accuracy achieved by the models. Therefore, metrics such as precision, recall, and F1 score were also determined for both training and testing stage. These values are listed in Table 3. Higher

TABLE 2: Experimental setup.

Component	Specifications
Processor	Intel(R)Core (TM) i7-3770
Hard disk	Seagate 1 TB
GPU	NVIDIA GeForce GTX 750
SSD	256 GB
Memory	10240 MB DDR3 1333 MHz
Operating system	Windows 10 professional
Computational software	Python 3.7
Interface	PyCharm

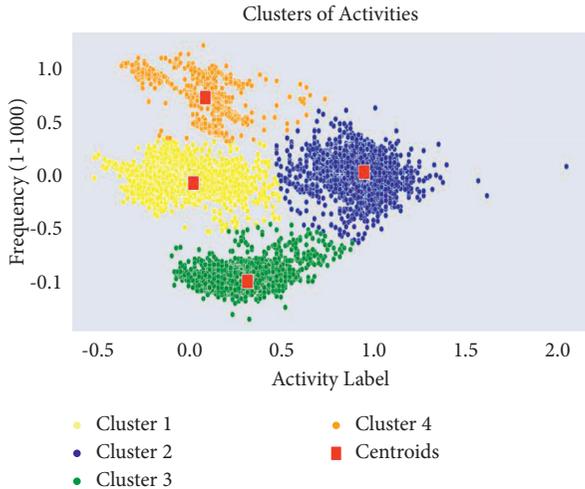


FIGURE 6: Clustering output produced by K-means++ algorithm.

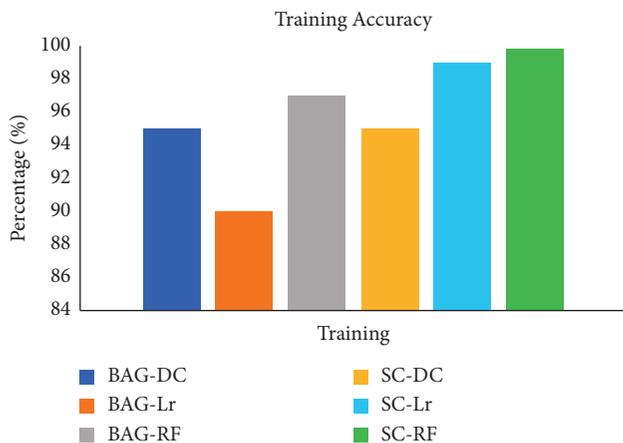


FIGURE 7: Comparison of model accuracy on training data.

values of precision, recall, and F1 score indicate better classification results.

From the values listed in Table 3, we concluded that both stacking and bagging classifiers produced recognition models with good accuracies when RF was used as the meta-model. However, when all the metrics are compared, the stacking-RF model performed much better than the bagging-RF classifier. Thus, we have determined experimentally the best classification model to recognize the activities from the data recorded by the sensors. As a part of the future

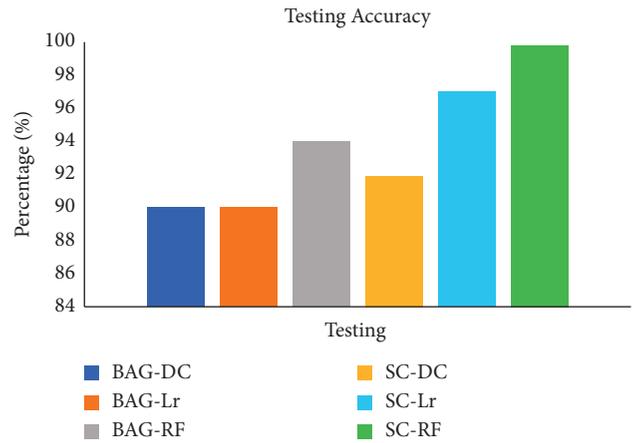


FIGURE 8: Comparison of model accuracy on testing data.

TABLE 3: Comparison of classification metrics for different models.

Model	Training			Testing		
	Precision	Recall	F1 score	Precision	Recall	F1 score
BAG-DC	0.975	0.972	0.973	0.949	0.943	0.946
BAG-Lr	0.947	0.941	0.944	0.947	0.941	0.944
BAG-RF	0.986	0.982	0.984	0.970	0.967	0.969
SC-DC	0.975	0.973	0.974	0.951	0.963	0.957
SC-Lr	0.995	0.995	0.995	0.990	0.980	0.985
SC-RF	0.999	0.996	0.999	0.999	0.999	0.999

research, we will implement the monitoring the smart wearable devices to track their operational status and the associated classification model will be developed.

