PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Acta Informatica Pragensia
2024, Volume 13, Issue 1, pp. 38–61
https://doi.org/10.18267/j.aip.228

Article Open Access

Dynamic Context-Aware Recommender System for Home Automation Through Synergistic Unsupervised and Supervised Learning Algorithms

Tahar Dilekh, Saber Benharzallah, Ayoub Mokeddem, Saoueb Kerdoudi

Department of Computer Science, University of Batna 2, Batna, Algeria

Corresponding author: Tahar Dilekh (tahar.dilekh@univ-batna2.dz)

Abstract

Home automation, supported by smart devices and the internet of things, works to enhance household control. However, the reliance on current systems with fixed rules poses challenges, which can be inflexible and anxiety-provoking for users who want control over their smart home devices, limit responsiveness to changing conditions and affect energy efficiency, comfort and security. To address this, the paper proposes a dynamic personalized recommender system that considers the user's current state and contextual preferences to suggest relevant automation services for smart home devices. The system uses an unsupervised algorithm to extract rules from past interactions and supervised algorithms to make recommendations based on those rules. The proposed context-aware recommender system for smart homes achieved a remarkable average accuracy of 86.99%, a recall of 76.06% and a precision of 82.67% on publicly available datasets, surpassing previous studies. It offers users an enhanced quality of life, energy efficiency and cost reduction, while providing service providers with increased engagement and valuable insights.

Keywords

Association rule; Generalized linear model; Machine learning; Predictive models; Recommender systems; Context-aware services; Home automation.

Citation: Dilekh, T., Benharzallah, S., Mokeddem, A., & Kerdoudi, S. (2024). Dynamic Context-Aware Recommender System for Home Automation Through Synergistic Unsupervised and Supervised Learning Algorithms. *Acta Informatica Pragensia*, 13(1), 38–61. https://doi.org/10.18267/j.aip.228

Academic Editor: Stanislav Vojir, Prague University of Economics and Business, Czech Republic

1 Introduction

Home automation systems have become increasingly popular in recent years, thanks to the widespread availability of smart devices and the internet of things (IoT) technology. These systems enable users to remotely monitor and control various aspects of their homes, such as lighting, heating, security and entertainment (Aldrich, 2003). This convenience has led to improved energy efficiency, comfort and safety for homeowners, as well as the potential for cost savings and environmental benefits.

However, the current generation of home automation systems has some limitations that prevent them from fully realizing their potential. One of the main challenges is that these systems often rely on pre-programmed schedules or user-defined rules to control devices, which can be inflexible and insufficiently responsive to changing circumstances (Brush et al., 2011; Mennicken & Huang, 2012; Randall, 2003; Woodruff et al., 2007). For instance, a heating system that is set to turn on at a certain time every day may not take into account variations in temperature or occupancy that occur during that time, leading to wasted energy or discomfort for the homeowners. Similarly, a security system that is triggered by a specific sensor may not respond appropriately to different types of intrusions or user preferences.

To address these limitations, researchers have proposed the concept of autonomous or self-controlled home automation, which would allow devices to adapt to the users' behaviour and desires without requiring explicit input from them. However, this approach faces several challenges, particularly in terms of predicting the users' intentions with 100% accuracy and providing a flexible and intuitive interface for users to interact with the system. Many users also express anxiety about losing control over their homes to an automated system (Eggen et al., 2003), which can lead to resistance or distrust towards these systems.

Hence, our primary contribution lies in proposing an innovative personalized recommendation system for smart home residents. This system addresses users' current states and contextual preferences, eliminating the need for manual description of often unconscious daily behaviour (Davidoff et al., 2006). Notably, we introduce a novel context-aware recommender system, integrating supervised and unsupervised learning concurrently, a unique approach in the field. The system captures user behaviour using an unsupervised FP-growth algorithm (Han et al., 2004) to extract behaviour rules from IoT-enabled sensor data. Subsequently, it applies supervised methods, specifically the generalized linear model (GLM) (Nelder & Wedderburn, 1972), to recommend automation services based on these rules, considering the user's current behaviour and context. This pioneering approach strikes a balance between autonomy and control, enabling users to delegate routine tasks while retaining the flexibility to override or modify recommendations as needed.

The proposed dynamic context-aware personalized recommender system in a smart home environment has been evaluated on publicly available datasets (Alshammari et al., 2018) and it was shown to produce highly correct recommendations with 86.99% average accuracy, outperforming the results of studies conducted in this field (Gupta, 2020; Rasch, 2014).

The proposed system has several potential benefits for users and providers of home automation systems. For users, the system can improve their quality of life by reducing the need to manually control devices, as well as by providing more tailored and responsive services. The system can also improve energy efficiency and reduce costs by optimizing device usage based on actual usage patterns and preferences. For providers, the system can increase user engagement and loyalty by offering a more personalized and user-friendly experience, as well as by generating valuable insights into user behaviour and preferences.

Overall, our approach offers a promising direction for the development of home automation systems that can balance the benefits of automation with the need for user control and flexibility. By leveraging the power of machine learning and IoT technologies, we provide a practical and scalable solution for improving the user experience and energy efficiency of smart homes.

The paper is organized into distinct sections for clarity and coherence. Section 2 offers an overview of current research in generating Activities of Daily Living (ADL) datasets, home automation and integrating contextual awareness, as well as the use of unsupervised and supervised learning in smart homes. Moving on, Section 3 introduces the conceptual framework of a dynamic context-aware personalized recommender system within a smart home environment, accompanied by a case study outlining the various layers and processes involved. Section 4 delves into the life cycle of the proposed dynamic personalized recommender system. The procedural details of an experimental case study are comprehensively outlined in Section 5, while Section 6 presents the results and provides an in-depth discussion. Lastly, Section 7 encapsulates the paper with conclusions and thoughtful suggestions for future work.

2 Literature Review

Research in this domain encompasses a spectrum of topics, spanning from the creation of datasets on home automation to the incorporation of contextual awareness and the utilization of unsupervised or supervised learning in smart homes. However, studies that effectively amalgamate these facets, harnessing their strengths and mitigating their shortcomings, as our study does, are exceptionally uncommon. In this section, we summarize some related works (shown in Table 1) touching on the different areas mentioned above.

Table 1. *Survey of works on datasets, home automation, contextual awareness and machine learning in smart homes.*

Work (brief description)	Creation of datasets	Home automation	Incorporation of contextual awareness	Utilization of unsupervised or supervised learning
Rasch (2014): Unsupervised recommendation system for smart homes based on user's previous situation and planned actions. Evaluated on two datasets, achieving 61% and 73% accuracy. Did not address personalization challenge.	Yes	Yes	No	Unsupervised
Belghini et al. (2016): Theoretical framework proposing use of physical sensors for context-aware services in smart homes. Considered user activity, time, location, temperature and events. Aimed to enhance user interaction with the environment.	No	Yes	Yes	Not specified
Engelmann et al. (2016): Presented a dataset for learning automation rules in smart homes using interaction-centric data. Emphasized the importance of interaction data for smart home functionality.	Yes	Yes	No	Not specified
Alshammari et al. (2018): Introduced two datasets for smart home studies, one for classification and the other for anomaly instances. Created using the OpenSHS simulation tool with data from seven individuals. Spanned 63 sample days.	Yes	Yes	No	Not specified
Reyes-Campos et al. (2021): Introduced a platform using historical records of home automation device usage to detect resident behaviour patterns. Employed C4.5 algorithm for data classification. Identified behaviour patterns among smart home residents.	No	Yes	Yes	Supervised (C4.5 algorithm)

Work (brief description)	Creation of datasets	Home automation	Incorporation of contextual awareness	Utilization of unsupervised or supervised learning
Gupta (2020): Proposed a context-aware recommender system for smart home automation using unsupervised algorithms to discover user preferences and supervised algorithms to predict user behaviour. Extracted contextual features from datasets.	Yes	Yes	Yes	Unsupervised and supervised (Apriori, deep learning, GLM, gradient boosted tree)
Ortiz-Barrios et al. (2022): Proposed a method for predicting duration of activities performed by dementia patients in smart sensing environments. Used synthetic data and partial least squares regression. Evaluated on a real smart home dataset.	Yes	Yes	Yes	Supervised (partial least squares regression)
Ali et al. (2023): Introduced a methodology for personalizing smart homes based on user preferences and needs. Involved user profiling, system configuration and system adaptation. Evaluated through a case study.	No	Yes	Yes	Not specified
Belaidouni et al. (2022): Proposed a hybrid approach using reinforcement learning (Q-learning) and case-based reasoning (CBR) for context-aware services in smart environments. Addressed computational complexity of RL by combining it with CBR.	No	Yes	Yes	Hybrid (reinforcement learning and case-based reasoning)
Osman et al. (2023): Developed an automatic control system for a smart building aimed at optimizing energy consumption and enhancing user comfort. Implemented the system in a real smart building environment.	No	Yes	No	Not specified
Miraoui (2018): Focused on development and implementation of a context-aware smart office system. Aimed to improve the comfort of office occupants and achieve energy savings by dynamically adjusting the office environment.	No	Yes	Yes	Not specified
Leonidis et al. (2019): Explored design and implementation of intelligent systems to enhance living room experience. Integrated ambient intelligence technologies to create an intelligent and responsive living room environment.	No	Yes	Yes	Not specified
Ramesh & Kannimuthu (2023): Proposed a context-aware practice problem recommender system for POJs based on learners' skill level navigation patterns. Integrated content-based, CF and SPM approaches. Outperformed traditional systems in accuracy.	Yes	No	Yes	Not specified

3 Conceptual Framework

The conceptual framework of the proposed recommender system comprises four layers and three interrelated processes. These layers and processes collaborate seamlessly to handle raw sensor data and produce recommendations, as depicted in Figure 1.

