



# A semantic blocks model for human activity prediction in smart environments using time-windowed contextual data

Rob Dunne<sup>1</sup> · Tim Morris<sup>1</sup> · Simon Harper<sup>1</sup>

Received: 6 June 2022 / Accepted: 5 November 2022 / Published online: 24 November 2022  
© The Author(s) 2022

## Abstract

Complex human activity prediction is a difficult problem for computer science. Simple behaviours can be mapped to sequence prediction algorithms with good results; however, real-world examples of activity are generally stochastic and much more computationally difficult to infer. One method for solving this problem is to utilise contextual data—clues surrounding the actual activity—to decipher what is about to happen next; in much the same way humans do. In this paper, we present the semantic blocks model (SBM), a method for using contextual data to infer the next activity in a smart home environment by augmenting the inference with contextual data, but also segmenting it into time-windowed sections of activity—or semantic blocks. Our proof-of-concept produces 74.55% accuracy on the CASAS smart home dataset, an increase on the comparable CRAFFT algorithm which produces 66.91% on the same dataset. We detail how our experimental prototype works using intersecting contextual data, and explore opportunities for further work by the research community.

**Keywords** Human activity prediction · Smart environments · Ambient intelligence

## 1 Introduction

In recent years, there have been great strides taken towards ubiquitous computing and Ambient Intelligence (AmI) [8], most notably with remarkable advances in pervasive computing technologies [1, 10, 24], machine learning and data mining [23], networking [14], and mobile devices [25, 27]. We have started to see the beginnings of this permeate through to consumer markets in the form of smart home assistants such as Amazon's Alexa and Google Home, Internet of Things (IoT) devices, and mobile applications that use edge computing, e.g. face recognition built into mobile apps. There is still much work to be done to join up all these dots to form coherent AmI, and also to make these systems into predictive rather than purely reactive request and response smart environments.

It is this latter challenge that forms the basis of this research. There are many ways to approach the problem of human activity prediction in smart environments. Machine Learning methods using time-series prediction with LSTMs are common in the literature [23], as are probabilistic approaches using Markov Model [21] and Bayesian Network variants [17]. These approaches all tackle the fundamental component of this problem which is inference of the next event in a sequence. Sometimes extending this to predict many steps ahead. However, this gives us a very simplistic form of computational inference which is difficult to map the complex, multi-faceted, stochastic nature of everyday life. To produce better prediction systems we need to enrich them with more data—contextual data that can enable more accurate analysis of a situation [5, 13, 20].

In this paper, we present the semantic blocks model (SBM). A method for using contextual time-windowed data for next activity prediction. This can be used as a standalone model or on conjunction with another predictive model that shortlists potential next activities, e.g. a Dynamic Bayesian Network (DBN). We show improved accuracy over a comparable system—the CRAFFT algorithm [17], and detail how the SBM works and can be extended to accommodate further contextual data.

---

✉ Rob Dunne  
rob.dunne@manchester.ac.uk

Tim Morris  
tim.morris@manchester.ac.uk

Simon Harper  
simon.harper@manchester.ac.uk

<sup>1</sup> The University of Manchester, Manchester, UK

## 1.1 Background

At The University of Manchester, we have been working on computational methods of human behaviour modelling and prediction. In this paper, we present work on the next step in that research, which goes beyond simple sequence prediction, and presents a proof-of-concept method for refining prediction results by creating an extensible model of contextual data to enable more accurate inference systems for smart environments. The SBM builds upon previous research which indicates that contextual data is key to improving the accuracy of predictive models in this field.

## 1.2 Related work

The state of the art in computational methods of human behaviour prediction falls into three categories: machine learning (ML), statistical forecasting and trend analysis, and Probabilistic Graphical Models (PGM). ML predictive models use sequence prediction with Recurrent Neural Networks (RNN), and more specialised subsets of RNNs such Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) [3,23]. These all produce good results for simple data with accuracies  $> 90\%$ . Statistical methods such as AutoRegressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Facebook's Prophet library tend to look at macro level behaviours for higher level trends in human behaviour [22]. Lastly, PGMs which are essentially graph data structures with probability functions on the graph edges representing state transitions from one vertex, or state, or another. These take many forms depending on the probability method and nuances of modelling behaviour. Dynamic Bayesian Network (DBN) is good for representing a chain of causal events [17]. If we only want the current state observable in determining the next state, then a Markov Model (MM) can be used. Hidden Markov models (HMM) allow for representation of hidden states in influencing state transition [15,21]. There are many variations of PGMs in the literature that represent the different facets of behaviour. Many of these have accuracy results of  $> 90\%$  for simple behaviours such as moving from one location. Human activities, however, are complex stochastic processes, which overlap, stop, start, and can be difficult to predict. Adding contextual data to a model can provide an insight into this complexity and disambiguate the inference process [1,9]. PGMs are adept at extension and augmentation with data due to the flexibility of the graph data structure. A significant research contribution in the field of Ambient Intelligence (AmI) and smart environments is the CRAFFT algorithm [17]. A PGM in the form of a Bayesian Network that uses current activity and features to predict next features, by augmenting with contextual data. Context aware ontologies and semantic representation frameworks for intelligent and smart environments exist in the literature and could