7. Conclusions

In this paper, we proposed a smart HAR system to identify the physical activities performed by elderly people and a SG monitoring system to track the wearable sensing device operational status for ensuring the data quality and sensing device as well as entire system reliability. A major challenge in the development of the HAR system was the creation of a more accurate activity recognition model. In this study, we used a combination of clustering and ensemble classification algorithms to increase the efficiency of the recognition model and provide accurate classification of predefined and abnormal activities, so that elderly people can be assisted promptly in emergencies.

Data Availability

The numeric data supporting this analysis are from previously reported studies and datasets, which have been cited.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the National Research Foundation of Korea Grant funded by Korean Government (MSIP, South Korea) Number: 2020R1C1C1007127.

References

- [1] J. Muangprathub, A. Sriwichian, A. Wanichsombat, S. Kajornkasirat, P. Nillaor, and V. Boonjing, "A novel elderly tracking system using machine learning to classify signals from mobile and wearable sensors," *International Journal of Environmental Research and Public Health*, vol. 18, pp. 12652–23, 2021.
- [2] D. Bouchabou, S. M. Nguyen, C. Lohr, B. Leduc, and I. Kanellos, "A survey of human activity recognition in smart homes based on iot sensors algorithms: taxonomies, challenges, and opportunities with deep learning," *Sensors*, vol. 21, no. 18, pp. 6037–18, 2021.
- [3] R. Kolkar and V. Geetha, "Human activity recognition in smart home using deep learning techniques," in *Proceedings of the 2021 13th International Conference on Information & Communication Technology and System (ICTS) 2021*, pp. 230–234, Surabaya, Indonesia, October 2021.
- [4] H. D. Mehr and H. Polat, "Human activity recognition in smart home with deep learning approach," in *Proceedings of the 2019 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*, pp. 149–153, Istanbul, Turkey, April 2019.
- [5] S. Mohsen, A. Elkaseer, and S. G. Scholz, "Industry 4.0-oriented deep learning models for human activity recognition," *IEEE Access*, vol. 9, pp. 150508–150521, 2021.
- [6] S. Staab, J. Luderschmidt, and L. Martin, "Recognition of usual similar activities of dementia patients via smartwatches using supervised learning," in *Proceedings of the 2021 IEEE International Conference on Progress in Informatics and Computing (PIC)*, pp. 397–402, Shanghai, China, December 2021.
- [7] M. Ullrich, A. Mucke, A. Kuderle et al., "Detection of unsupervised standardized gait tests from real-world inertial sensor data in Parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 2103–2111, 2021.
- [8] I. Alrashdi, M. H. Siddiqi, Y. Alhwaiti, M. Alruwaili, and M. Azad, "Maximum entropy Markov model for human activity recognition using depth camera," *IEEE Access*, vol. 9, pp. 160635–160645, 2021.
- [9] H. Manoharan, Y. Teekaraman, I. Kirpichnikova, R. Kuppusamy, S. Nikolovski, and H. R. Baghaee, "Smart grid monitoring by wireless sensors using binary logistic regression," *Energies*, vol. 13, no. 15, p. 3974, 2020.
- [10] L. J. Awalim and M. K. Rahmat, "A recent development of monitoring devices on smart grid," in *Proceedings of the E3S Web of Conferences*, vol. 186, pp. 02004–02008, OSAKA, Japan, March 2020.
- [11] R. L. Shinmoto Torres, R. Visvanathan, D. Abbott, K. D. Hill, and D. C. Ranasinghe, "A battery-less and wireless wearable sensor system for identifying bed and chair exits in a pilot trial in hospitalized older people," *PLoS One*, vol. 12, no. 10, pp. e0185670–25, 2017.