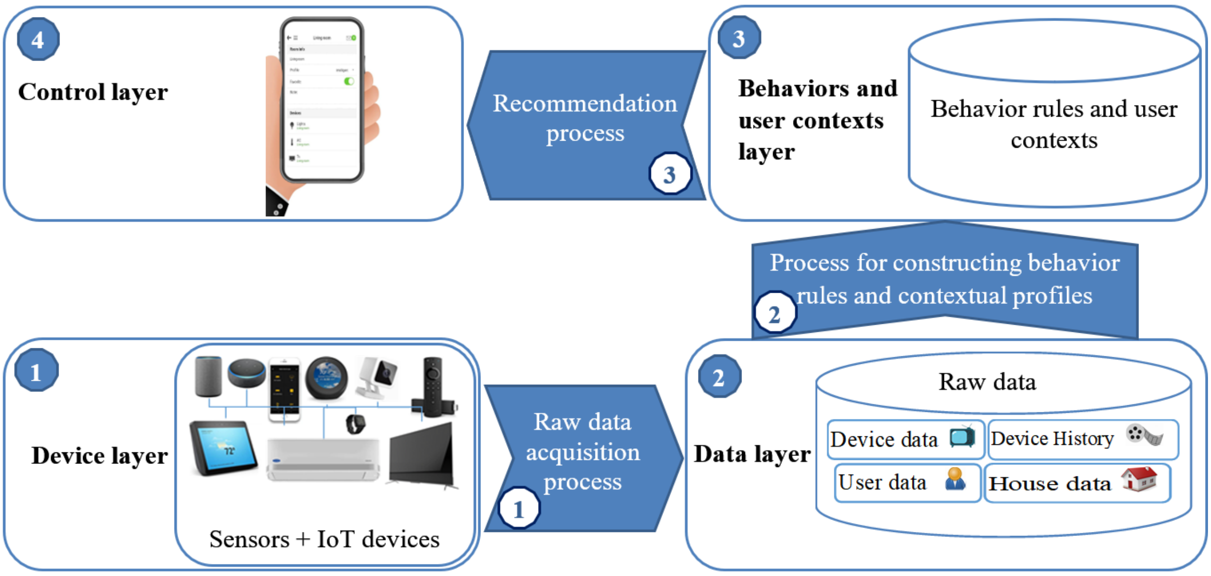


***Figure 1.** Conceptual framework of dynamic context-aware personalized recommender system in smart home environment.*

In the following section, we will describe the functional framework of our system and the role of each phase in the recommendation process.

3.1 Four layers

3.1.1 Device layer

The first layer of our system comprises communication tools and technologies required to manage and supervise home automation devices. This layer includes sensors that detect and capture events in diverse environmental conditions such as temperature, light, population movement and others in smart home systems. Additionally, this layer consists of IoT devices that employ IoT technology to connect with home automation systems.

3.1.2 Data layer

This layer contains valuable information for both smart home control and the various modules of the system. It encompasses four main types of data: (1) device data, which pertain to the stored information of each device connected to the system; (2) device history, containing historical data on all status changes. These data are crucial for analysing and identifying people's behaviour patterns and are used in the modelling phase; (3) user data, which are the stored information of the smart home inhabitants; and (4) house data, which represent the stored data of each room in the house. The raw data generated from this layer serve as the output dataset.

3.1.3 Behaviour and user context layer

This layer is produced through the application of an unsupervised algorithm. This algorithm serves as an analytical process, which identifies frequent patterns or associations from raw data with user contexts and determines the behaviour rules. These rules are then used to generate appropriate recommendations during a later phase.

3.1.4 Control layer

The control layer is the output of the modelling process, where a model is trained using supervised learning algorithms. This model predicts and recommends the actions that the inhabitant should take in a given context, based on the learned rules. The recommender system can then be implemented to highlight interesting choices in centralized interfaces, such as smart wearable devices. This allows users to trigger desired actions through voice, touch or gestures, making the smart home ecosystem more convenient.

3.2 Three processes

3.2.1 Raw data acquisition process

This process involves collecting raw data from the user's physical context and interactions with home automation devices. The information is then classified based on sensor types and user interactions. This enables the detection of user behaviour in the subsequent process.

3.2.2 Process for constructing behaviour rules and contextual profiles

The process of constructing behaviour rules involves analysing the raw data that represent the user's interactions with home automation devices. This analysis allows us to recognize user preferences and shape them into behaviour rules. Simultaneously, the contextual profiling process enables us to determine the user's context based on the raw data and understand their current situation. Since user preferences are influenced by context, these profiles are context-dependent. For instance, a user may prefer listening to the radio instead of watching TV on weekends, highlighting the close relationship between user preferences and context. Therefore, the context and profile have a significant impact on the behaviour rule construction process.

To create these behaviour rules, an unsupervised algorithm is applied to the user interaction preferences. Then, these preferences are combined with contextual information.

3.2.3 Recommendation process

The process involves using a supervised learning algorithm to train a model that predicts and recommends resident behaviour in a given contextual setting. By considering behaviour rules and user context, this process enables the system to generate more accurate patterns to apply in a specific context. The recommendations can then be implemented in centralized interfaces such as smart wearable devices to highlight interesting choices for the user.

4 Life Cycle

The life cycle of the proposed personalized recommender system can be defined by a cycle diagram that has various states representing the execution status of the process at various times and transitions. The workflow is shown in Figure 2. Initially, there is a need for constant access to the local devices of the smart home, involving IoT technology devices, to obtain interesting information, as well as data corresponding to the history of the device use.

Then, the data received from the device are stored and analysed using the FP-Growth algorithm, which is commonly used for frequent pattern miningand is used as an analytical process to find frequent patterns or associations in datasets. This analysis allows us to recognize user preferences and shape them into behaviour rules.

The recommendation phase begins by processing the raw data to extract important contextual information from the dataset and adding it to the rules identified in the previous phase. Contextual information is additional important knowledge that extends the meaning of the data as a whole (Rasch, 2014; Verkasalo & Karjalainen, 2010). Therefore, it is important to incorporate this knowledge to gain an edge in decision-making, which means recommending more accurate automated services to inhabitants.

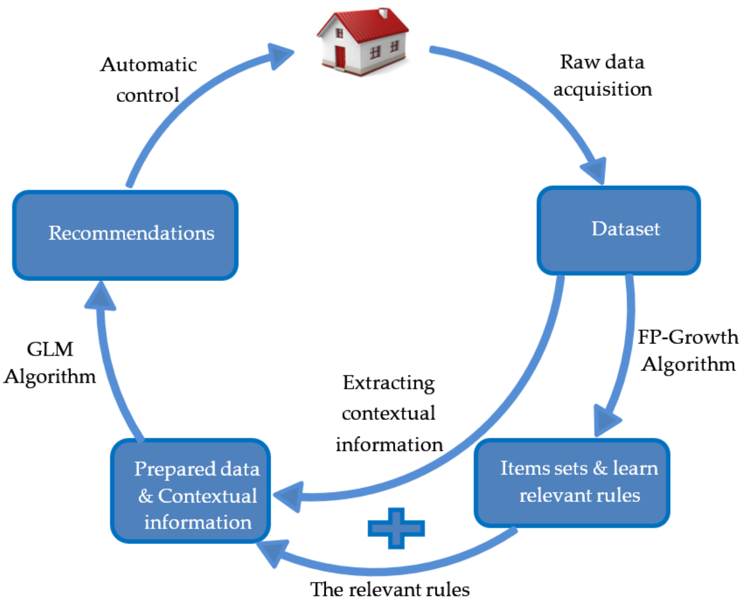


Figure 2. *Life cycle of proposed dynamic personalized recommender system.*

After the usage patterns have been discovered, a supervised learning algorithm GLM is applied to predict user behaviour in a given contextual environment and provide recommendations that serve as a base for constructing custom comfort setting rules for each specific home's history. As a result, depending on the circumstances of the occupants of each smart home, it is controlled according to perfectly proportioned rules.

Finally, the system will send the appropriate recommendations to residents of the house and implicitly to the devices, making smart homes more convenient in terms of device control.

5 Experimental Case Study

We propose the general algorithm mentioned below for the experimental case study, as this general algorithm provides a framework for building a context-aware personalized recommendation system for home automation, which includes both unsupervised learning for behaviour rule extraction and supervised learning for real-time recommendation generation. Specific implementation details and algorithm parameters may vary based on smart home environment requirements and data characteristics.

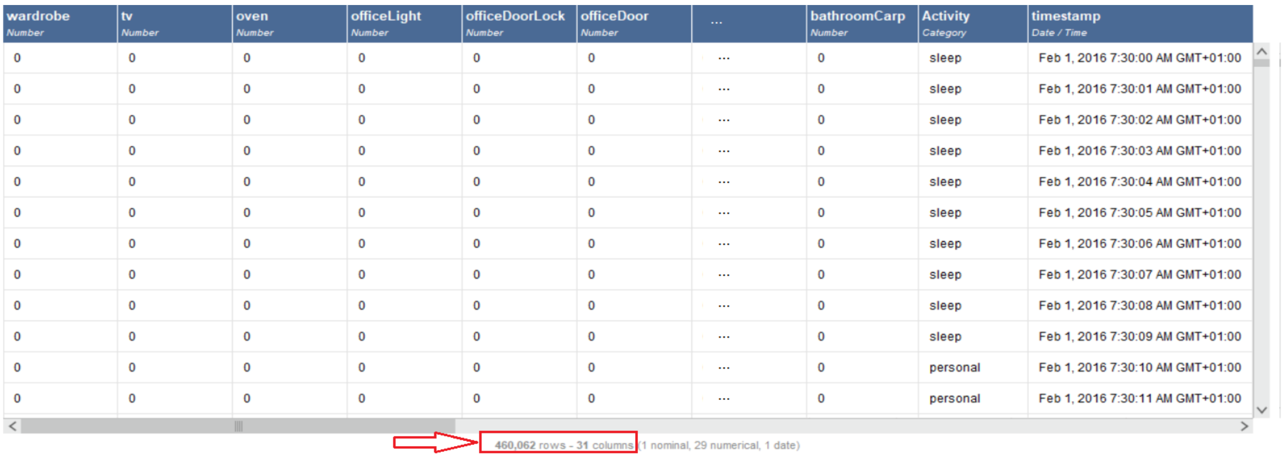
```
Algorithm Context-aware personalized recommendation system
Input:
Historical data on user interactions with smart home devices.
Current contextual information (time, occupancy, temperature, etc.)
Output:
Personalized recommendations for automation services
1:   Unsupervised phase (behaviour rule extraction)
2:        Collect and preprocess historical data on user interactions with
          smart home devices.
3:        Apply the FP-Growth algorithm to discover closed frequent itemsets
          and maximal frequent itemsets.
4:        Extract association rules from the frequent itemsets, representing
          behavioural patterns.
5:        Build a behavioural model based on the learned association rules.
```

```
6:   Context-aware supervised phase (recommendation)
7:         Acquire real-time contextual information (time of day, occupancy,
           ambient temperature, etc.).
8:         Combine real-time context with the historical behavioural model.
9:         Implement a closed frequent itemset maximal strategy to filter
           relevant rules based on the current context.
10:        Apply the GLM (generalized linear model) algorithm to learn the
           relationship between contextual features and recommended automation
           services.
11:        Generate personalized recommendations using the learned model.
12:  End.
```

Moreover, we commence our algorithm under the assumption of having the initial dataset, progressing directly to the subsequent phases, specifically the unsupervised and supervised stages. Nevertheless, a question persists: from where was the initial dataset obtained, and what steps are involved in its preparation for the subsequent unsupervised and supervised learning phases? What are the details of the unsupervised and supervised learning stages mentioned in the algorithm? These aspects will be addressed in the upcoming sections.