**Table 1** CRAFFT prediction accuracies using the CASAS smart home datasets

CASAS dataset apartment id	Accuracy
HH101	64.18%
HH102	74.57%
HH103	61.98%
<b>Averaged</b>	<b>66.91%</b>

be leveraged in combination with the ML, PGM, and statistical approaches to provide improved contextual data support to a predictive system [2,6,11,18].

The CRAFFT algorithm is the closest benchmark to the semantic blocks model and achieves an average accuracy of 66.91% over three of the CASAS smart home datasets, as shown in Table 1.

CRAFFT differs from many of the inference models by augmenting a well known algorithm with added contextual data. The SBM model presented here further extends this novelty by applying data clustering via time windows to add another dimension to the inference model.

## 1.3 Motivation

The shift from reactive to predictive smart environment systems is a rapidly growing area of research, particularly with Ambient Assistive Living (AAL) use cases, such as determining behavioural trends or anomalous activity [12], but also in supporting activities of daily living (ADL) such as using electronic devices, cooking, and dressing [16]. The research findings presented here build on previous human activity prediction research questions conducted by this research team at the University of Manchester in the Interaction Analysis and Modelling Laboratory (<https://archive.is/cQZ98>) and Advanced Interfaces Group Laboratory (<https://archive.is/igPa>), which is focussed on developing these computational capabilities in the context of realising the potential of Ambient Intelligence [7,19] by improving smart environment technology.

There are many approaches to Human Activity Prediction (HAP) including Bayesian Networks [17], Markov Model variants [26], and Deep Learning such as RNNs and LSTMs [23]. All of which are capable of producing results with varying degrees of accuracy depending on the complexity of the activity and number of occupants in the observed space. The motivation for our work here is that a broad range of different methods have been applied to this prediction problem, and what is now needed is a refinement of the inference capability to make this technology more accurate and robust. To enable more accurate predictive models, we can make better use of secondary data outside of the event or activity sequences. This contextual data support the inference pro-

cess by enabling differentiation between similar sequences of data. The SBM proposes one method which can be used towards this research goal.

#### 1.4 Contribution to research

We present a novel proof-of-concept model which improves prediction accuracy of human activity in a smart home, demonstrated using the CASAS smart home datasets [4], and benchmarked against an established method in this area; see Table 1. The model calculates the probability intersection of the next activity label, and sets of contextual labels, based on previously observed data from the same environment. We demonstrate improved prediction accuracy against a leading peer in this area—the Bayesian Network CRAFFT model—and also demonstrate a novel contribution to research by building upon CRAFFT and adding time window data clustering to improve the inference model.

#### 1.5 Data availability

The experimentation code and results data that support the findings of this study are available in the Mendeley Data repository: a semantic blocks model for contextually enhanced human activity prediction, with the identifier DOI 10.17632/jm28kgt7tm.1 available online and open access<sup>1</sup>.

## 2 Methods

The methodology for testing and analysing our SBM proof-of-concept includes several stages. We describe the hypothesis we will test—namely that we can improve prediction accuracy with a probability model based on contextual data. We then outline the research method used, our choice of dataset for testing, and benchmarks using existing algorithms against that dataset. Finally, we detail the experimentation method used to verify the conjectures made in our hypothesis.

### 2.1 Hypothesis

Human activity prediction in smart environments can be improved by enriching the data with contextual labels and a time window, to enable differentiation between similar events. We hypothesise that adding more contextual data will improve the prediction accuracy of the next activity label.

In this paper, we propose the semantic blocks model (SBM), a deep contextual data probability model, which will improve inference accuracy by taking a list of next activity

labels, which could be the causal probability output of a predictive model, e.g. a Dynamic Bayesian Network (DBN) and then calculating the intersection of the set of current contextual labels, the current set of time (current time window), and the set of next activity labels. The “deep” descriptor refers to the layers of contextual data, as opposed to a “shallow” model which only uses the event sequence data. We posit that this method is extensible by adding more sets of contextual data which would improve the accuracy further.

