Heliyon 10 (2024) e29398

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Semi-supervised ensemble learning for human activity recognition in casas Kyoto dataset

Ariza-Colpas Paola Patricia [a], Pacheco-Cuentas Rosberg [a], Shariq Butt-Aziz [b,*], Piñeres-Melo Marlon Alberto [c], Morales-Ortega Roberto-Cesar [a], Urina-Triana Miguel [d], Sumera Naz [e]

[a] Universidad de la Costa, Department of Computer Science and Electronics, Barranquilla, Colombia
[b] School of Systems and Technology, Department of Computer Science, University of Management and Technology, Lahore, Pakistan
[c] Universidad del Norte, Department of Systems Engineering, Barranquilla, Colombia
[d] Universidad Simón Bolívar, Faculty of Health Sciences, Barranquilla, Colombia
[e] Department of Mathematics, Division of Science and Technology, University of Education, Lahore, Pakistan

ARTICLE INFO

Keywords:
Human activity recognition
Activities of daily living
Ensemble learning
Classification methods
Smart home
Clustering
Semi-supervised

ABSTRACT

—The automatic identification of human physical activities, commonly referred to as Human Activity Recognition (HAR), has garnered significant interest and application across various sectors, including entertainment, sports, and notably health. Within the realm of health, a myriad of applications exists, contingent upon the nature of experimentation, the activities under scrutiny, and the methodology employed for data and information acquisition. This diversity opens doors to multifaceted applications, including support for the well-being and safeguarding of elderly individuals afflicted with neurodegenerative diseases, especially in the context of smart homes. Within the existing literature, a multitude of datasets from both indoor and outdoor environments have surfaced, significantly contributing to the activity identification processes. One prominent dataset, the CASAS project developed by Washington State University (WSU) University, encompasses experiments conducted in indoor settings. This dataset facilitates the identification of a range of activities, such as cleaning, cooking, eating, washing hands, and even making phone calls. This article introduces a model founded on the principles of Semi-supervised Ensemble Learning, enabling the harnessing of the potential inherent in distance-based clustering analysis. This technique aids in the identification of distinct clusters, each encapsulating unique activity characteristics. These clusters serve as pivotal inputs for the subsequent classification process, which leverages supervised techniques. The outcomes of this approach exhibit great promise, as evidenced by the quality metrics' analysis, showcasing favorable results compared to the existing state-of-the-art methods. This integrated framework not only contributes to the field of HAR but also holds immense potential for enhancing the capabilities of smart homes and related applications.

* Corresponding author.
E-mail address: Shariq2315@gmail.com (S. Butt-Aziz).

https://doi.org/10.1016/j.heliyon.2024.e29398
Received 22 July 2023; Received in revised form 31 March 2024; Accepted 8 April 2024

1. Introduction

Currently, the Internet of Things (IoT) has enabled the development of various applications that support decision-making processes. One of the sectors where these applications have had a significant impact is the health sector, where it is now possible to carry out remote monitoring of sensitive variables, aiding in the treatment and monitoring of various types of diseases. The pathologies that significantly impactf the global health system include those related to dementia, especially in the case of older adults. Dementia is conceptually defined as a syndrome that, depending on its state (early, intermediate, or advanced), brings side effects such as loss of memory, reasoning ability, behavior, and the ability to perform daily activities.

HAR has established itself as a line of research that has enabled the combination of the Internet of Things using different types of sensors (wearable, object, and environmental) and artificial intelligence, which allow for the monitoring of various activities carried out in both indoor and outdoor environments. The health sector stands out among the wide application sectors, providing a new scenario for experimentation to support the monitoring of adult patients with neurodegenerative diseases, such as dementia. Throughout the investigative processes, different experimentation scenarios have been developed specifically in indoor environments, generating diverse datasets. These datasets allow the incorporation of various techniques and models based on artificial intelligence to predict the different human activities carried out, support the monitoring, and evolution of patients.

The Casas Kyoto dataset was developed by WSU and is part of the Center for Advanced Studies in Adaptive Systems (CASAS) project. It compiled a data repository related to activities of daily living in indoor environments, allowing for the evaluation of an individual's daily activities in an intelligent environment. It contains data representing events from sensors that detect the movement of individuals in the indoor environment. The paper is organized as follows: In Section 2, the Casas Kyoto dataset and its implementation with different sensors are described. Furthermore, in Section 3, we present the related work and applications of supervised and unsupervised techniques in the recognition of human activity. Afterwards, Section 4 describes the adopted methodology, and finally, Section 5 explains the experimentation.

This article aims to contribute to the field of human activity recognition using a semi-supervised ensemble learning approach on the Casas Kyoto dataset [1]. It proposes to develop a system that combines multiple machine-learning models using ensemble techniques to enhance accuracy and robustness in activity identification. These contributions include the design and implementation of an innovative ensemble-based approach, as well as a comprehensive evaluation of performance using standard metrics such as precision, recall, and F1-score. The main objective is to enhance activity recognition capabilities in the context of the Casas Kyoto dataset and provide results that drive advancements in this research area, with potential applications in developing home monitoring systems and improving the quality of life for individuals.

To address concerns about robustness and generalization, this article explicitly details how the proposed models and ensemble techniques improve the system's ability to function effectively under various conditions and with different types of data. The methodology outlined in this work focuses on how the selected and combined models can generalize learnings from one dataset to another, maintaining a high degree of accuracy and reliability in different scenarios. This discussion focuses on the implementation of strategies to mitigate model overtraining, a critical consideration in machine learning models. The emphasis is on ensuring that these models exhibit robust performance not only during training but also demonstrate the ability to adapt and make accurate predictions when presented with new, previously unseen data [2,3]. This approach reflects accurate robustness and generalization in the practice of human activity recognition.