- [12] R. Shinmoto Torres, R. Visvanathan, S. Hoskins, A. Van den Hengel, and D. Ranasinghe, "Effectiveness of a batteryless and wireless wearable sensor system for identifying bed and chair exits in healthy older people," *Sensors*, vol. 16, no. 4, pp. 546–4, 2016.
- [13] R. L. Shinmoto Torres, D. C. Ranasinghe, Q. Qinfeng Shi, and A. P. Sample, "Sensor enabled wearable RFID technology for mitigating the risk of falls near beds," in *Proceedings of the 2013 IEEE International Conference on RFID (RFID) 2013*, pp. 191–198, Johor Bahru, Malaysia, September 2013.
- [14] "UCI activity recognition with healthy older people using a batteryless wearable sensor Data Set," 2017, <https://archive.ics.uci.edu/ml/datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor>.
- [15] T.-H. Tan, J.-Y. Wu, S.-H. Liu, and M. Gochoo, "Human activity recognition using an ensemble learning algorithm with smartphone sensor data," *Electronics*, vol. 11, no. 3, pp. 322–417, 2022.
- [16] A. Prasad, A. K. Tyagi, M. M. Althobaiti, A. Almulihi, R. F. Mansour, and A. M. Mahmoud, "Human activity recognition using cell phone-based accelerometer and convolutional neural network," *Applied Sciences*, vol. 11, no. 24, pp. 12099–24, 2021.
- [17] M. Usman Sarwar, A. Rehman Javed, F. Kulsoom, S. Khan, U. Tariq, and A. Kashif Bashir, "PARCIV: recognizing physical activities having complex interclass variations using semantic data of smartphone," *Software: Practice and Experience*, vol. 51, no. 3, pp. 532–549, 2021.
- [18] Z. Chen, C. Jiang, S. Xiang, J. Ding, M. Wu, and X. Li, "Smartphone sensor-based human activity recognition using feature fusion and maximum full a posteriori," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 3992–4001, 2020.
- [19] J. Wannenburg and R. Malekian, "Physical activity recognition from smartphone accelerometer data for user context awareness sensing," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 12, pp. 3142–3149, 2017.
- [20] A. Approach, "Classifying human activity patterns from smartphone collected," vol. 20, no. 6, pp. 869–886, 2017.
- [21] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson, and G. J. Norman, "Classification accuracies of physical activities using smartphone motion sensors," *Journal of Medical Internet Research*, vol. 14, no. 5, pp. e130–9, 2012.
- [22] A. Esmaeili Kelishomi, A. H. S. Garmabaki, M. Bahaghighat, and J. Dong, "Mobile user indoor-outdoor detection through physical daily activities," *Sensors*, vol. 19, no. 3, pp. 511–3, 2019.
- [23] Y. Nan, N. H. Lovell, S. J. Redmond, K. Wang, K. Delbaere, and K. S. van Schooten, "Deep learning for activity recognition in older people using a pocket-worn smartphone," *Sensors*, vol. 20, no. 24, pp. 7195–7214, 2020.
- [24] K. Amarasinghe, C. Wickramasinghe, D. Marino, C. Rieger, and M. Manicl, "Framework for data driven health monitoring of cyber-physical systems," *2018 Resilience Week (RWS)*, vol. 2018, pp. 25–30, 2018.
- [25] C. Pereyda, N. Raghunath, B. Minor, G. Wilson, M. Schmitter-Edgecombe, and D. J. Cook, "Cyber-physical support of daily activities," *ACM Transactions on Cyber-Physical Systems*, vol. 4, no. 2, pp. 1–24, 2020.
- [26] M. I. Bocicor, D. C. Frau, I.-C. Draghici et al., "Cyber-physical system for assisted living and home monitoring," in *Proceedings of the 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP) 2017*, pp. 487–493, Cluj-Napoca, Romania, September 2017.
- [27] S. Bergweiler, "Intelligent manufacturing based on self-monitoring cyber-physical systems," in *Proceedings of the Ninth International Conference of Mobile Ubiquitous*