The set of sequential time values indicating the current time window, simply the start and end for the time window:

$$T = \{t_1, t_2\}$$

$$N = \{n_1, \dots, n_n\}$$

The set of potential activities where each element is an activity that could occur next. This could be a shortlist output from a sequential prediction model, e.g. a set of causal activities from a DBN. Or a less optimal list of all activity labels:

$$C = \{c_1, \dots, c_n\}$$

The set of contextual labels present in the current time window, as obtained from the smart environment data:

$$P = \{p_1, \dots, p_n\}$$

The set of person metadata for the person being observed, as obtained from the smart environment data:

$$X = \{x_1, \dots, x_n\}$$

Any other sets of data, so that the system is extensible:

$$\mathcal{D} = [[t_n, n \in N, c \in C, P, X]]$$

Previously observed data, the historic dataset of the smart environment:

We then select the event data from the specified time window, and calculate the intersection value of the other sets of contextual data within that time window, to determine the likelihood of the next activity label occurring after the current activity label.

$$\begin{aligned} f(\mathcal{D}, N, T, C, P, X) &= \prod_{n \in N} (E \subset \mathcal{D} : e_1 \geq t_1 \wedge e_n \leq t_2) \\ &= Pr(n \cap E \cap C \cap P \dots \cap X) \end{aligned}$$

In the equation above, where  $N$  is the set of next activities (e.g. a shortlist of known activity labels output by a DBN), and  $C$  is the set of current contextual labels, and  $T$  is the current time window (start and end time), it is possible to take the set  $N$  of next activity labels and return a set of probabilities of the likelihood that the activity label will be next, by

<sup>1</sup> SBM data and code, <https://data.mendeley.com/v1/datasets/jm28kgt7tm>

calculating the probability intersection of contextual labels, next activity, and person metadata. A large overlap or high intersection between the sets outputs a high probability, and inversely lower overlap produces a lower probability. In plain English, if these groups of data are not usually seen together at this time then the activity label has a low probability of being the next activity—but if all the contextual labels are seen at this time and location with this person, then it has a high probability that activity label is going to occur next. The activity  $n \in N$  with the highest probability is returned as the inferred next activity.

This model can be extended by adding more sets of different kinds of data to the probability intersection. Lastly, if the next activity labels  $N$  do not intersect with the set of time values (the current time window)  $T$ , then the probability must be zero,  $Pr(N \cap T) = 0$ , as those labels have never been observed in the time window previously.

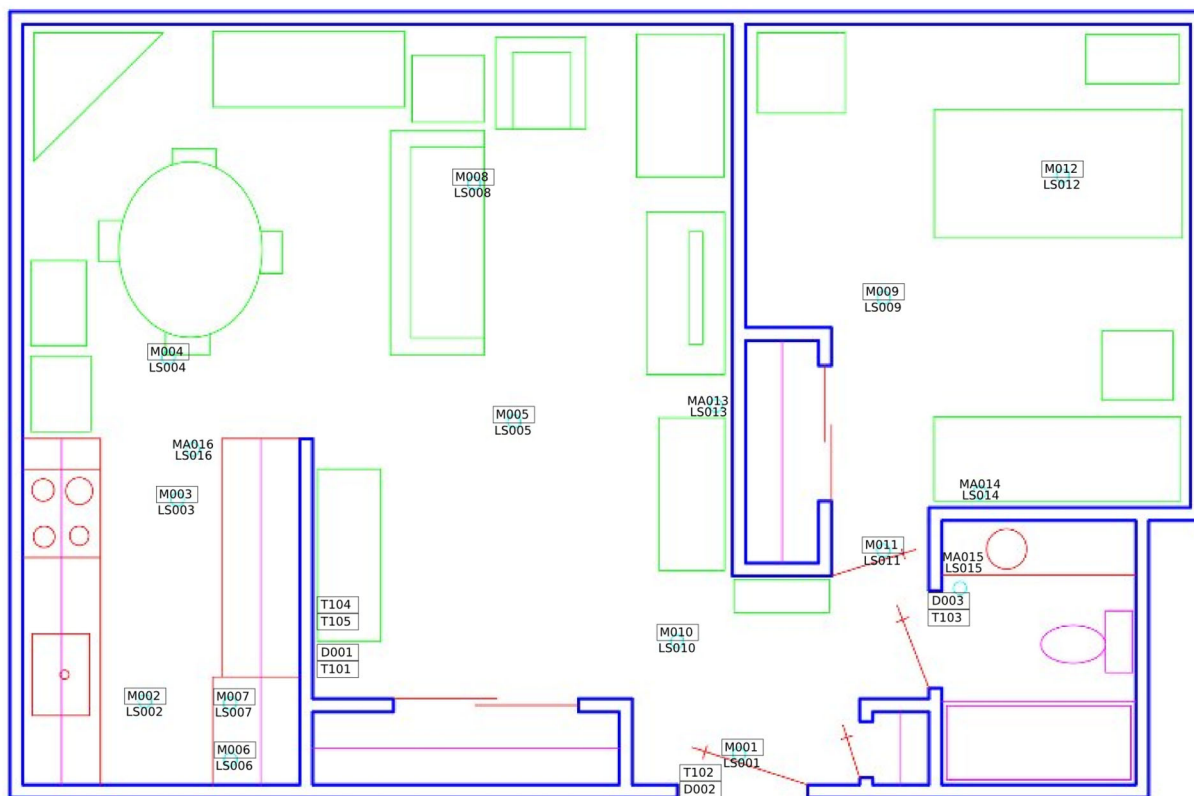
We will test this hypothesis using the original CASAS datasets from the CRAFFT model experiments, by comparing the accuracy of next activity inferences to the actual next activity label in the dataset.