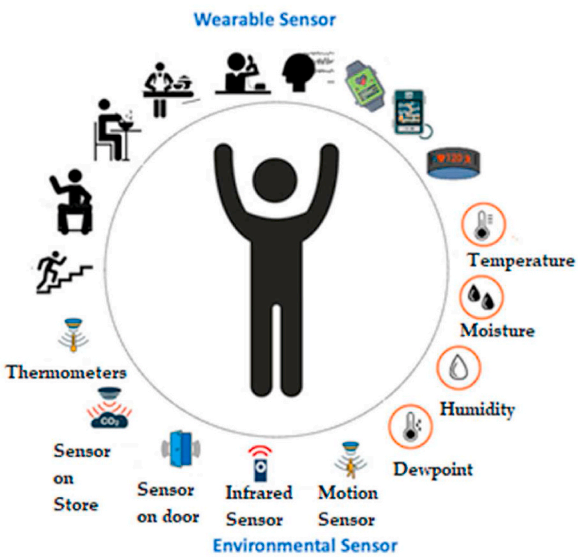

Fig. 1. Sensors in human activity recognition [1].

2. Conceptual information

2.1. Human activity recognition

Systems based on the recognition of human activities make use of both signals and images from various types of sensors found in a given physical space in the indoor environment (room, bathroom, living room, kitchen, closet, etc.) or can also be used by individuals. The cellular technology has become highly relevant due to the amount of data that can be processed and transmitted by smartphones in their continuous use by individuals [4]. The high usability of smartphones in various datasets is supported by the fact that, in addition to their processing capacity, they have become another personal accessory for people due to their portability. They also enable interaction with different applications and the mass sending of information anytime and anywhere [2]. Some other HAR applications focus on the use of integrated cameras in indoor environments that can capture different activities carried out by individuals and sensors located in different parts of the body that denote the activity that is being carried out [3] as mentioned in Fig. 1.

Both in mobile devices and indoor environments, HAR-based systems utilize various types of sensors. Among these sensors, we can highlight positioning, proximity, temperature, accelerometers, gyroscopes, magnetometers, and microphones, among others. One of the sensors widely employed for detecting human activities is inertial sensors, based on the principles of inertia, which dictates the inclination of bodies to maintain their velocity, typically in uniform rectilinear motion. Among the array of sensors conventionally utilized in HAR, the accelerometer stands out, facilitating the measurement of acceleration in m/s^2, alongside the gyroscope, which measures or tracks rotational motion [5]. This set of sensors allows for capturing highly relevant information from the monitoring of both physical and physiological components of people. This has become a field of research for the scientific community that, through data analysis and image processing, strives to identify the different activities that are carried out in both indoor and outdoor environments.

2.2. Casas Kyoto dataset

The Center for Advanced Studies in Adaptive Systems (CASAS) project was developed by Washington State University [1]. This project compiled a repository of data related to activities of daily living in indoor environments. One of the datasets within this repository is Daily Life 2010–2012 Kyoto, which is widely used to evaluate the daily activities of an individual in a smart home (SH). The dataset contains sensor events that detect the movement of individuals in the indoor environment.

These datasets represent sensor events collected in the WSU smart apartment testbed. The data represent participants performing five Activities of Daily Living (ADL). These are routine activities that people tend to do every day as part of their regular self-care and independent living. ADLs are a key measure in assessing an individual's functional status, especially in contexts such as healthcare and gerontology. The ability to perform ADLs is often used to gauge an individual's level of independence and to identify any difficulties they might be experiencing in their daily life. In the apartment, these activities are as follows:

1. Make a phone call. The participant moves to the phone in the dining room looks for a specific number in the phone book, dials the number, and listens to the message.
2. Wash hands. The participant moves into the kitchen sink and washes his/her hands in the sink, using hand soap and drying their hands with a paper towel.
3. Cook. The participant cooks a pot of oatmeal according to the directions given in the phone message.
4. Eat. The participant takes the oatmeal and a medicine container to the dining room and eats the food.
5. Clean. The participant takes all the dishes to the sink and cleans them with water and dish soap in the kitchen.

Among the sensors that were used to carry out the experimentation process are, see Table 1:

3. Related works: applications of supervised and unsupervised techniques in the recognition of human activities

Table .2 summarizes key findings from various studies focused on human activity recognition using different datasets and

Table 1
Summary of sensors used in the dataset.

Sensors Name	Description
M01.M26	Motion Sensors.
I01.I05	Item sensors for oatmeal, raisins, brown sugar, bowl, measuring spoon.
I06	Medicine container sensor
I07	Pot sensor
I08	Phone book sensor
D01	Cabinet sensor
AD1-A	Water sensor
AD1-B	Water sensor
AD1-C	Burner sensor
Asterisk	Phone usage

Table 2
Supervised and unsupervised techniques in the recognition of human activity.