5.1 Dataset

The datasets selected for this study were provided by Alshammari et al. (2018), who aimed to generate realistic datasets using simulation tools (OpenSHS). In the simulation, seven participants performed their ADL over a period of 63 days. The simulation tool recorded the state of sensors placed in various parts of the house, such as the bedroom, kitchen, living room, bathroom, office and hallway passage, as shown in Table 2. The smart home was equipped with 29 binary sensors placed on various objects, which indicated the state of the object on which they were installed.



wardrobe Number	tv Number	oven Number	officeLight Number	officeDoorLock Number	officeDoor Number	...	bathroomCarp Number	Activity Category	timestamp Date / Time
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:00 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:01 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:02 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:03 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:04 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:05 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:06 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:07 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:08 AM GMT+01:00
0	0	0	0	0	0	...	0	sleep	Feb 1, 2016 7:30:09 AM GMT+01:00
0	0	0	0	0	0	...	0	personal	Feb 1, 2016 7:30:10 AM GMT+01:00
0	0	0	0	0	0	...	0	personal	Feb 1, 2016 7:30:11 AM GMT+01:00

460,062 rows - 31 columns (1 nominal, 29 numerical, 1 date)

Figure 3. *Raw data snippet from seventh participant's simulation.*

The study involved seven randomly selected participants, who had prior experience with first-person games, facilitating their use of the simulation tool. Each participant performed a simulation that was intended to represent the participant's ADL. Figure 3 shows a raw data snippet from the seventh participant's simulation with a total record of 460,062 rows in 31 columns.

The subsequent section outlines the proposed approach for constructing a system that acquires knowledge of inhabitant behaviour during the initial unsupervised phase, integrates this with contextual information during the subsequent supervised phase, and ultimately provides recommendations for pertinent automation services.

Table 2. *All smart home sensors.*

#	Name	Active/ passive	Description	Type
1	bathroomCarp, bathroomDoor, bathroomDoorLock, bedTableLamp, bedroomDoor, bedroomDoorLock, kitchenDoor, kitchenDoorLock, mainDoor, mainDoorLock, officeDoor, officeDoorLock, oven, tv, wardrobe	Active	Sensor	Binary
2	bedroomCarp, bed, kitchenCarp, livingCarp, office, officeCarp	Passive		
3	bathroomLight, bedroomLight, hallwayLight, kitchenLight, livingLight, officeLight	Active	Light	
4	Fridge	Active	Kitchen fridge	
5	Couch	Passive	Living room couch	-
6	Activity	-	Current participant activity	String
7	Timestamp	-	Timestamp every second	-

5.2 Unsupervised phase (learning)

After loading the dataset and confirming that the required data format (one-hot encoded transaction format converts nominal categorical data into features with binary values, where True represents the presence of an item and False represents the absence of the item) matches our needs, we proceed to discover the user's frequent behaviour and learn association rules through the use of an association analysis algorithm known as FP-Growth. However, what exactly are association analysis algorithms and what is the FP-Growth algorithm specifically?

5.2.1 Algorithms of association rule analysis

Association rules measure the level of correlation between two events or items (Han et al., 2004). These algorithms take a list of previous transactions as input and provide a rule in the form of a tuple of items that frequently appear together as output. In other words, if event A happens, event B will likely also occur. The main objective of any association rule algorithm is to identify frequently occurring events from all potential events.

The following example concerns common events that occur simultaneously:

{Couch, TV} → {Lights Dim}, which means the lights are dimmed while watching TV from the couch.

These events occur frequently and suggest that these are the resident's preferences in certain contexts.

The following concepts and techniques are described below to measure the relationship between items.

- **Definition 1 (itemset)**

A collection of one or more items or events, for example: {Oven, Kitchen Cupboard}, {Radio}.

- **Definition 2 (support)**

This indicates the popularity of the itemset as measured by the percentage of transactions in which the itemset appears. If the occurrence of some items in a transaction above a certain percentage is found to significantly affect the occurrence of other items, then that percentage should be used as the support threshold. An itemset with support values above this threshold can be identified as an important itemset.

- **Definition 3 (frequent itemset)**

An itemset with support greater than or equal to the specified minimum support threshold.

- **Definition 4 (association rule)**

An implicit expression of the form X → Y (where X and Y are itemset), for example: {Couch, TV} → {Lights Dim}.

- **Definition 5 (premises and conclusion)**

These are sets of itemsets that are related to each other by association rules. In a rule X → Y, X is the premises and Y is the conclusion.

- **Definition 6 (confidence)**

It indicates the likelihood that all items of Y will occur if all items of X occur, denoted by {X → Y}. This is measured by the proportion of transactions with items of X, in which items of Y also appear. Confidence gives the reliability of the rule.

$$Confidence(X \to Y) = \frac{Support(X \cup Y)}{Support(X)} \tag{1}$$

- **Definition 7 (lift)**

In the lift, the confidence ignores the frequency of the conclusion Y, and if the value is closer to the value of one (1), the premises and conclusion are independent and the rule is not interesting. Otherwise, lift values higher than 1 means the rule is interesting. The lift may find very strong associations for less frequent items.

$$Lift(X \to Y) = \frac{Support(X \cup Y)}{Support(X) \times Support(Y)} \tag{2}$$

- **Definition 8 (conviction)**

Conviction is the ratio of the expected frequency of a premise X occurring despite a conclusion Y to the observed frequency of an incorrect prediction.

5.2.2 FP-Growth algorithm

Han et al. (2004) introduced a novel method, distinct from level-based approaches, to extract frequent itemsets from databases without candidate generation, called FP-Growth (frequent pattern growth). It involves compressing the database into an FP-tree (frequent pattern tree) structure and segmenting it into conditional bases. Each base is linked to a frequent item, enabling independent extraction of frequent itemset. FP-Growth offers a solution to mining frequent patterns in sizable transactional databases. Storing frequent elements compactly eliminates the need for repetitive scans, while sorting them expedites pattern search. The FP-Growth method demonstrates efficiency and scalability in mining both long and short frequent patterns, notably outperforming the Apriori algorithm and other recent frequent-pattern mining techniques in terms of processing speed.

FP-growth pseudo-code

The various steps of the FP-growth algorithm can be summarized as shown in the following algorithm.

```
Algorithm FP-Growth
1:   Scan the transaction dataset T for the first time
Construct L, the list of frequent items with their support
            Sort L in descending order of support
```

```
Construct the TreeB containing a root labelled" null"
Called FP-Growth(TreeB, null)
Procedure FP-Growth(Tree, α)
If (tree contains a single path P) then
      For (each β = combination in P) do
             pattern = β ∪ α
support_min = min (support of the nodes of β)
Generate itemset β ∪ α with support = support_min
      End For
Else
      For (each a_i in the header of Tree) do
             support = support_ai
             patternβ = a_i ∪ α with support
             Construct conditional FP-tree of β
 TreeB = Construct conditional FP-tree of β
             If TreeB ≠ null then
                    Called FP-Growth(TreeB, β)
             End If
     End For
End If
End
```

The tree comprises nodes that are labelled with the items they represent, the number of times the item appears in the transactions and a link to the next node in the tree. If there is no such node, the link value is "null".

The index is a header table containing a list of frequent items and pointing to the first occurrence of each item. Each entry in the table includes the item name and the head pointer of the sequence of nodes that have the same name.

Constructing the FP-tree involves two passes over the transaction base (T). In the first pass, frequent items are determined based on the minimum support provided, and they are sorted in descending order of support in a list (L). Items in this list are processed in that order. The second pass over (T) follows, during which each transaction is sorted based on the order of the items in (L). The root node of the null tree is created first, and for each transaction, a branch is built. Transactions with the same prefix share the same branch start in the tree, meaning that a single branch represents two identical transactions.

Implementation

Once the theoretical aspects of association rule analysis algorithms, and specifically the FP-Growth algorithm, are understood, the first step in the unsupervised phase is to identify frequent events. This step is crucial in gaining insights into user preferences to develop effective recommendation systems. There are various solutions available to discover frequent events, including the maximal frequent itemset and closed frequent itemset techniques, as presented in (Verma & Chen, 2012). In this study, the closed frequent itemset technique is found to be appropriate for identifying frequent events and is utilized for generating association rules.

a. Closed frequent itemset maximal

Maximal and closed itemset both are a subset of a frequent itemset, but maximal is more compact. "Closed frequent itemsets are more widely used than maximal frequent itemsets because when efficiency is more important than space, they provide us with the support of the subsets so no additional pass is needed to find this information" (Verma & Chen, 2012).

This is a more mechanical and nescient way, but may be suitable in many cases. In this study, the theoretical closed frequent itemset approach works better because of its innate and conscious methodology.

A frequent itemset that does not have a superset with the same support count as the original itemset is called a closed frequent itemset. After identifying all frequent itemsets, it has to be ensured that each superset has the same support number as the original itemset. If so, they are disqualified, and those that do not are often closed.

The closed frequency is less sparse than the maximum frequency. Table 3 shows a snippet of the result of the closed frequent itemset (fifth participant's dataset).