## 2.2 Dataset

For our experiment, we used the CASAS smart home dataset [4] made publicly available and open access by the Centre of Advanced Studies in Adaptive Systems at Washington State University (<https://archive.md/EQRrK>). This is the same dataset used by the CRAFFT research and we are able to use the data from three apartments to compare our results to the CRAFFT benchmark. The CASAS dataset uses several smart home apartments fitted with sensors and single occupants who continuously generate data. We selected three single occupancy apartments (HH101, HH102, and HH103) to simplify our inference model, as multiple occupancy prediction is out of scope of this research. In Fig. 1, we can see the floorplan of one of the apartments with sensor locations throughout the rooms. These are the data points we use for our event states and constitute the activity labels in our dataset, e.g. Leave\_Home, Enter\_Home, Watch\_TV, Cook.

Each dataset subset is split into a 90/10 train test split, where the ‘training’ is done on-the-fly by using the train dataset as the historic observable data for the smart home, and the test dataset as the ‘live’ data to assess accuracy of next activity prediction.

Table 2 shows a sample of the CASAS dataset for the HH101 apartment shown in Fig. 1, which is floor plan of the



**Fig. 1** CASAS smart home HH101 floor plan with sensor locations. Credit: CASAS; casas.wsu.edu

apartment, with the event timestamp, the sensor code—which can also be seen on the floor plan, and the event code for that sensor, e.g. a door opening or closing, or a device being switched on. These data are the input for both the CRAFFT study, and our SBM proof-of-concept. The contextual data extracted and used as input variables in the SBM are: Current activity, Preceding event, Time window, Day of the week, Weekend or weekday, Location in the apartment.

### 2.3 Research method

To test our hypothesis, we used an experimental prototyping methodology to develop a proof-of-concept. To enable this we created a test framework, as shown in Fig. 2, in which we loaded the CASAS smart home datasets in a 90/10 train test split into separate tables. Using Python we coded the SBM to read the events sequentially from the test table, calculate the intersections of the data groups based on the train table, then output the event with the highest value as the predicted next activity. This value was logged in a results table in the

database for comparison with the test table for accuracy at the end.

The SBM is designed to take any list of possible next activities. We envisage this as a secondary model to existing predictive models, such as a DBN—that would shortlist potential next activities. However, the SBM can take any number of next activities as input and will calculate the probability of occurrence for each. As there was no requirement for performance in this initial experiment for the proof-of-concept (i.e. this isn't yet a production system) we can simply pass all known activities for each prediction step and calculate the probabilities and select the highest as the predicted next activity.