Reference	Summary of methods	Dataset	Results	Identified Weaknesses	Improvements Proposed in This Article
[6]	This paper investigates and compares a range of classical machine learning and deep learning techniques for human activity recognition. Through a comprehensive analysis, various techniques are evaluated to identify the classifier that exhibits the most effective recognition performance. The experimental findings highlight the superiority of the established Deep Neural Network (DNN) model.	UCI Machine Learning dataset.	Accuracy is 75.7 % and 90.84 %.	These studies primarily focus on classic machine learning methods and feature engineering techniques, which could limit their effectiveness in more complex or varied scenarios. Additionally, they do not deeply explore the application of deep learning techniques.	This article suggests the integration of deep learning approaches and a more exhaustive evaluation of advanced feature engineering techniques. This could provide a deeper understanding and more accurate classification of human activities.
[7]	This paper introduces a wearable sensor-based continuous fall monitoring system designed not only to detect fall incidents but also to identify falling patterns and associated activities, addressing the need for comprehensive fall risk prevention. These proposed models employ a two-layer multi-granularity framework and an emergent paradigm with marker-based stigmergy, enhancing context-aware information aggregation and generating a two-dimensional activity pheromone trail.	Various datasets	Varying accuracy (65 %–96.71 %)	Although these studies make advancements in fall detection and sensor-based activity recognition, they may lack generalization to different environments and types of activities.	The current article proposes the use of semi-supervised learning techniques and the exploration of new algorithms that enhance robustness and generalization in various scenarios, including those outside the health context.
[8]	This investigation introduces a hierarchical framework called HierHAR, designed to infer ongoing activities through a multi-stage process. This framework aims to enhance the differentiation of similar activities and overall performance improvement. Notably, It advocates for a data-driven approach, minimizing dependence on extensive prior domain knowledge by automatically determining relationships among activities.	UCI Machine Learning datasets	KNN 95 %, KNN-Decision Tree – 87.01 %	These works propose a hierarchical approach to human activity recognition that can be complex and specific to the dataset, limiting its applicability	It is recommended to simplify the proposed models and optimize the algorithms to allow for a broader and more efficient application in various contexts.
[9]	They created a model for predicting activities like Cleaning, Cooking, Eating, Washing hands, and Phone Calls. The suggested approach introduces a novel method that includes preprocessing and segmenting the dataset through the use of sliding windows. Additionally, carried out experiments with different classifiers to identify the most effective choice for the model.	UCI Machine	KNN-SVM 96.71 % and 82.24 %	It conducts a comparison between classic methods and deep learning techniques but without an effective integration of both approaches.	The current article advocates for an integration of classic methods and deep learning, leveraging the strengths of each approach to improve accuracy and efficiency in the classification of human activities.
[10]	This study introduces various deep learning (DL) models designed for classifying human activities. Specifically, the utilization of long short-term memory (LSTM) is explored for modeling spatiotemporal sequences captured by smart home sensors.	Various datasets	Decision Tree 60 %–96.3 %	It focuses on the use of the Decision Tree algorithm, showing variability in performance across different datasets.	It is proposed to improve the selection and optimization of algorithms to increase consistency and accuracy in activity classification, especially in more varied contexts.
[11]	This paper introduces an open-source, automatic, and highly configurable framework. Its purpose is to establish a baseline for the definition, standardization, and development of Human Activity	MHealth dataset	99.79 % y 99.89 %	The focus on mobile health systems using IoT for human activity recognition might be limited to health scenarios.	Expand the scope of the application to include different contexts and environments, thereby improving the generalization and utility of the

(*continued on next page*)

Table 2 (*continued*)

Reference	Summary of methods	Dataset	Results	Identified Weaknesses	Improvements Proposed in This Article
	Recognition (HAR) methodologies, facilitating the evaluation and comparison of different approaches.				approach in a variety of applications.
[12]	The suggested method aligns with an unsupervised learning approach, offering several benefits. These include streamlining future replication, enhancing control, and deepening understanding of the system's internal mechanisms. The ultimate goal of this system is to simplify the adoption of this approach in a broader range of households.	SisFall dataset	Precision 96.82 %, Precision 79.99 %	Although the study shows high accuracy in fall detection, it primarily focuses on a specific type of activity.	Expand the research to cover a wider range of activities and contexts, thereby improving the system's ability to recognize and classify various human activities.
[13]	In this study, the author introduced a customizable Human Activity Recognition (HAR) system that relies on an affordable and user-friendly Body Area Network (BAN).	Various datasets	Precision 63.27 %–82.79 %	The focus on activity recognition in smart homes shows variability in accuracy across different datasets.	Implement more sophisticated and accurate algorithms, especially in the realm of deep learning, to improve the detection and classification of activities in smart home environments.

methodologies. These studies address the challenges of accurately classifying human activities, presenting results, identifying weaknesses, and proposing improvements. As we delve into the details of each study, it becomes evident that there is a diverse range of datasets and techniques employed. To facilitate a better understanding of the landscape, Table .1 provides a comprehensive overview of the referenced studies, highlighting their contributions and areas for enhancement.

The highlighted studies contribute significantly to the understanding and enhancement of Human Activity Recognition (HAR) methodologies, encompassing classical machine learning and deep learning techniques. However, it is essential to acknowledge certain limitations within the existing literature that the authors aim to address through their proposed solutions. The current body of research often tends to concentrate on specific methodologies or datasets, potentially restricting the generalizability of findings to diverse real-world scenarios. Furthermore, some studies may lack a comprehensive exploration of the application of deep learning techniques, hindering their effectiveness in more complex or varied contexts. Additionally, hierarchical approaches proposed for human activity recognition may demonstrate complexity and specificity to particular datasets, limiting their broader applicability. The authors, in response to these limitations, propose insightful improvements such as the integration of deep learning approaches, exploration of new algorithms, and simplification of hierarchical models. Through these suggestions, the authors aim to not only contribute to overcoming the identified limitations but also advance the field by providing more adaptable, efficient, and widely applicable solutions for accurate human activity classification across diverse scenarios and environments.

4. Methodology

For the development of this experimentation, three important phases are contemplated, which are detailed in Fig. 2.

Phase 1 - Semi-Supervised Dataset: In this phase, the data set is prepared to be included in the experimentation process,

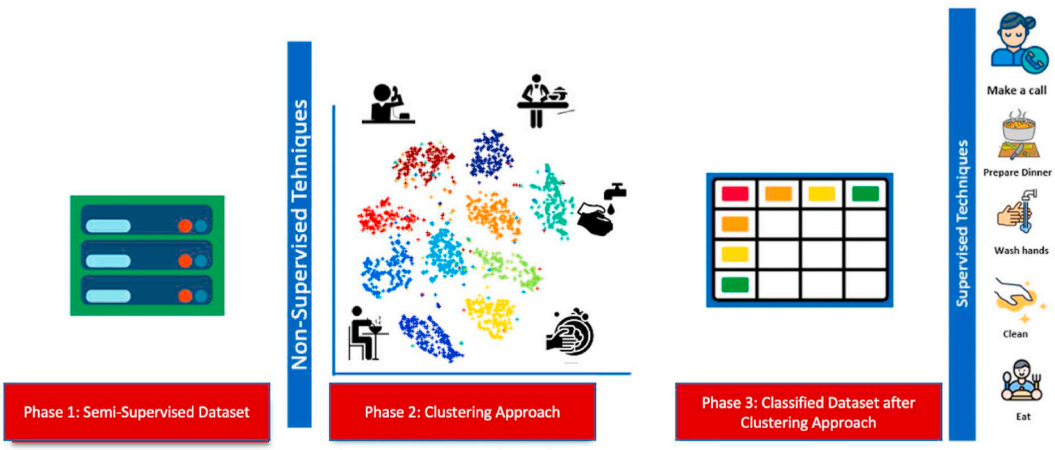

Fig. 2. Ensembled method for human activity recognition.

managing to consolidate a dataset with the following structure. An event-based characterization process was used, considering the application of rank, standard deviation, bias, and kurtosis.