Table 3. *Snippet of result of closed frequent itemset (fifth participant's dataset).*

Size	Support	Item1	Item2	Item3
1	0.984391	mainDoorLock		
1	0.466075	livingLight		
1	0.452460	bedroomDoor		
2	0.452175	bedroomDoor	mainDoorLock	
1	0.430237	livingCarp		
1	0.410316	tv		
...				
2	0.001450	bathroomDoor	fridge	
3	0.001408	wardrobe	bedroomCarp	hallwayLight
2	0.001397	wardrobe	bedTableLamp	
3	0.001355	bedroomLight	fridge	bedTableLamp
3	0.001355	bedroomLight	bathroomDoor	fridge

b. Learning association rules

The next step in implementing the unsupervised phase is to learn the association rules from these frequent itemsets. A frequent itemset with n items can create ($3n$–$2n$+1+1) rules (Tan et al., 2005). This can create hundreds of rules and devalue important and interesting rules. Most of them are irrelevant and not suitable for our recommender system.

It is crucial to set appropriate thresholds for metrics such as confidence, support and lift to filter out unimportant rules (Han et al., 2004). Careful consideration of these metric thresholds is essential to learn the relevant rules. In this study, confidence and lift are selected and implemented to extract interesting rules from infrequent events (Azevedo & Jorge, 2007). By doing so, the number of rules generated to learn user preferences is reduced and the number of rules varies across different users. Table 4 displays the snippet of the rules generated for the fifth participant's preferences.

Table 4. *Snippet of result of association rules to learn fifth participant's preferences.*

Premises	Conclusion	Support	Confidence	Lift	Conviction
- wardrobe	bedroomCarp	0.004664	0.975453	22.469578	38.969566
- bed	bedroomCarp	0.011824	0.907356	20.900958	10.325391
- bathroomCarp	bathroomLight	0.141845	0.926002	6.211755	11.49927
wardrobe, hallwayLight	bedroomCarp	0.001409	0.923077	21.263095	12.435642
tv, oven	fridge	0.051737	0.989418	7.127977	81.382673
tv, kitchenLight	fridge	0.052567	0.974359	7.019489	33.5865
tv, kitchenCarp	fridge	0.055334	0.950324	6.846335	17.336176
oven, hallwayLight	fridge	0.051737	0.989418	7.127977	81.382673
hallwayLight, kitchenLight	fridge	0.052567	0.974359	7.019489	33.5865
kitchenDoor, hallwayLight	fridge	0.056729	0.962448	6.933682	22.933595
kitchenCarp, hallwayLight	fridge	0.055334	0.966184	6.960591	25.466687
bedroomLight, hallwayLight	bedroomCarp	0.002465	0.913043	21.031975	11.00076
bedroomDoor, hallwayLight	bedroomCarp	0.004309	0.92446	21.294964	12.663401
bedroomLight, bathroomCarp	bathroomLight	0.099272	0.912411	6.12059	9.715047
bathroomDoor, bathroomLight	bathroomCarp	0.099272	0.940133	6.137418	14.145104
bathroomDoor, bedroomLight, bathroomLight	bathroomCarp	0.009287	0.989283	6.458276	79.012955
bedroomDoor, bathroomDoor, bathroomLight	bathroomCarp	0.00946	0.989477	6.459544	80.471366

5.3 Supervised phase (recommendation)

To initiate the recommendation phase, the first step is to process the raw data and extract important contextual information. In addition to the activity attributes present in the data, three additional contexts were derived, namely the time of day, the day of the week and the user's location within the house (Gupta, 2020).

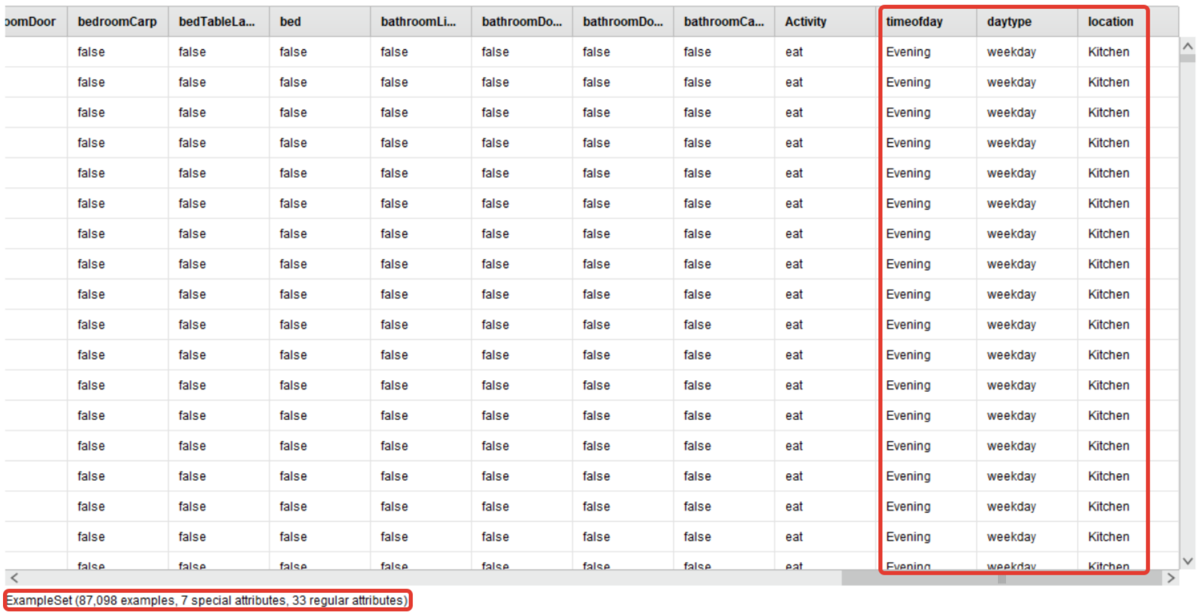


Figure 4. *Snippet of extraction and representation of contextual information as attributes.*

The unsupervised phase learns rules and the premises and conclusions of these rules are extracted and converted into a usable format. The raw data and the rules are then combined to form a new data frame, where duplicates are removed and the context is logically included as shown for example in Figure 4. The resulting data frame contains only the required attributes for modelling, which are the sensor data, contextual information and user preferences. These results are combined into an attribute called "label", which indicates the preferred action to take in a given contextual setting by the evaluator.

This dataset is ready to export for modelling where the label will be the consequent attribute.

5.3.1 Modelling

Using a supervised learning algorithm, we train a model to predict and recommend inhabitant behaviour in a given contextual setting based on the rules and datasets constructed in the unsupervised learning phase, which contains the desired user preferences stored in the conclusion column, as shown in Table 4, and will be treated as a label for a supervised learning classification problem. GLM, DL, GBT and other algorithms are qualified for use to build the predictive model. Here, GLM is used to generate a predictive model using rapid prototyping in Rapid Miner (https://rapidminer.com/) (Kotu & Deshpande, 2014).

Generalized linear model: GLM

A GLM, or generalized linear model, is a statistical method that extends linear regression to accommodate various output distribution functions that can describe the variance between predicted and observed values. The method uses a link function that specifies the relationship between inputs and outputs. Unlike classical linear regression, where changes in input variables lead to constant changes in output values, GLMs allow for non-linear relationships between inputs and outputs. An example of a commonly used link function is the exponential function of the response density function. This allows the variance of each observation to be modelled as a function of its predicted value (Zheng & Agresti, 2000).

If the underlying data generation process is binomial or polynomial (proportions), Poisson (counts), or exponential (time-to-event), then the responses given to the set of predictors are not normally distributed. In such situations, ordinary linear regression can predict ratios outside the range [0, 100], or negative counts or times.

The GLM solves this problem by modelling a function f of the expected value of y.

The GLM consists of three components: the random component is the probability distribution of the response variable (normal, binomial, Poisson, etc.); the systematic component is the explanatory variable $X\beta$; and the link function μ provides a link between the random and systematic components by transforming the response domain to $[-\infty, +\infty]$.

Linear regression is therefore a special case of the GLM whose link function is the identity function:

$$f(E(X \vee Y)) = E(X \vee Y) \quad (3)$$

For logistic regression, where the data generation process is binomial, the link function is:

$$f(E(Y \vee X)) = ln\left(\frac{E(Y|X)}{1 - E(Y|X)}\right) = ln\left(\frac{\pi}{1-\pi}\right) = logit(\pi) \quad (4)$$

where π is the event probability.

For a Poisson regression, the link function is:

$$f(E(Y \vee X)) = ln(E(Y \vee X)) = ln(\lambda) \quad (5)$$

where λ is the expected event rate.

https://doi.org/10.18267/j.aip.228

For an exponential regression, the link function is:

$$f(E(Y \vee X)) = -(E(Y \vee X)) = -\lambda \quad (6)$$

where λ is the expected event rate.

The GLM uses maximum likelihood estimation (MLE) rather than ordinary least squares (OLS) to estimate the parameters, and thus relies on large-sample approximations.

6 Results and Discussion

6.1 Results

This study demonstrated the possibility of identifying usage patterns and recommending personalized services in smart homes, which have gained paramount importance due to the improvement of comfort to the inhabitant based on the use of machine learning techniques in monitoring and analysing usage datasets collected from IoT devices connected to home automation systems.

In the unsupervised phase, we carefully processed the raw data to learn the inhabitants' behaviour. FP-Growth association analysis algorithms were used to generate association rules that could be considered similar to real user preferences. In fact, meaningful rules were discovered. These user rules are also marked by high values of confidence, meaning high co-occurrence of premises and conclusion, high values of lift, meaning interesting rules, and high conviction, meaning high dependence on the antecedent. During the supervised phase, contextual information such as location and time of day was extracted from the same raw data and merged with the rules generated during the unsupervised phase. After the necessary preparation, the data were modelled in Rapid Miner to generate results.

From the summary of the GLM performance shown in Table 5, we can see that the models have high classification scores for average accuracy, average precision and average recall, with little variation in their values.

Table 5. *Performance (classification).*

Participant	Accuracy	Classification error	Mean precision	Mean recall
d1	92.32%	7.68%	50.29%	58.25%
d2	81.77%	18.23%	67.00%	85.71%
d3	93.11%	6.89%	85.54%	91.56%
d4	80.01%	19.99%	79.07%	88.26%
d5	97.34%	2.66%	97.97%	91.8%
d6	82.13%	17.87%	66.67%	75.52%
d7	82.25%	17.75%	85.88%	87.56%
Average	86.99%	13.01%	76.06%	82.67%

The confusion matrix is indispensable for assessing the performance of a classification model. It offers a detailed breakdown of predictions for each class, allowing a thorough analysis of precision, recall and other metrics. This information is essential for fine-tuning models, identifying areas of improvement and making informed decisions about the model deployment.