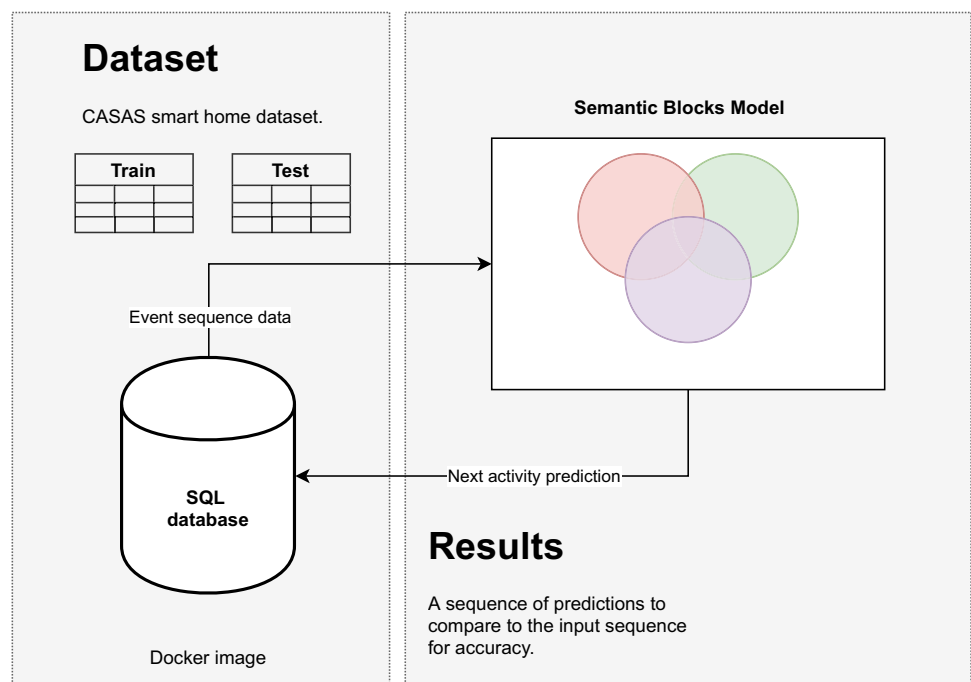
The time complexity of our model is  $O(n)$  linear time dependent on the size of the input, e.g. the length of the list of next activities. Combining the SBM with another model, e.g. a DBN, would optimise the performance by shortlisting or reducing the list of next activities.

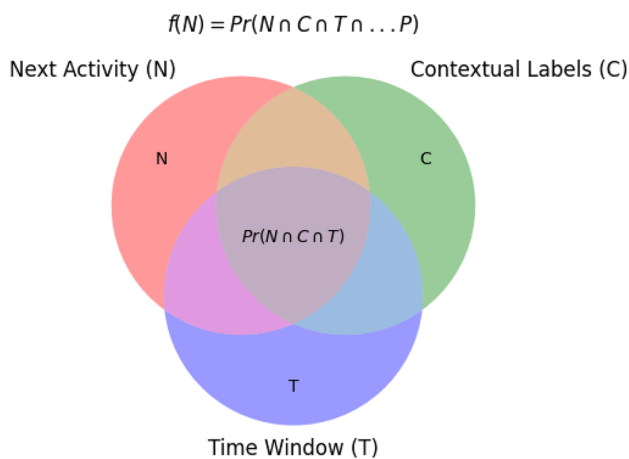
The probability intersection for our sets of contextual data and activities implements the equation in Sect. 2.1 in Python.

**Table 2** CASAS smart home HH101 dataset sample

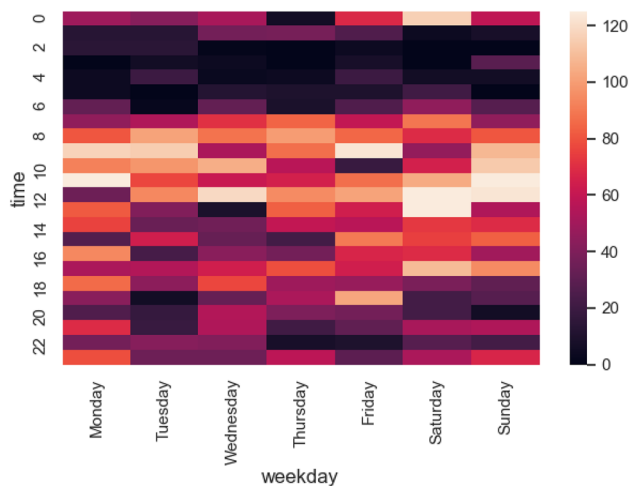
Event timestamp	Sensor	Event	Label	Location
2012-07-20 11:09:59.128578	MA015	OFF	Toilet="end"	Bathroom
2012-07-20 11:50:04.029817	M008	ON	Relax="begin"	Lounge
2012-07-20 11:55:40.958827	M008	OFF	Relax="end"	Lounge
2012-07-20 12:00:43.445554	MA015	ON	Personal_Hygiene	Bathroom
2012-07-20 12:00:44.634408	MA015	OFF	Personal_Hygiene	Bathroom
2012-07-20 12:00:45.129838	MA015	ON	Personal_Hygiene="begin"	Bathroom
2012-07-20 12:05:30.659402	MA015	OFF	Personal_Hygiene="end"	Bathroom