Phase 2 – Clustering Approach: In this phase, three algorithms based on clustering were applied, namely: Fuzzy Clustering, Agglomerative Clustering, and K-means.

Phase 3 – Classification Techniques Approach: Regarding the predictive techniques, two techniques have been used that serve to support the identification process after the cluster: Bagging and J48.

Phase 4- Data Exploration and Preprocessing: Before hyperparameter selection, exploratory data analysis was conducted to understand the characteristics of the dataset. This process aided in identifying initial values for hyperparameters and establishing an appropriate search range.

Hyperparameter selection:

The process of hyperparameter selection was carried out following the typical phases of the knowledge discovery process, which include:

Problem Understanding: In this initial phase, a solid understanding of the human activity recognition problem in the Casas Kyoto dataset was established. Key hyperparameters requiring adjustment were identified, such as the number of estimators in the ensemble, the maximum depth of decision trees, and learning rates.

Feature Engineering: Feature engineering was performed in advance to enhance data representation. The selected or generated features had an impact on hyperparameters, influencing model complexity and the need for regularization.

Model and Hyperparameter Selection: In this phase, model selection was performed, and initial hyperparameters were defined. An initial hyperparameter search was conducted, involving a broad exploration of possible values for each hyperparameter.

Cross-Validation and Hyperparameter Tuning: Cross-validation was employed to assess model performance with different combinations of hyperparameters. This phase involved fine-tuning hyperparameters using performance metrics such as accuracy, recall, and F1 score in multiple iterations.

Evaluation and Deployment: Once the optimal hyperparameters were selected, the model was evaluated on an independent test dataset to obtain a final performance assessment.

Results Interpretation: The results of hyperparameter tuning were interpreted to understand how they influenced the model's performance and how these findings translate into the human activity recognition task.

Table 3
Clustering metrics.

Metric	Formula	Description.
Silhouette Index	$S(C) = \frac{\sum_i sil_i}{n}$ and $sil_i = \frac{(b_i - a_i)}{max(a_i, b_i)}$,	where a_i is the average distance of object x_i to all other objects in the same cluster, and bi is the minimum average distance of object x_i to all objects in other clusters. High values of this index indicate a good clustering structure.
Jaccard Index	$J(C, C^*) = \frac{(z-n)}{\sum_p n_p^2 + \sum_p n_q^{*2} - z - n}$,	where $z = \sum_{pq} n_{pq}^2$ A value near 1 indicates a strong similarity between C^* and C.
Rand Index	$R(C, C^*) = 1 + \frac{2\sum_{pq}\binom{n_{pq}}{2} - \sum_p \binom{n_p}{2} - \sum_q \binom{n_q^*}{2}}{\binom{n}{2}}$	A value near 1 indicates a strong similarity between C^* and C.
Completeness Score	$c = 1 - \frac{H(Y_{pred} \vert Y_{true})}{H(Y_{pred})}$	The objective of this index is to take a sample belonging to the class and observe that it can be assigned to the same cluster. It is determined using the conditional entropy $H(K/C)$, which denotes the uncertainty of determining the correct group associated with a class.
Davies and Bouldin Index.	$DB(C) = \frac{1}{k}\sum_{p=1}^{k} \underset{q \neq p}{max} \left\{ \frac{d_{intra(C_p)} + d_{intra(C_q)}}{d_{inter(C_p, C_q)}} \right\}$	Small values of this index indicate a good clustering structure.
Fowlkes and Mallows Index.	$FM(C, C^*) = \frac{\frac{1}{2}(z-n)}{\left(\sum_p \binom{n_p}{2} \sum_q \binom{n_q^*}{2}\right)^{\frac{1}{2}}}$	where $z = \sum_{pq} n_{pq}^2$ A value near 1 indicates a strong similarity between C^* and C.
Homogeneity Score.	$h = 1 - \frac{H(Y_{true} \vert Y_{pred})}{H(Y_{true})}$	A clustering result satisfies homogeneity if all its clusters contain only data points that are members of a single class.

5. Experimentation

5.1. Phase 1: supervised experimentation

As described in the previous section, in the first phase of the experimentation supervised learning will be used, using algorithms based on conglomerate analysis. Clustering or grouping in machine learning is an unsupervised learning technique in which an algorithm is given to group data sets with similar characteristics. There is, in turn, a set of metrics that can measure the quality of the grouping that results in clustering, among which we highlighted in Table .3:

C and C*: These terms are used in the context of clustering indices such as the Corrected Rand Index (RCC*). A value close to 1 in this index indicates a strong similarity between C and C*. These terms likely represent different groupings or sets of clusters in clustering analysis.

H: In the context of the Completeness Index (c = 1-H(Y_pred|Y_true)/H(Y_pred)), H refers to conditional entropy (HK/C), which denotes the uncertainty of determining the correct group associated with a class. Conditional entropy is used to calculate how complete a clustering is in terms of assigning all samples of a class to the same cluster.