Table 6. *Fifth participant's GLM confusion matrix values.*

	True fridge	True couch	True tv	True officeLight	True office	Class precision
Pred. fridge	18786	0	0	0	0	100.00%
Pred. couch	0	46404	9028	0	0	83.71%
Pred. tv	0	0	0	0	0	0.00%
Pred. officeLight	0	0	0	6440	6440	50.00%
Pred. office	0	0	0	0	0	0.00%
Class recall	100.00%	100.00%	100.00%	100.00%	0.00%	

The confusion matrix breaks down the model predictions into the following categories: true positives, true negatives, false positives, false negatives. For example, in the GLM confusion matrix for the fifth participant (as illustrated in Table 6):

- True fridge: 18,786 instances were correctly predicted as fridge.
- True couch: 46,404 instances were correctly predicted as couch.
- True tv: None of the instances were correctly predicted as tv.
- True officeLight: 6,440 instances were correctly predicted as officeLight.
- True office: None of the instances were correctly predicted as office.

Table 6 shows that certain classes, such as fridge, couch and officeLight, exhibit both high precision and recall, indicating that the model is effective in accurately predicting these classes. Notably, the model failed to predict the office and TV classes, resulting in low precision and recall scores for these categories, suggesting that the model struggles to correctly identify instances of these classes.

On the other hand, context awareness is crucial in various technological applications, and the importance of specific contextual attributes, such as location, activity, day type and time of day, is highlighted in the study. The emphasis on context-aware technology, as indicated by the attribute weights in Figure 5 for the sixth participant (for the attribute weights for the remaining participants, see Figures A1-A6 in the Appendix A), underscores the value of incorporating contextual information into technological systems.

Location: Knowing the user's location provides essential context for tailoring services or recommendations. For instance, a smart home system might adjust lighting, temperature or security settings based on the occupant's location within the home.

Activity: Understanding the user's current activity allows more personalized and adaptive interactions. For a smart home, this could involve adjusting the environment based on whether the user is working, relaxing, exercising or engaged in other activities.

Day type: Different days may have distinct patterns or requirements. For example, weekdays might involve work-related activities, while weekends could be associated with leisure and relaxation. Contextualizing services based on the type of day enhances the system's ability to meet user needs.

Time of day: The time of day is a critical contextual factor influencing user behaviour and preferences. Smart systems can use this information to automate processes such as adjusting lighting conditions, suggesting activities or managing energy consumption based on the time of day.

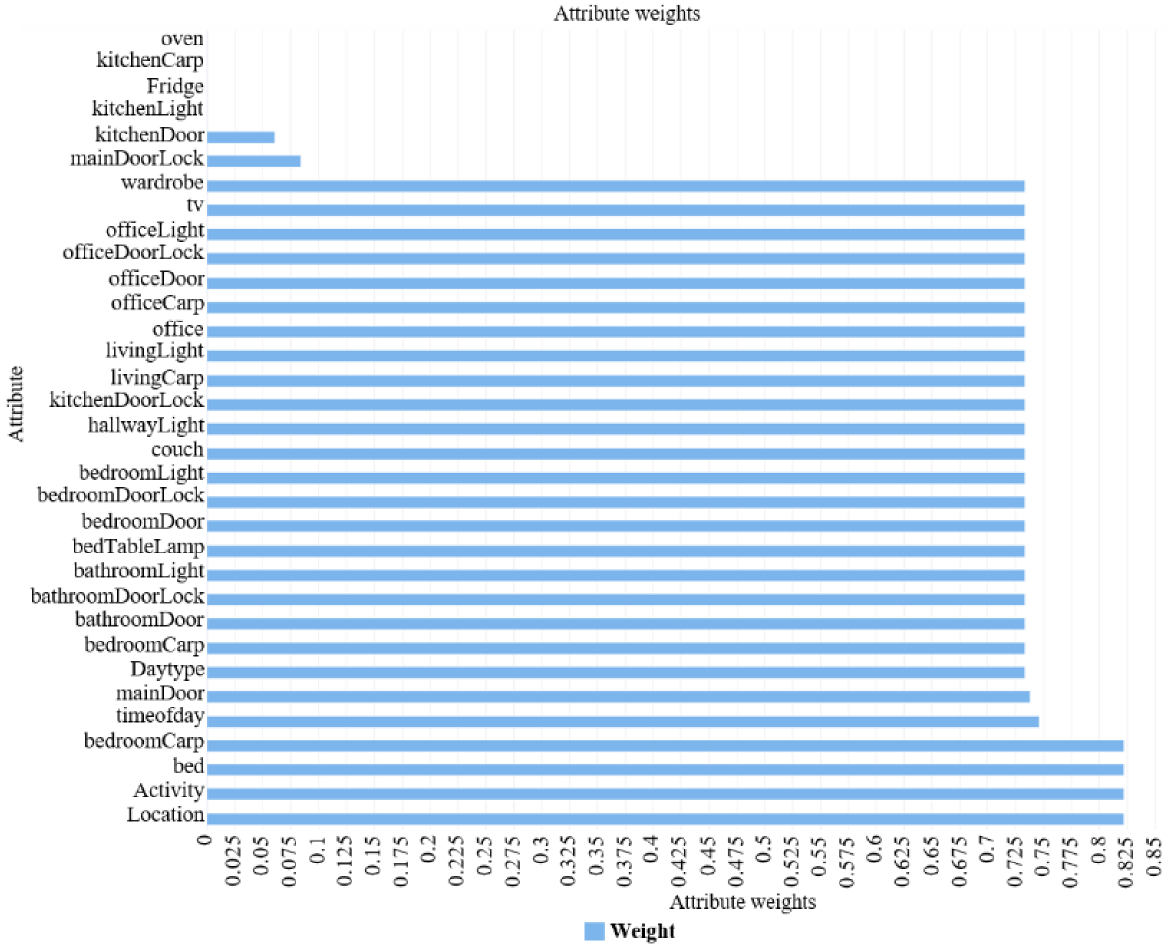


Figure 5. *Assigning weights to various attributes in GLM – 6th participant's attribute weights.*

Assigning greater weight to these contextual attributes indicates their higher significance in shaping the user experience. This prioritization recognizes that users' needs, preferences and behaviour are intricately linked to their context. By giving more importance to location, activity, day type and time of day, the system can offer more tailored and relevant services, leading to enhanced user experience and improved system performance. Context awareness, in this context, ensures that technology adapts to the dynamic nature of users' lives, making it more intuitive, responsive and ultimately more valuable.

6.2 Discussion

The development of a personalized recommendation system for smart homes is an essential aspect of enhancing the quality of life for residents. This study presents an intelligent framework that achieves this by capturing the behaviour patterns of residents through the use of contextual information and machine learning techniques. Unsupervised observation was utilized to identify resident behaviour patterns, and the FP-Growth association analysis algorithm was employed to extract frequent events and user preferences as rules. These rules were then integrated with contextual information to train a supervised learning algorithm, which predicts the recommended action for a given situation, resulting in the creation of a predictive model. The system's ability to offer personalized recommendations based on user behaviour patterns and contextual information is a significant step in providing an efficient and enjoyable daily routine for smart home residents. Furthermore, the framework's potential to be applied to other domains where personalized recommendations are required is an exciting prospect.

During the unsupervised phase, we observed that the frequent itemset and resulting rules can become unwieldy if not managed properly. We found that utilizing the closed frequent itemset method to reduce itemset size was effective in this study, producing compact and logically appropriate rules.

In the second phase, the supervised learning model used to predict user-preferred actions showed a high accuracy rate. However, for imbalanced classes, such as in this case, precision and recall are better indicators of prediction success. Precision measures the relevance of results, while recall measures the number of truly relevant results returned. We observed that the GLM produced excellent results for these two parameters.

We discovered that the supervised model that performed well on precision/recall measures assigned higher weights to contextual attributes. Previous studies have emphasized that systems that account for contextual intelligence, such as Gupta (2020) and Rasch (2014), tend to perform better on these measures. In the GLM model, we identified four contextual attributes with higher weights than other attributes for most participants. Therefore, we recommend a separate comparative study to examine the impact of contextual information on recommender system performance in the smart home environment.

Furthermore, when comparing our findings with other studies and analysing the outcomes as presented in Table 7, it becomes evident that our system surpasses the GUPTA (Gupta, 2020) system performance on the Alshammari dataset (Alshammari et al., 2018) in terms of recall, precision and accuracy. Moreover, our system demonstrates higher recall scores and greater accuracy compared to the Rasch Katharina system (Rasch, 2014), both in the Kasteren houseA (Van Kasteren et al., 2008) and Kasteren houseB (Van Kasteren et al., 2010) datasets. These results underscore our system's superior ability to predict and recommend context-aware services within a smart home environment.

Table 7. *Comparison of results of our system with other systems.*

	Dataset	**Recall**	**Precision**	**Accuracy**
Rasch, Katharina system	Kasteren houseA dataset	61.00%	63.00%	61.00%
	Kasteren houseB dataset	73.00%	78.00%	73.00%
GUPTA system	Alshammari dataset	75.67%	60.03%	80.22%
Our system	Alshammari dataset	82.67%	76.06%	86.99%

7 Conclusion

The article describes a context-aware personalized recommendation system designed for a smart home. The study consisted of two phases: unsupervised and supervised. In the unsupervised phase, the dataset was selected, loaded and pre-processed to discover frequent behaviour rules and establish association rules using the FP-Growth algorithm. In the second phase, the raw data were processed to extract relevant contextual information, which was combined with the association rules to form a new data frame. An experimental case study was conducted to verify the effectiveness of the platform. The FP-Growth algorithm was effective in detecting repeated user actions and establishing resident behaviour in the form of association rules. By filtering these rules, meaningful resident behaviour was identified. The supervised learning model (GLM) used in the second phase had a high accuracy rating, and the test for precision/recall produced favourable results. The study found that assigning higher weights to contextual information resulted in better performance on these metrics.

Exploring the extension of the recommendation system to accommodate multiple inhabitants represents a promising avenue for future research. This endeavour introduces a set of challenges and considerations that merit careful exploration and investigation.