**Fig. 2** Semantic blocks model: test framework

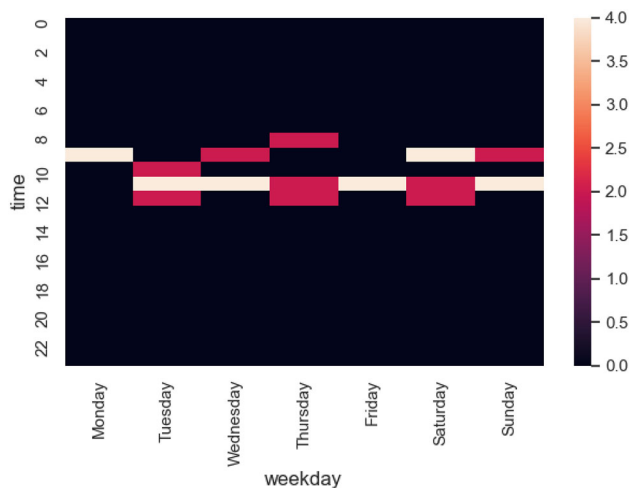




**Fig. 3** A simplified representation of intersecting contextual data



**Fig. 4** CASAS smart home HH101 heatmap of all activities



**Fig. 5** CASAS smart home HH101 heatmap of activity label ‘bathe’

**Table 3** Apartment activity time windows

Time window	Range
Morning	0700–1100
Afternoon	1200–1700
Evening	1800–2100
Night	2200–0600

For each activity observed in the smart home sequence, a time window is looked up, as shown in Table 3. The data from the train table (the historic dataset) is selected within this time window (semantic block). Any contextual data present with the current activity is passed to the model. The SBM then calculates for each known activity (or a shortlist if used in conjunction with another predictive model) how often each potential next activity occurs after the current one, how often it occurs in that time window, and how often it is present with the contextual labels. All these bits of intersecting data together, as illustrated in Fig. 3, produce a probability score. The highest score of all processed activities is the predicted value.

The CRAFFT algorithm uses the current state (activity) along with the contextual data: activity location, activity time, activity day, and preceding activity from the CASAS dataset to augment its prediction calculation. Similarly we use as our contextual data: location, time, day, and also preceding activity. We chose to base our model on CRAFFT for its interesting use of contextual data. We determined that it can be used to create a model that extends the concept of CRAFFT; improving its performance but also designing it to be coupled with DBN or DL methods to give even better predictive output than the standalone SBM itself.

To determine where to demarcate the time windows we generated heatmaps of each dataset to elucidate any patterns in the occupant’s behaviour. In Fig. 4, we include all the activity events in the dataset, and we can clearly see bands of activity and inactivity. For example, the dark band between 0100 and 0600 would indicate sleeping. In the morning, there is the maximum amount of activity, which then reduces during the day. If we single out one activity, such as ‘bathe’ we can see in Fig. 5 how this activity always occurs within the same time window between 0800 and 1200. Ideally we would dynamically generate time windows for each individual occupant perhaps using a clustering algorithm such as K-means or K-Nearest Neighbour. However, this is out of scope for our proof-of-concept, so we manually selected time windows based on the average demarcation of the activity heatmaps, which are shown in Table 3.

To summarise, the SBM is a three stage process:

1. Calculate environment specific contextual time windows for activities—what we term ‘semantic blocks’.

**Table 4** Semantic blocks model sample output

Dataset	Time_window	Predicted_next_label	Inf_score	Actual_next_label
hh101	22:00:00–06:59:59	Personal_Hygiene	0.348281	Evening_Meds="begin"
hh101	22:00:00–06:59:59	Evening_Meds="end"	1.05856	Evening_Meds="end"
hh101	22:00:00–06:59:59	Personal_Hygiene	0.305946	Dress="begin"
hh101	22:00:00–06:59:59	Dress="end"	1.09628	Dress="end"
hh101	22:00:00–06:59:59	Personal_Hygiene	0.44534	Personal_Hygiene
hh101	22:00:00–06:59:59	Personal_Hygiene	0.899305	Personal_Hygiene
hh101	22:00:00–06:59:59	Personal_Hygiene	0.899305	Personal_Hygiene="begin"

2. Extract the necessary data from the previously observed smart home dataset: data with a time window based on the current activity time.
3. Intersect the sets of data: contextual data, time window, current activity, and previous activity. The greater the intersection, the higher is the score. The next activity with the highest score is the outputted predicted value.

### 3 Results

When testing the SBM proof-of-concept, we first ran it with no contextual data, and only used the time window to narrow down the next activity. This gave a baseline indication of how the time window worked on its own. As you can see in Table 5, the average accuracy for this across the three apartment datasets was 13.456%. A very low accuracy but slightly better than random selection. When we run the experiment with all the contextual data added (time window, preceding event, day of the week, weekend or weekday, location in the apartment) we can see, again Table 5, that we have much better accuracy. As high as 82.3991% for the HH102 apartment dataset, and an overall average accuracy of 74.552%.