5.1.1. Scenario No.1

Each scenario could involve varying parameters, datasets, or methodologies, providing a comprehensive understanding of how clustering techniques perform in different situations. In summary, scenarios set the stage for experimentation and testing as a clustering technique, and are employed for both testing its capabilities under different conditions and implementing it in real-world applications based on the insights gained from testing. In this scenario, three clustering techniques were applied: AgglomerativeClustering, Fuzzy C-Means (FCM), and KMeans working first on the detection of the ideal number of clusters, resulting in the number 2 as stated in Figs. 3 and 4.

When analyzing the results of the silhouette index, it can be identified that values close to 1 describe that there is a good performance regarding the intercluster relationship (cohesion), but it has difficulties for the extra clustering work (separability). Regarding the analysis of the Davis Bouldin index, it can be analyzed that the resulting clusters are compact, and their centers are well separated from each other as mentioned in Figs. 5–7.

5.1.2. Scenario No.2

In the second scenario, the number of clusters is increased, that is, 5 to carry out a more exhaustive evaluation of different quality metrics through the implementation of the three algorithms defined in the previous scenario: AgglomerativeClustering, FCM, and KMeans. In this experiment, it was possible to extract another group of quality metrics from the cluster grouping as described in Table 4.

When analyzing the results of the clustering quality metrics, the following interpretations can be made. Regarding the Silhouette Index, which identifies good behavior in internal clusters (cohesion) and external clusters (separability), the Agglomerative-Clustering algorithm presents a better result in terms of internal clusters but has difficulty with external clusters: ' separability. As for the FCM and KMeans algorithms, they have better behavior in terms of separability but can improve their results in terms of cohesion. Regarding the Jaccard index, it allows determining the similarity between the different array data, where it can be noted that the K-means algorithm, belonging to the partitional category, has better performance due to its closer proximity to number 1, in contrast to the other two algorithms.

Regarding the Completeness index, it allows to identification of the integrity of the points given that they are members of the same group, in this case again the AgglomerativeClustering algorithm has a better performance compared to the algorithms based on fuzzy and partitional logic. The Davis-Bouldin index determines whether the clusters are compact among themselves, generating a positive separation between the clusters. In the experiment conducted, the FCM-based algorithm performs better in terms of its relationship with the clusters. The Fowlkes-Mallows index is defined as an external evaluation method of clustering that seeks to find similarities between two clusters. In this specific case, the higher the value of the index, the better the result. In this experimental scenario, Agglomerative Clustering shows better results compared to other algorithms such as K-means and FCM.

Regarding homogeneity, this index can identify whether all groups in the cluster belong to the same class. In this specific case, the

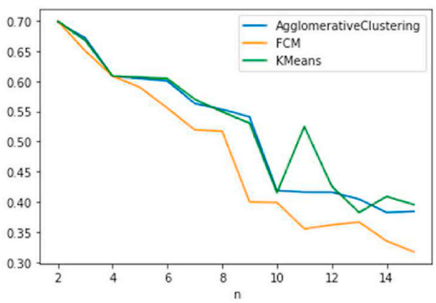

Fig. 3. Clustering analysis scenario.

algorithm	AgglomerativeClustering	KMeans	FCM
n	2.000000	2.000000	2.000000
silhouette	0.697039	0.699389	0.699389
davies bouldin	0.497228	0.492305	0.492305

Fig. 4. Results of clustering metrics of scenario 1.

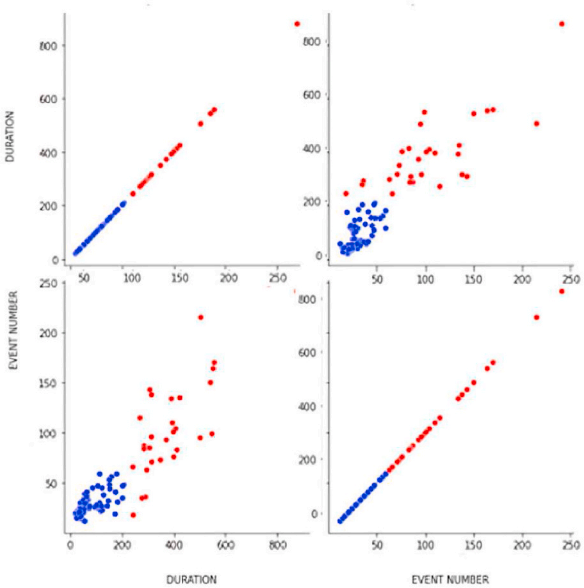

Fig. 5. Agglomerative clustering results.

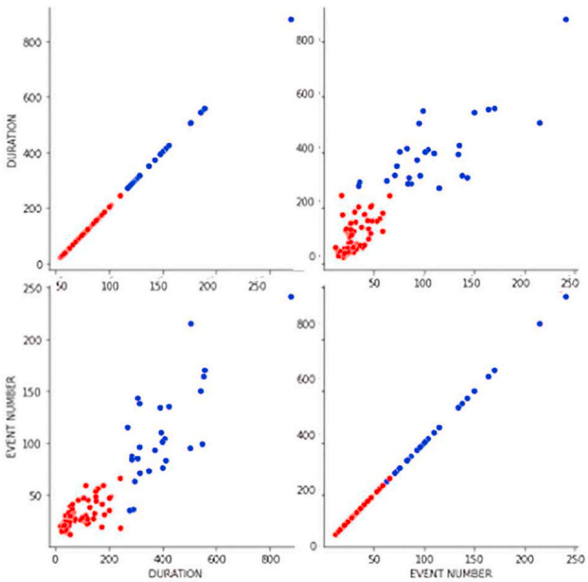

Fig. 6. Fcm results.