Behavioural models for multiple residents

Implementing the recommendation system for multiple inhabitants necessitates the development of distinct behavioural models for each resident. Each individual possesses unique preferences, habits and contextual considerations. Future work should delve into methodologies to effectively capture, analyse and model the diverse behaviour of multiple residents within a shared smart home environment.

User interactions and preferences

Understanding how the recommendation system can adapt to and differentiate between the interactions and preferences of various inhabitants is crucial. Future research should focus on designing algorithms that can discern individual user preferences amid shared spaces, ensuring that recommendations align with the specific needs and desires of each resident.

Privacy and personalization balance

As the recommendation system becomes more tailored to individual users, privacy concerns may arise. Future work should explore strategies to strike a balance between providing personalized recommendations and respecting the privacy of each inhabitant. Implementing user-controlled privacy settings or employing privacy-preserving machine learning techniques could be areas of exploration.

Dynamic household environments

Households are dynamic and inhabitants' preferences may evolve over time. Future research could investigate adaptive models that can accommodate changes in user behaviour and preferences, ensuring that the recommendation system remains relevant and effective in dynamic household environments.

Evaluation metrics for multi-user systems

Establishing appropriate evaluation metrics becomes crucial when dealing with multiple inhabitants. Future work should develop metrics that reflect the overall system performance while considering individual user satisfaction and the system's ability to cater to the diverse needs of the household residents.

User feedback and system learning

Implementing mechanisms for users to provide feedback on recommendations can enhance the learning capabilities of the system. Future research could explore ways to incorporate user feedback into the recommendation algorithms, facilitating continuous learning and improvement of the system over time.

Scalability and resource efficiency

Scaling the recommendation system for multiple inhabitants requires consideration of resource efficiency. Future work should address the scalability challenges associated with increasing numbers of users and devices, ensuring that the system remains efficient and responsive as the smart home environment expands.

Integration with smart home ecosystems

Future research could focus on seamlessly integrating the recommendation system with broader smart home ecosystems. This involves exploring interoperability with various devices and platforms, providing comprehensive and cohesive smart living experience for all inhabitants.

By addressing these facets in future work, researchers can contribute to the development of recommendation systems that not only accommodate the complexities of multi-user environments but also enhance the overall quality and adaptability of smart home automation.

Additional Information and Declarations

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: T.D.: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. S.B.: Supervision, Writing – review & editing. A.M.: Writing – review & editing. S.K.: Writing – review & editing.

Data Availability: The datasets used in this article are also available from https://zenodo.org/records/1185172

Appendix A

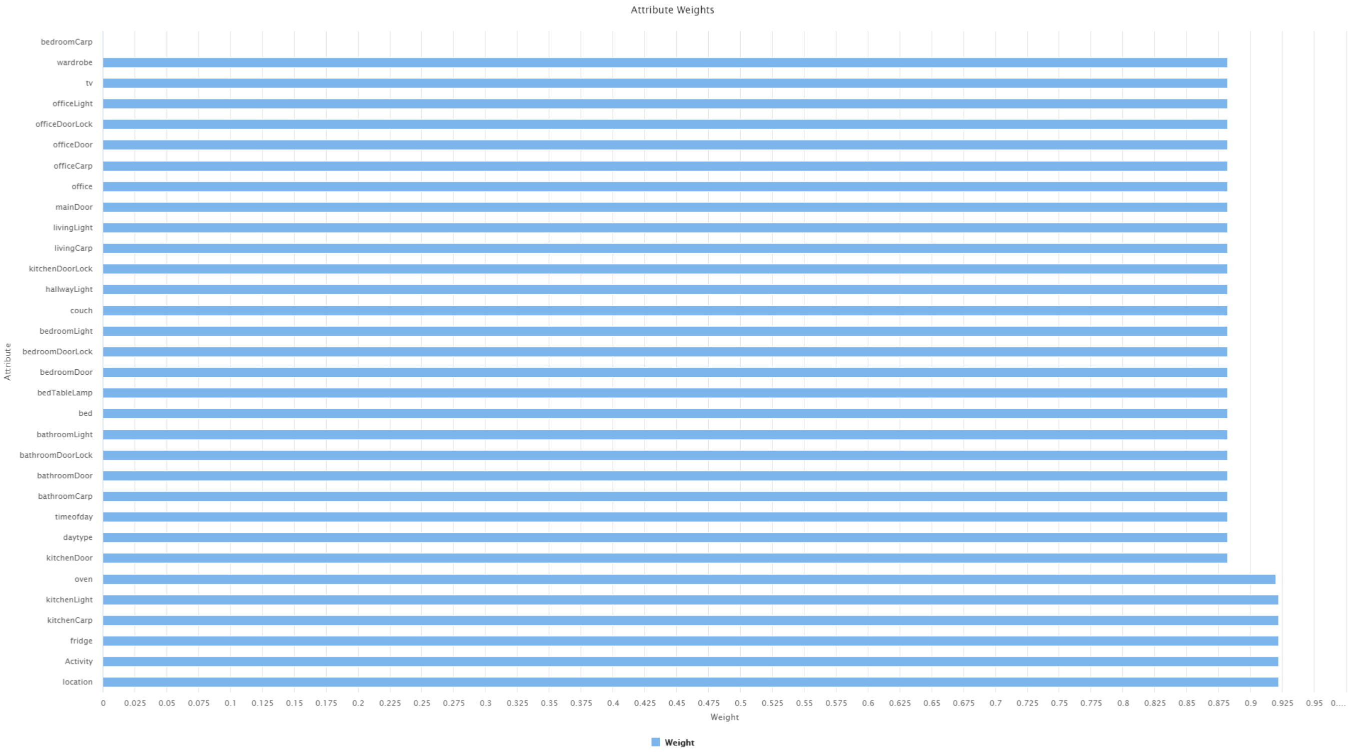


Figure A1. *Assigning weights to various attributes in GLM – 1st participant's attribute weights.*

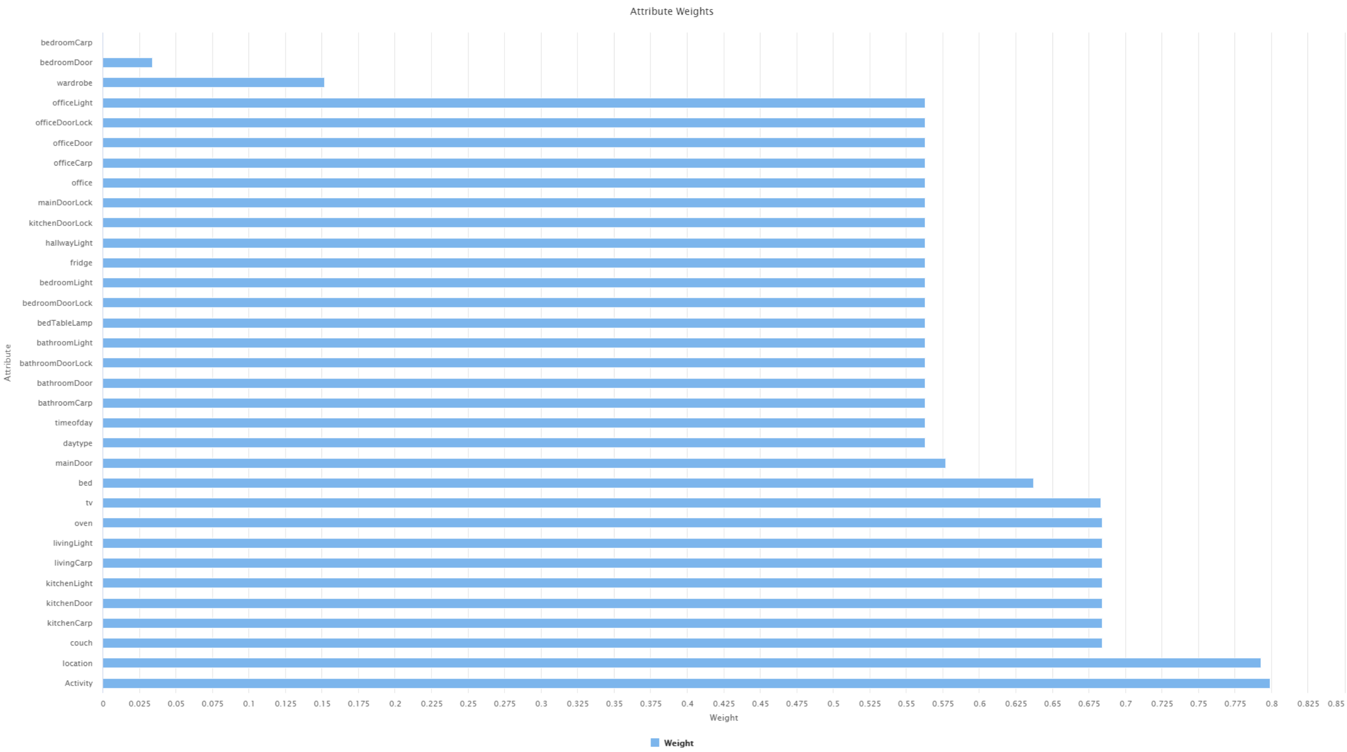


Figure A2. *Assigning weights to various attributes in GLM – 2nd participant's attribute weights.*