Table 4 shows sample output of the SBM with the predicted next activity label, the inference score generated internally where a higher score is greater confidence, and the actual next label from the dataset. The inference score, or *inf\_score*, is generated by calculating the how much the contextual data and each possible next activity label intersects. The higher the score is, the more the sets of contextual data intersect, and, therefore, the more frequently these contextual conditions are observed together in the historic data from the smart home. The SBM uses this score to determine which activity has the highest likelihood to be next.

The results data demonstrate that further work could be done to generalise the model for higher level activity events. In Table 4, we can see that some of the *predicted\_next\_label* and *actual\_next\_label* data have a partial match, e.g. *Personal\_Hygiene* and *Personal\_Hygiene="begin"*. These have been classed as an incorrect prediction for this experiment; however, it shows that the SBM is very close to being correct

**Table 5** Semantic blocks model experimental results

Test phase	Apartment dataset	Accuracy (%)
Time window with no contextual data	HH101	3.2557
	HH102	27.2411
	HH103	9.8706
	<b>Average accuracy: 13.456</b>	
Time window with added contextual data	HH101	65.0407
	HH102	82.3991
	HH103	76.2162
	<b>Average accuracy: 74.552</b>	

in a large number of cases, and highly encouraging that the model could be further improved.

### 4 Conclusion

The CRAFT benchmark in Table 1 produced an average accuracy of 66.91%. Our semantic blocks model outperformed this over the same datasets with an average accuracy of 74.552%. A significant increase. Indeed, as you can see in the results comparisons, in Table 6, the SBM improved on the prediction accuracies of each dataset it was run against. These results are highly encouraging, and we believe that SBM accuracy could be improved and refined even more with added contextual data, or when coupled with another prediction model such as a Dynamic Bayesian Network. In Sect. 5, we detail improvements that could be made in future experiments. Dynamic clustering of the data over longer observed time periods for apartments occupants would generate the most optimal time windows for that smart environment. We have used fixed time windows for the purposes of our proof-of-concept which may not have been tuned perfectly to the data.

To reduce overfitting of the models the training data from multiple different environments could be combined to produce a generalised model for activity prediction. Similarly, to

**Table 6** Semantic blocks model (SBM) vs CRAFFT benchmark prediction accuracy results on the CASAS dataset

	CRAFFT benchmark	Semantic blocks model (SBM)
HH101	64.18	65.0407
HH102	74.57	82.3991
HH103	61.98	76.2162
Average	<b>66.91</b>	<b>74.552</b>

produce greater flexibility for the model, a probability distribution of potential next activities could be output to allow upstream systems to select a group of possible activities. Internally, the SBM already produces this distribution, but simply selects the one with the highest score to output as the prediction.

There will always be a limit to how accurate prediction can be for stochastic events—there will almost always be a small element of randomness that can't be accounted for. However, by combining DBN or DL models with a model like the SBM which uses time windowing and contextual data, plus the optimisations through ML clustering, it appears possible to create a system which is able to infer activity with a high level of accuracy.

To conclude, we can confidently state that the hypothesis laid out in Sect. 2.1, that human activity prediction in smart environments can be improved by enriching the data with contextual labels and a time window, has been supported. When comparing our results to the closest benchmark available—the CRAFFT algorithm—which uses similar contextual data on the same dataset, but without a time window, we yield a higher prediction accuracy. The proof-of-concept developed and tested here clearly demonstrates that clustering data into time windows, or semantic blocks, is a valid method for improving accuracy of human activity prediction.

## 5 Future work

Our successful results indicate that further work could be done in this area to refine and improve upon this concept. Areas for further work by the ourselves and the research community include:

- Automated dynamic clustering of the smart environment data, so that the most optimal time windows are generated, bespoke, for each environment based on previously observed data. Exploring the use of machine learning clustering algorithms would be a good starting point for this work.
- Adding more contextual data. The SBM has been designed to be extensible. Such data may include person

metadata (age, adult/child, gender, religion, state/emotions), calendar metadata (daily events, calendar events, ritual/cultural/geographical events) that could influence a person's activities, and multi-occupancy environment data where interactions with other people are taken into account.

- Experiment with a data completeness coefficient, that would alter the prediction calculation based on how many of the contextual sets of data were present and complete.
- Pairing with a DBN to account for causality, and shortlisting of possible next activities.
- Pairing with a Deep Learning (DL) model for shortlisting of possible next activities.
- Comparing the SBM results to a modified CRAFFT model that includes time-windowed data.
- Analyse the SBM performance when transferring and applying it to other smart home datasets.
- Determine the SBM accuracy for a range of parameters including size of the time window, and frequency of activity events.