- Computer System Service Technology*, pp. 108–113, İstanbul, Turkey, July 2015.
- [28] A. Ganesan, A. Paul, G. Nagabushnam, and M. J. J. Gul, “Human-in-the-Loop predictive analytics using statistical learning,” *Journal of Healthcare Engineering*, vol. 2021, pp. 1–14, 2021.
- [29] S. Lim, L. Chung, O. Han, and J.-H. Kim, “An interactive cyber-physical system (CPS) for people with disability and frail elderly people,” in *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication-ICUIMC’11 2011*, Seoul Korea, February 2011.
- [30] S. Ohta, H. Nakamoto, Y. Shinagawa, and T. Tanikawa, “A health monitoring system for elderly people living alone,” *Journal of Telemedicine and Telecare*, vol. 8, no. 3, pp. 151–156, 2002.
- [31] S. Chernbumroong, S. Cang, A. Atkins, and H. Yu, “Elderly activities recognition and classification for applications in assisted living,” *Expert Systems with Applications*, vol. 40, no. 5, pp. 1662–1674, 2013.
- [32] Z. Wang, Z. Yang, and T. Dong, “A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time,” *Sensors*, vol. 17, no. 2, p. 341, 2017.
- [33] A. Sriwichian and J. Muangprathub, “Comparison of algorithm selection to analyze elderly activity recognition based on sensor data using r program,” in *Proceedings of the 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON) 2019*, pp. 377–380, Pattaya, Thailand, July 2019.
- [34] A. Papagiannaki, E. Zacharaki, G. Kalouris et al., “Recognizing physical activity of older people from wearable sensors and inconsistent data,” *Sensors*, vol. 19, no. 4, pp. 880–919, 2019.
- [35] D. Castro, W. Coral, C. Rodriguez, J. Cabra, and J. Colorado, “Wearable-based human activity recognition using an IoT Approach,” *Journal of Sensor and Actuator Networks*, vol. 6, no. 4, pp. 28–4, 2017.
- [36] E. Y. Song, G. J. Fitzpatrick, and K. B. Lee, “Smart sensors and standard-based interoperability in smart grids,” *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7723–7730, 2017.
- [37] G. Luca, M. Benedetta, F. Nardecchia, F. Bisegna, and G. Chiara, “Home smart grid device for energy saves and failure monitoring,” in *Proceedings of the 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC) 2015*, pp. 671–676, Rome, Italy, June 2015.
- [38] S. Joshi, T. Usha, and S. Ata, “Data cleaning using clustering based data mining technique,” vol. 4, no. 2, pp. 2–5, 2015.
- [39] W. Hu, A. Zaveri, H. Qiu, and M. Dumontier, “Cleaning by clustering: methodology for addressing data quality issues in biomedical metadata,” *BMC Bioinformatics*, vol. 18, no. 1, pp. 1–13, 2017.
- [40] P. Yildirim and V. Radevski, “Comparative analysis of ensemble learning methods for signal classification,” pp. 1–4, 2018.
- [41] P. Lohumi, S. Garg, T. P. Singh, and M. Gopal, “Ensemble learning classification for medical diagnosis,” in *Proceedings of the 2020 5th International Conference on Computing, Communication and Security (ICCCS) 2020*, Patna, India, October 2020.
- [42] B. J. Kim, “Ensemble methods applied to classification problem,” *Int. J. Internet, Broadcast*, vol. 11, no. 1, pp. 47–53, 2019.
- [43] E. A. Alshdaifat, M. Al-hassan, and A. Aloqaily, “Effective heterogeneous ensemble classification: an alternative approach for selecting base classifiers,” *ICT Express*, vol. 7, no. 3, pp. 342–349, 2021.
- [44] N. A. B. S. Noor, I. Elamvazuthi, and N. Yahya, “Classification of diabetes mellitus using ensemble algorithms,” in *Proceedings of the 2020 8th International Conference on Intelligent and Advanced Systems (ICIAS)*, pp. 1–6, Kuala Lumpur, Malaysia, July 2021.
- [45] V. O. Khilwani, V. Gondaliya, S. Patel, J. Hemnani, B. Gandhi, and S. K. Bharti, “Diabetes prediction, using stacking classifier,” in *Proceedings of the 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, pp. 1–6, Gandhinagar, India, September 2021.
- [46] P. Ri, A. U. Berliana, and A. Bustamam, “Implementation of stacking ensemble learning for classification of COVID-19 using image dataset CT scan and lung x-ray,” in *Proceedings of the 2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, pp. 148–152, Yogyakarta, Indonesia, November 2021.
- [47] M. S. Bin Alam, M. J. A. Patwary, and M. Hassan, “Birth mode prediction using bagging ensemble classifier: a case study of Bangladesh,” in *Proceedings of the 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD) 2021*, pp. 95–99, Bangladesh, Dhaka, February 2021.
- [48] X. Li, T. Shi, P. Li, and W. Zhou, “Application of bagging ensemble classifier based on genetic algorithm in the text classification of railway fault hazards,” in *Proceedings of the 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD) 2019*, pp. 286–290, Chengdu, China, May 2019.