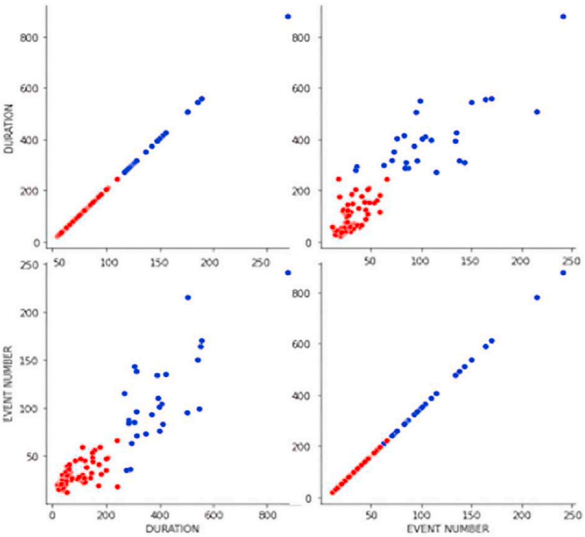

Fig. 7. KMeans results.

Table 4
Results of clustering metrics of scenario 2.

Algorithm	AgglomerativeClustering	K-means	FCM
n	5	5	5
Jaccard	0.604337	0.606676	0.589671
silhouette	0.016053	0.259136	0.021693
Adjusted rand	0.274492	0.2469	0.226937
completeness	0.440904	0.419267	0.386574
Davies Bouldin	0.446221	0.452908	0.603184
Fowlkes mallows	0.460056	0.443775	0.41732
Homogeneity	0.34592	0.323507	0.3284

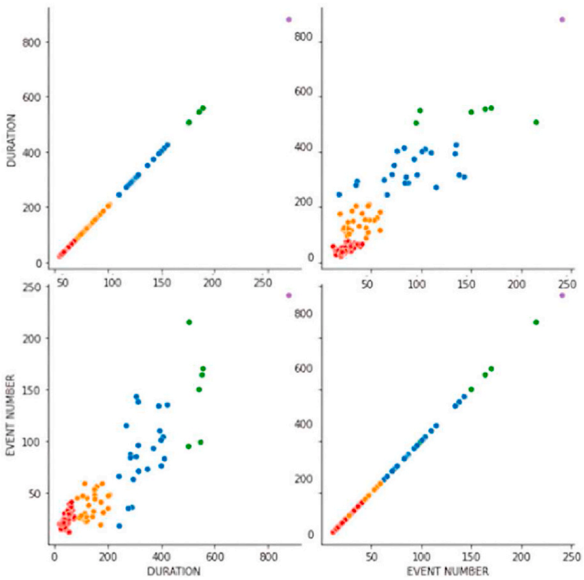

Fig. 8. Results of AgglomerativeClustering of scenario 2.

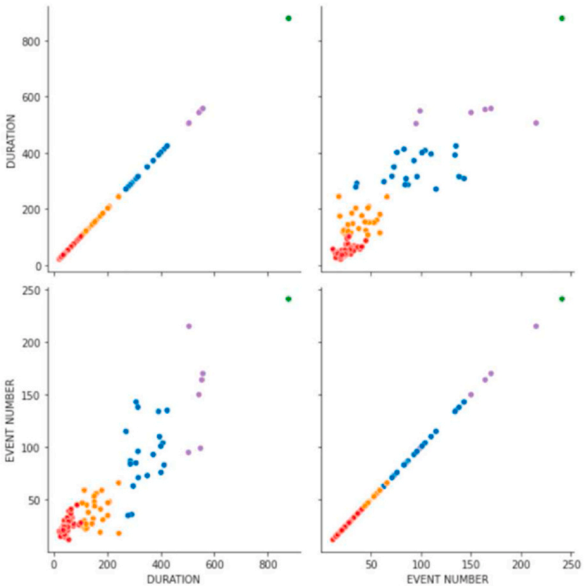

Fig. 9. Results of FCM of scenario 2.

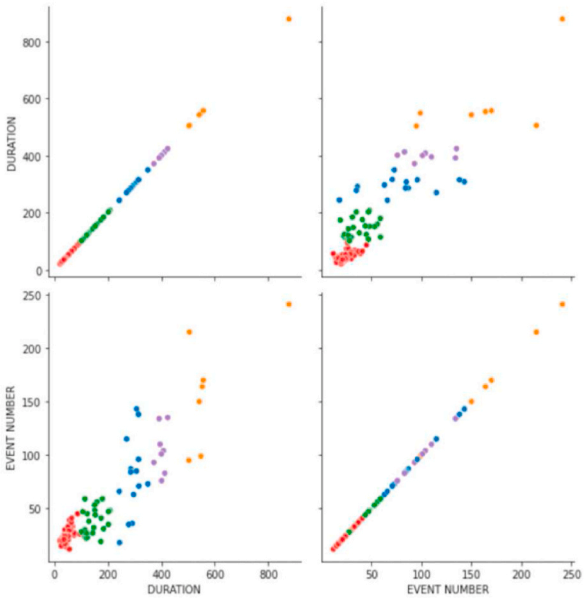

Fig. 10. Results of KMeans of scenario 2.

clustering-based algorithm performs better than the other algorithms, see Figs. 8–10.

5.2. *Phase 2 and 3: clustering and classification approach*

After performing the unsupervised experimentation process based on the two previously defined scenarios, the second scenario was selected as the one that provides the best quality metrics for the formed clusters. With this selection, supervised techniques are implemented using hierarchy-based and partition-based algorithms, which performed the best in terms of quality metrics. After the class analysis process, supervised algorithms such as Bagging and J48 were used.

5.2.1. *AgglomerativeClustering with bagging*

When experimentation is performed after the grouping process, it can be observed that the Bagging algorithm, which is based on

Table 5
Quality metrics of AgglomerativeClustering with bagging.