## Declarations

**Conflict of interest** We declare that there are no conflicts of interest and no financial contributions or funding were received for the production of this research.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Benmansour A, Bouchachia A, Feham M (2015) Multioccupant activity recognition in pervasive smart home environments. *ACM Comput Surv* 48:1–36
2. Chen H, Finin T, Joshi A (2003) An ontology for context-aware pervasive computing environments. *Knowl Eng Rev* 18:197–207
3. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. *ArXiv Computer Science*, pages 1–9,
4. Cook DJ, Crandall AS, Thomas BL, Krishnan NC (2013) Casas: a smart home in a box. *Computer* 46:62–69
5. Danner UN, Aarts H, De Vries NK (2008) Habit vs. intention in the prediction of future behaviour: the role of frequency, context stability and mental accessibility of past behaviour. *Br J Soc Psychol* 47:245–265



6. Dinh TL, Thi TTP, Pham-Nguyen C, Nam LNH (2022) A knowledge-based model for context-aware smart service systems. *J Inf Telecommun* 6:141–162
7. Ducatel K, Bogdanowicz M, Scapolo F, Leijten J, Burgelman J-C (2001) Istag scenarios for ambient intelligence in 2010. Society, page 58
8. Dunne R, Morris T, Harper S (2022) A survey of ambient intelligence. *ACM Comput Surv* 54:1–27
9. Elbayoudi A, Lotfi A, Langensiepen C, Appiah K (2016) Determining behavioural trends in an ambient intelligence environment. pages 1–8. ACM Press
10. Elbayoudi A, Lotfi A, Langensiepen C, Appiah K (2017) Trend analysis techniques in forecasting human behaviour evolution. Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments-PETRA '17, pages 293–299
11. Faieq S, Front A, Saidi R, Ghazi HE, Rahmani MD (2019) A context-aware recommendation-based system for service composition in smart environments. *SOCA* 13:341–355
12. Forkan ARM, Khalil I, Tari Z, Fofou S, Bouras A (2015) A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living. *Pattern Recogn* 48:628–641
13. Je TM, Chao lin W, Sipun Kumar P, Yifei X, Li Ting Y, Li chen F, Yi chong Z (2016) Context aware activity prediction using human behavior pattern in real smart home environments
14. Jiang H, Cai C, Ma X, Yang Y, Liu J (2018) Smart home based on wifi sensing: a survey. *IEEE Access* 6:13317–13325
15. Kang W, Shin D, Shin D (2010) Detecting and predicting of abnormal behavior using hierarchical markov model in smart home network. pages 410–414. IEEE
16. Machado A, Maran V, Augustin I, Wives LK, de Oliveira JPM (2017) Reactive, proactive, and extensible situation-awareness in ambient assisted living. *Expert Syst Appl* 76:21–35
17. Nazerfard E, Cook DJ (2015) Crafft: an activity prediction model based on bayesian networks. *J Ambient Intell Humaniz Comput* 6:193–205
18. Paola AD, Ferraro P, Gaglio S, Re GL, Das SK (2017) An adaptive bayesian system for context-aware data fusion in smart environments. *IEEE Trans Mob Comput* 16:1502–1515
19. Sadri F (2011) Ambient intelligence: a survey. *ACM Comput Surv* 43:1–66
20. Shin D, Lee G, Shin D, Shin D (2016) System architecture using human interaction markup language for context awareness in home network. *Multimed Tools Appl* 75:15199–15209
21. Sánchez VG, Lysaker OM, Skeie NO (2019) Human behaviour modelling for welfare technology using hidden markov models. *Pattern Recogn Lett* 137:71–79
22. Survadevara NK, Mukhopadhyay SC, Rayudu RK (2012) Applying sarima time series to forecast sleeping activity for wellness model of elderly monitoring in smart home. In: Proceedings of the International Conference on Sensing Technology, ICST, pages 157–162
23. Tax N (2018) Human activity prediction in smart home environments with lstm neural networks. In: Proceedings-2018 International Conference on Intelligent Environments, IE 2018
24. Tonkin Emma L, Woznowski Przemyslaw R (2018) Activities of daily living ontology for ubiquitous systems. 2018 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2018, pages 573–578
25. Vengatesan K, Kumar Abhishek, Karuppuchamy V, Shaktivel R, Singhal A (2019) Face recognition of identical twins based on support vector machine classifier. In: Proceedings of the 3rd International Conference on I-SMAC IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2019, pages 577–580
26. Wu E, Zhang P, Lu T, Gu H, Gu N (2016) Behavior prediction using an improved hidden markov model to support people with disabilities in smart homes. pages 560–565. IEEE
27. Ye J, Stevenson G, Dobson S (2016) Detecting abnormal events on binary sensors in smart home environments. *Pervasive Mobile Comput*

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.