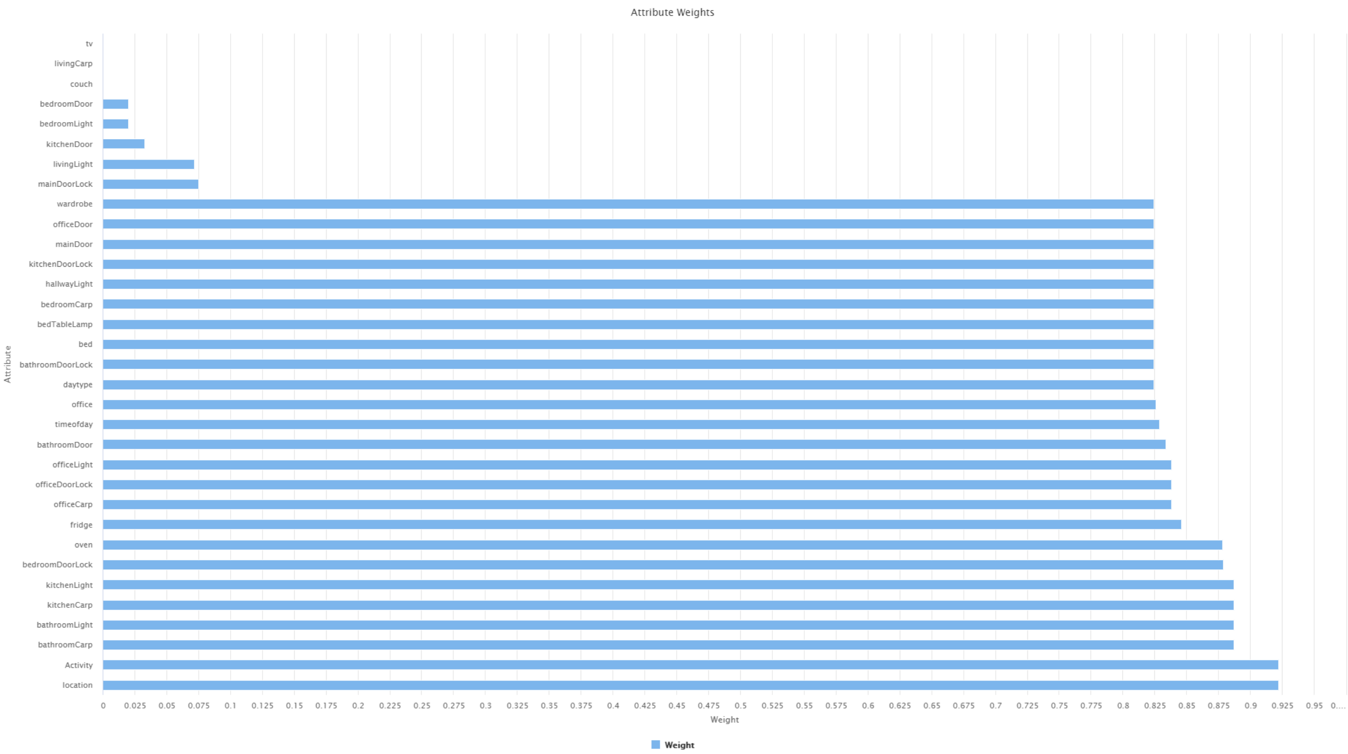


Figure A3. *Assigning weights to various attributes in GLM – 3rd participant's attribute weights.*

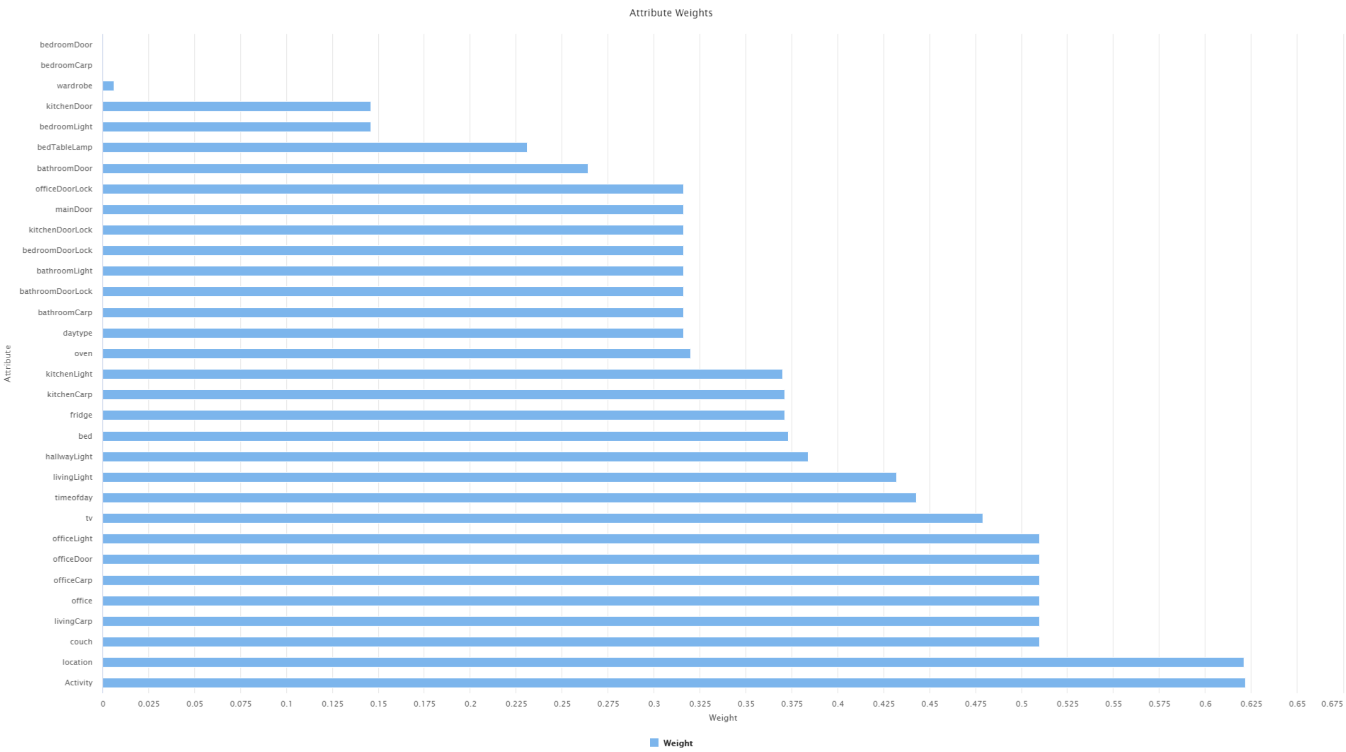


Figure A4. *Assigning weights to various attributes in GLM – 4th participant's attribute weights.*

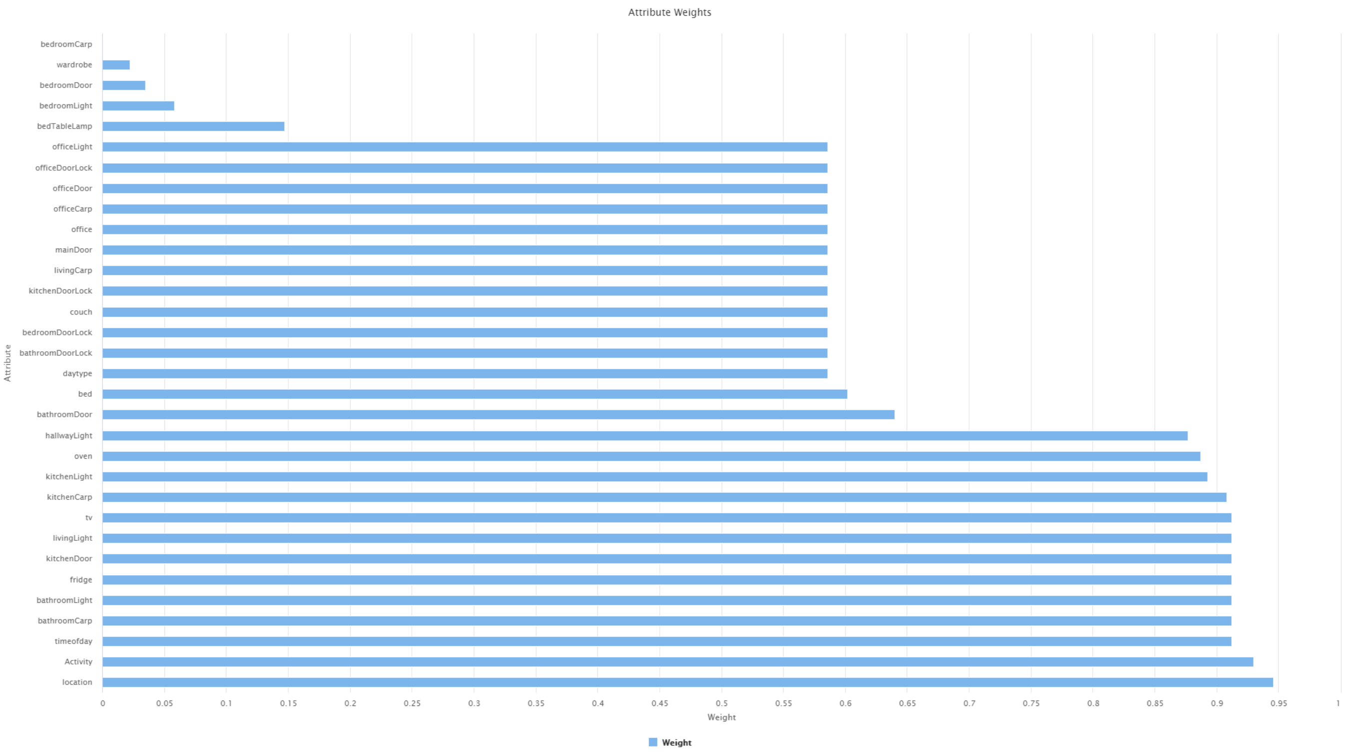


Figure A5. *Assigning weights to various attributes in GLM – 5th participant's attribute weights.*