TP Rate	FP Rate	Precision	Recall	F-measure	MCC	ROC Area	PRC Area	Class
0.955	0.000	1.000	0.955	0.977	0.971	0.999	0.998	0
0.963	0.027	0.929	0.963	0.945	0.925	0.997	0.994	1
1.000	0.011	0.857	1.000	0.923	0.921	0.989	0.735	2
0.977	0.018	0.977	0.977	0.959	0.959	0.999	0.999	3

TP Rate: True Positive Rate.
FP Rate: False Positive Rate.
Precision: Precision.
Recall: Recall.
F-measure: F-Measure.
MCC: Matthews Correlation Coefficient.
ROC Area: Receiver Operating Characteristic Area.
PRC Area: Precision-Recall Curve Area.

the analysis of multiple independent models and is built upon the Agglomerative Clustering algorithm, achieves a classification accuracy of 95.55 % for the classes grouped in the previous stage as like described in Table 5. The acronyms used in Tables are as follows:

In the confusion matrix, it can be identified that there is high cohesion in the identification of each of the classes associated with the prediction process, as can be observed in the main diagonal as stated in Table 6.

5.2.2. *Agglomerative Clustering with J48*

In the experimentation of combining the Agglomerative Clustering algorithm with the J48 tree-based algorithm, the identification processes for each of the classes were identified to show good performance in predicting class 3. As for classes 0 and 1, there were also good performances in prediction. However, there were difficulties in identifying class 2 as like mentioned in Table 7.

When analyzing the results reported by the confusion matrix, it can be identified that there are good results in the process of defining the classes, even though there are difficulties in identifying class 2 as explain in Table 8.

5.2.3. *K-means with bagging*

Concerning the experimentation of partitioning algorithms such as K-means, it can be identified that it has good prediction processes based on partitioning algorithms, although it also has difficulties in classifying class 2, where it could improve the results such as Table 9.

The verification of the confusion matrix denotes good class prediction processes. However, opportunities for improvement are observed regarding class 2 in Table 10.

5.2.4. *K-means with J48*

Concerning the combination of partition-based algorithms and J48 decision tree-based algorithms, good results can be identified. However, better results are obtained using Bagging in Table 11.

In the confusion matrix, good prediction process results for each of the different classes can be identified, showing opportunities for improvement in predicting classes 2 and 4, as stated in Table 12.

6. Discussion

Through this implementation process, the assembled models based on unsupervised and supervised learning provide significant advantages in data processing, as explained in the improvement of predictive processes. We propose evaluating the model by hybridizing selection and classification techniques with the segmented and balanced dataset labeling of activities (not identified) and recognizing activities using a multi-level classifier approach. This approach integrates genetic algorithms for feature selection and Growing Hierarchical Self-Organizing Maps for classification, based on proposals and using classifying multi-classes. One classifier evaluates categories of activities with a low number of instances and a low level of interactions, while another classifier evaluates categories of activities with a high number of instances and a high level of interactions separately. Another future work would be deploying a system that collects individuals' interactions with the indoor environment of UJAmI Smart Lab. After processing the data,

Table 6
Confusion matrix of AgglomerativeClustering with bagging.

a	b	c	d	e	Classified as
21	1	0	0	0	a = 0
0	26	0	1	0	b = 1
0	0	6	0	0	c = 2
0	1	0	43	0	d = 3
0	0	1	0	0	e = 4

Table 7
Quality metrics of AgglomerativeClustering with J48.

TP Rate	FP Rate	Precision	Recall	F-measure	MCC	ROC Area	PRC Area	Class
0.955	0.013	0.955	0.955	0.955	0.942	0.971	0.921	0
0.963	0.027	0.929	0.963	0.945	0.925	0.998	0.904	1
1.000	0.021	0.750	1.000	0.857	0.857	0.984	0.750	2
0.955	0.000	1.000	0.955	0.977	0.960	0.977	0.975	3

Table 8
Confusion matrix of AgglomerativeClustering with J48.

a	b	c	d	e	Classified as
21	0	1	0	0	a = 0
1	26	0	1	0	b = 1
0	0	6	0	0	c = 2
0	2	0	42	0	d = 3
0	0	1	0	0	e = 4

Table 9
Quality metrics of K-means with bagging.

TP Rate	FP Rate	Precision	Recall	F-measure	MCC	ROC Area	PRC Area	Class
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	0
1.000	0.014	0.963	1.000	0.981	0.975	0.999	0.999	1
0.000	0.000	–	0.000	–	–	0.424	0.010	2
0.979	0.000	1.000	0.979	0.989	0.980	1.000	1.000	3
1.000	0.011	0.857	1.000	0.923	0.921	0.989	0.735	4
0.980	0.004	–	0.980	–	–	0.993	0.974	

Table 10
Confusion matrix of K-means with bagging.

a	B	c	d	e	Classified as
20	0	0	0	0	a = 0
0	26	0	0	0	b = 1
0	0	0	0	1	c = 2
0	1	0	46	0	d = 3
0	0	0	0	6	e = 4

Table 11
Quality metrics of K-means with J48.

TP Rate	FP Rate	Precision	Recall	F-measure	MCC	ROC Area	PRC Area	Class
0.950	0.013	0.950	0.950	0.950	0.938	0.969	0.913	0
0.962	0.014	0.962	0.962	0.962	0.948	0.974	0.935	1
0.000	0.000	–	0.000	–	–	0.465	0.010	2
0.979	0.000	1.000	0.979	0.989	0.980	0.989	0.989	3
1.000	0.021	0.750	1.000	0.857	0.857	0.984	0.750	4
0.960	0.007	–	0.960	–	–	0.976	0.935	

Table 12
Quality metrics of K-means with J48.

a	B	c	d	e	Classified as
19	0	0	0	1	a = 0
1	25	0	0	0	b = 1
0	0	0	0	1	c = 2
0	1	0	46	0	d = 3
0	0	0	0	6	e = 4

it predicts in real-time what activity the inhabitant is doing, based on the implementation of the proposed model.

In this study, the semi-supervised ensemble learning approach has been applied for human activity recognition using the Casas Kyoto dataset. Supervised algorithms, such as Bagging and J48, were used, followed by clustering-based algorithms, including AgglomerativeClustering, FCM, and KMeans with 2 and 5 clusters, respectively. The results obtained and the implications of this study will now be discussed. Firstly, the results of the supervised algorithms Bagging and J48 showed promising performance in classifying human activities in the Casas Kyoto dataset. This indicates that the supervised approach is effective for human activity recognition when available labeled data. Bagging and J48 have widely used algorithms in machine learning, and their performance in this study supports their efficacy in the task of activity recognition.