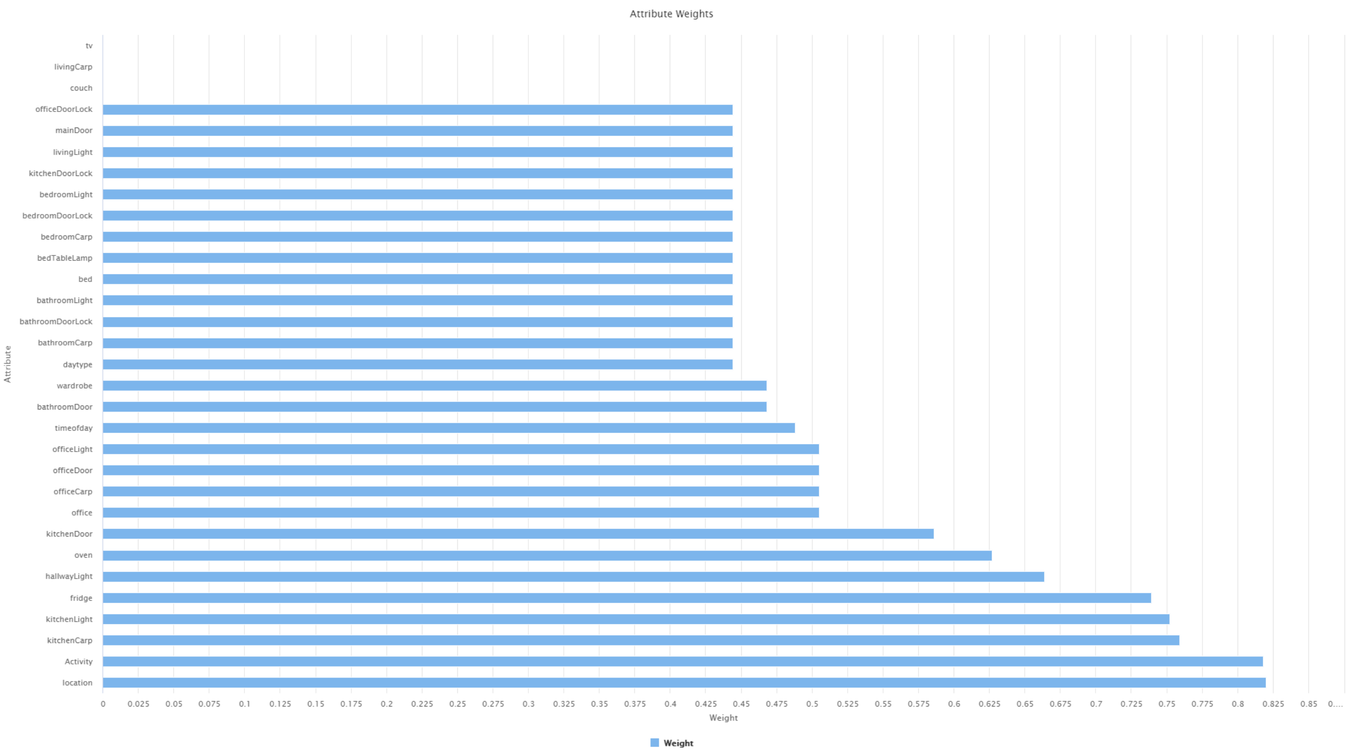


Figure A6. *Assigning weights to various attributes in GLM – 7th participant's attribute weights.*

References

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. In *Proc. 20th Int. Conf. Very Large Data Bases*, (pp. 487–499). VLDB. https://www.vldb.org/conf/1994/P487.PDF

Aldino, A. A., Pratiwi, E. D., Setiawansyah, Sintaro, S., & Dwi Putra, A. (2021). Comparison Of Market Basket Analysis To Determine Consumer Purchasing Patterns Using Fp-Growth And Apriori Algorithm. In *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE),* (pp. 29–34). IEEE. https://doi.org/10.1109/ICOMITEE53461.2021.9650317

Aldrich, F. K. (2003). Smart Homes: Past, Present and Future. In *Inside the Smart Home* (pp. 17–39). Springer-Verlag. https://doi.org/10.1007/1-85233-854-7_2

Ali, S. M. M., Augusto, J. C., Windridge, D., & Ward, E. (2023). A user-guided personalization methodology to facilitate new smart home occupancy. *Universal Access in the Information Society*, 22(3), 869–891. https://doi.org/10.1007/s10209-022-00883-x

Alshammari, T., Alshammari, N., Sedky, M., & Howard, C. (2018). SIMADL: Simulated Activities of Daily Living Dataset. *Data*, 3(2), 11. https://doi.org/10.3390/data3020011

Azevedo, P. J., & Jorge, A. M. (2007). Comparing rule measures for predictive association rules. In *European Conference on Machine Learning*, (pp. 510–517). https://doi.org/10.1007/978-3-540-74958-5_47

Belaidouni, S., Miraoui, M., & Tadj, C. (2022). QL-CBR Hybrid Approach for Adapting Context-Aware Services. *Computer Systems Science and Engineering*, 43(3), 1085–1098. https://doi.org/10.32604/csse.2022.024056

Belghini, N., Gouttaya, N., Bouab, W., & Sayouti, A. (2016). Pervasive Recommender System for Smart Home Environment. *International Journal of Applied Information Systems*, 10(9), 1–7. https://doi.org/10.5120/ijais2016451528

Brush, A. J. B., Lee, B., Mahajan, R., Agarwal, S., Saroiu, S., & Dixon, C. (2011). Home automation in the wild: Challenges and opportunities. In *Conference on Human Factors in Computing Systems - Proceedings*, (pp. 2115–2124). ACM. https://doi.org/10.1145/1978942.1979249

Davidoff, S., Lee, M. K., Yiu, C., Zimmerman, J., & Dey, A. K. (2006). Principles of Smart Home Control. In *UbiComp 2006: Ubiquitous Computing: 8th International Conference,* (pp. 19–34). Springer. https://doi.org/10.1007/11853565_2

Eggen, B., Hollemans, G., & van de Sluis, R. (2003). Exploring and enhancing the home experience. *Cognition, Technology & Work*, 5(1), 44–54. https://doi.org/10.1007/s10111-002-0114-7

Engelmann, K. F., Holthaus, P., Wrede, B., & Wrede, S. (2016). An interaction-centric dataset for learning automation rules in smart homes. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation* (LREC'16), (pp. 1432–1437). European Language Resources Association. https://aclanthology.org/L16-1228/

Gupta, S. (2020). *A context-aware personalized recommender system for automation in IoT based smart home environment*. Master thesis. Dublin Business School. https://esource.dbs.ie/handle/10788/4240

Han, J., Pei, J., Yin, Y., & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Mining and Knowledge Discovery*, 8, 53–87. https://doi.org/10.1023/B:DAMI.0000005258.31418.83

Hunyadi, D. (2011). Performance comparison of Apriori and FP-Growth algorithms in generating association rules. In *Proceedings of the 5th European Conference on European Computing Conference*, (pp. 376–381). ACM.

Kotu, V., & Deshpande, B. (2014). *Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner*. Morgan Kaufmann.

Leonidis, A., Korozi, M., Kouroumalis, V., Poutouris, E., Stefanidi, E., Arampatzis, D., Sykianaki, E., Anyfantis, N., Kalligiannakis, E., Nicodemou, V. C., Stefanidi, Z., Adamakis, E., Stivaktakis, N., Evdaimon, T., & Antona, M. (2019). Ambient intelligence in the living room. *Sensors*, 19(22), 5011. https://doi.org/10.3390/s19225011

Mennicken, S., & Huang, E. M. (2012). Hacking the natural habitat: An in-the-wild study of smart homes, their development, and the people who live in them. In *Pervasive Computing 2012*, (pp. 143–160). Springer. https://doi.org/10.1007/978-3-642-31205-2_10

Miraoui, M. (2018). A Context-Aware Smart Office for Improved Comfort and Energy Saving. In *Intelligent Environments 2018*, (pp. 455–465). IOS Press. https://doi.org/10.3233/978-1-61499-874-7-455

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A,* 135(3), 370–384. https://doi.org/10.2307/2344614

Ortiz-Barrios, M., Järpe, E., García-Constantino, M., Cleland, I., Nugent, C., Arias-Fonseca, S., & Jaramillo-Rueda, N. (2022). Predicting Activity Duration in Smart Sensing Environments Using Synthetic Data and Partial Least Squares Regression: The Case of Dementia Patients. *Sensors*, 22(14), 5410. https://doi.org/10.3390/s22145410

Osman, K., Štefić, M., Alajbeg, T., & Perić, M. (2023). Approach in development and implementation of automatic control system in a smart building. *Energy and Buildings*, 293, 113200. https://doi.org/10.1016/j.enbuild.2023.113200

Ramesh, P. N., & Kannimuthu, S. (2023). Context-Aware Practice Problem Recommendation Using Learners' Skill Level Navigation Patterns. Intelligent Automation and Soft Computing, 35(3), 3845–3860. https://doi.org/10.32604/iasc.2023.031329

Randall, D. (2003). Living Inside a Smart Home: A Case Study. In *Inside the Smart Home* (pp. 227–246). Springer-Verlag. https://doi.org/10.1007/1-85233-854-7_12

Rasch, K. (2014). An unsupervised recommender system for smart homes. *Journal of Ambient Intelligence and Smart Environments*, 6(1), 21–37. https://doi.org/10.3233/AIS-130242

Reyes-Campos, J., Alor-Hernández, G., Machorro-Cano, I., Olmedo-Aguirre, J. O., Sánchez-Cervantes, J. L., & Rodríguez-Mazahua, L. (2021). Discovery of Resident Behavior Patterns Using Machine Learning Techniques and IoT Paradigm. *Mathematics*, 9(3), 219. https://doi.org/10.3390/math9030219

Tan, P.-N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Wesley Longman.

Van Kasteren, T. L. M., Englebienne, G., & Kröse, B. J. A. (2010). Transferring knowledge of activity recognition across sensor networks. In *Pervasive Computing*, (pp. 283–300). Springer. https://doi.org/10.1007/978-3-642-12654-3_17

Van Kasteren, T., Noulas, A., Englebienne, G., & Kröse, B. (2008). Accurate activity recognition in a home setting. In *UbiComp 2008 - Proceedings of the 10th International Conference on Ubiquitous Computing,* (pp. 1–9). ACM. https://doi.org/10.1145/1409635.1409637

Verkasalo, H., & Karjalainen, J. (2010). *System and method for behavioural and contextual data analytics*. International patent, https://patents.google.com/patent/WO2010128198A1/ja

Verma, R. M., & Chen, P. (2012). A data mining hypertextbook: Design, implementation and experience. *Journal of Computing Sciences in Colleges*, 27(3), 22–28.

Woodruff, A., Augustin, S., & Foucault, B. (2007). Sabbath day home automation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 527–536). ACM. https://doi.org/10.1145/1240624.1240710

Zheng, B., & Agresti, A. (2000). Summarizing the predictive power of a generalized linear model. *Statistics in Medicine*, 19(13), 1771–1781. https://doi.org/10.1002/1097-0258(20000715)19:13%3C1771::AID-SIM485%3E3.0.CO;2-P

Editorial record: The article has been peer-reviewed. First submission received on 19 September 2023. Revision received on 13 December 2023. Accepted for publication on 18 December 2023. The editor in charge of coordinating the peer-review of this manuscript and approving it for publication was Stanislav Vojir.

Acta Informatica Pragensia is published by Prague University of Economics and Business, Czech Republic.

ISSN: 1805-4951