This study demonstrates the effectiveness of applying semi-supervised ensemble learning for human activity recognition using the Casas Kyoto dataset. The combination of supervised and semi-supervised approaches shows promise in improving classification accuracy. Further research could explore other algorithms and configurations, as well as investigate the performance of different datasets and incorporate additional features or contextual information to enhance activity recognition.

7. Conclusions

This study demonstrates the effectiveness of the semi-supervised ensemble learning approach in human activity recognition using the Casas Kyoto dataset. The supervised algorithms Bagging and J48 achieved good classification of activities with labeled data. The clustering-based algorithms AgglomerativeClustering, FCM, and KMeans, with 2 and 5 clusters, proved useful in activity classification with unlabeled data. Combining supervised and semi-supervised approaches enhances human activity recognition. The ensemble approach leverages both labeled and unlabeled data, improving model performance when labeled data is limited.

Regarding the clustering-based algorithms, FCM and KMeans with 5 clusters outperformed AgglomerativeClustering in classification accuracy. Maximizing membership and minimizing distances between data points and cluster centroids are effective approaches for activity classification in the Casas Kyoto dataset. It's important to note that these results are specific to the Casas Kyoto dataset and may vary in other datasets. However, they provide a solid foundation for future research in human activity recognition.

Applying supervised algorithms (Bagging and J48) followed by clustering-based algorithms (AgglomerativeClustering, FCM, and KMeans) on the Casas Kyoto dataset proves effective in human activity recognition. These findings support the importance of exploring hybrid and semi-supervised approaches in developing machine-learning models for human activity recognition in different contexts and datasets. They contribute to advancing the understanding and application of artificial intelligence in the field of human activity recognition.

Data availabitly

Data is available online n at link: https://casas.wsu.edu/datasets/. We have used the dataset Kyoto 20 ADL Activities.

CRediT authorship contribution statement

Ariza-Colpas Paola Patricia: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pacheco-Cuentas Rosberg:** Investigation, Formal analysis, Conceptualization. **Shariq Butt-Aziz:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Piñeres-Melo Marlon Alberto:** Writing – original draft, Validation, Methodology, Data curation. **Morales-Ortega Roberto-Cesar:** Resources, Methodology, Investigation, Formal analysis. **Urina-Triana Miguel:** Writing – original draft, Validation, Methodology. **Sumera Naz:** Writing – review & editing, Validation, Methodology, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Florenc Demrozi, Graziano Pravadelli, Azra Bihorac, Parisa Rashidi, Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey, IEEE Access 8 (2020) 210816–210836.

[2] C.A. Ronao, S.-B. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, Expert Syst. Appl. 59 (Oct. 2016) 235–244, https://doi.org/10.1016/j.eswa.2016.04.032.

[3] V. , A. A., D.J. Gudivada, Data quality considerations for big data and machine learning: going beyond data cleaning and transformations, International Journal on Advances in Software 10 (1) (2017).

[4] I.U.-T.M.A.P.P.M.M. Echeverri-Ocampo, El trabajo colaborativo entre ingenieros y personal de la salud para el desarrollo de proyectos en salud digital: una visión al futuro para lograr tener éxito, Archivos Venezolanos de Farmacología y Terapéutica, 2018.

[5] T., H.P. Cover, Nearest neighbor pattern classification, IEEE Trans. Inf. Theor. 13 (1) (1967) 21–27.

[6] F. Bozkurt, A comparative study on classifying human activities using classical machine and deep learning methods, Arabian J. Sci. Eng. 47 (2) (Feb. 2022) 1507–1521, https://doi.org/10.1007/s13369-021-06008-5.

[7] M. Tasmin, et al., Comparative study of classifiers on human activity recognition by different feature engineering techniques, in: 2020 IEEE 10th International Conference on Intelligent Systems (IS), IEEE, Aug. 2020, pp. 93–101, https://doi.org/10.1109/IS48319.2020.9199934.

[8] F. Hussain, F. Hussain, M. Ehatisham-ul-Haq, M.A. Azam, Activity-aware fall detection and recognition based on wearable sensors, IEEE Sensor. J. 19 (12) (Jun. 2019) 4528–4536, https://doi.org/10.1109/JSEN.2019.2898891.
[9] Z. Comas-González, J. Mardini, S.A. Butt, A. Sanchez-Comas, K. Synnes, A. Joliet, D. Ospino-Balcázar, Sensors and Machine Learning Algorithms for Location and POSTURE Activity Recognition in Smart Environments, Automatic Control and Computer Sciences 58 (1) (2024) 33–42.
[10] A. Wang, S. Zhao, C. Zheng, H. Chen, L. Liu, G. Chen, HierHAR: sensor-based data-driven hierarchical human activity recognition, IEEE Sensor. J. 21 (3) (Feb. 2021) 3353–3365, https://doi.org/10.1109/JSEN.2020.3023860.
[11] K. Maswadi, N.A. Ghani, S. Hamid, M.B. Rasheed, Human activity classification using decision tree and naïve bayes classifiers, Multimed. Tool. Appl. 80 (14) (Jun. 2021) 21709–21726, https://doi.org/10.1007/s11042-020-10447-x.
[12] Raúl Gómez-Ramos, Jaime Duque-Domingo, Eduardo Zalama, Jaime Gómez-García-Bermejo, An unsupervised method to recognise human activity at home using non-intrusive sensors, Electronics 12 (23) (2023) 4772.
[13] A. Subasi, M. Radhwan, R. Kurdi, K. Khateeb, IoT based mobile healthcare system for human activity recognition, in: 2018 15th Learning and Technology Conference (L&T), IEEE, Feb. 2018, pp. 29–34, https://doi.org/10.1109/LT.2018.8368507.