Interpretable Machine Learning for In-Home Mild Cognitive Impairment Detection

Budhitama Subagdja[1], Shanthoshigaa D[1], Ah-Hwee Tan[1], Iris Rawtaer[2]
[1]School of Computing and Information Systems
Singapore Management University, Singapore
[2]Department of Psychiatry, Sengkang General Hospital
Singhealth Duke NUS Academic Medical Centre, Singapore
{budhitamas,shanthod,ahtan}@smu.edu.sg, iris.rawtaer@singhealth.com.sg

Abstract

This paper introduces a novel system for in-home cognitive health assessment using ambient sensors and a machine learning technology that can robustly detect mild cognitive impairment (MCI) despite limited available data. The learned model can explain the aspects of individuals' daily lives led to the prediction, while reliably predicting MCI, providing more insights to healthcare workers for further clinical interventions. We developed the robust transparent machine learning model, based on fusion adaptive resonance theory (Fusion ART) neural network to learn individuals' daily patterns of activity from continuous sensor data in terms of a suite of digital biomarkers reflecting four key domains: physical, daily activity, cognitive engagement, and sleep patterns. Based on a longitudinal study of over one hundred participants, deployed with non-intrusive sensors in their homes to undergo parallel clinical evaluation across a period of five years, our model successfully identified individuals with MCI, achieving high predictive accuracy regardless the noisy and sparse availability of data. As a transparent neural network, the learned model can also be interpreted as classification rules to distinguish MCI from normal cognition (NC) cases based on the digital biomarkers. These results demonstrate that passively collected, sensor-derived digital biomarkers can be leveraged to indicate cognitive status and potentially providing clinically meaningful insights on the impairment conditions. We also discuss the practical challenges and lessons learned from this real-world deployment to inform future large-scale implementations of such AI-driven health monitoring systems.

Introduction

Sensor-based systems combined with machine learning are a feasible solution for the early detection of cognitive decline by monitoring individuals' daily routines in the community (Quek, Heikkonen, and Lau 2023). However, a significant gap exists between research and real-world application. Many studies on digital biomarkers are conducted in controlled lab environments (Dawadi et al. 2013) or rely on specific tasks like games (Gielis et al. 2021), which do not fully capture daily life. Even promising approaches that analyze daily activities to detect cognitive changes (Lee et al. 2020; Narasimhan et al. 2022; Lee et al. 2023) often require participants to provide manual activity logs for model training. This dependency on self-reporting is time-consuming and error-prone, hindering the practical deployment of these technologies in real-world settings. However, despite the potentials of sensors-based systems, continuously integrating heterogeneous data from individuals' daily routines remains challenging as they can be noisy, containing missing information, and creating highly irregular data distribution. Moreover, some people may concern about their privacy related to recording of images, audio, or video data, limiting the kind of modality to be captured.

This may add complications in ensuring that the model learned can still be useful for clinical use. On the other hand, how the changes in cognitive states are reflected in individuals' daily routines also remains understudied. Beyond predicting cognitive decline, the learned model should also be leveraged to explain the relationship between the cognitive condition and the patterns in the individual's daily routines. This allows healthcare or medical workers to gain more insight into individuals' cognitive conditions and timely conduct clinical interventions when necessary.

This paper presents an in-home cognitive health assessment system for early detection of mild cognitive impairment (MCI)—a key window for intervention—that robustly distinguishes daily routines of participants with MCI from the ones with normal cognition (NC), and transparently indicating the aspects of daily lives that base the prediction. The system is developed as a part of the SINEW (Sensors IN-home for Elder Wellbeing) project (Rawtaer et al. 2020), a large-scale study leverages non-intrusive ambient sensors and artificial intelligence (AI) in a real life home-based setting to analyze cognitive health. As the first longitudinal study of its kind, SINEW captures sensor data from a sizable cohort of real homes surpassing the scale of similar studies (Teh, Rawtaer, and Tan 2022; Kwon et al. 2021). Spanning almost five years, since the beginning of data collection (2020-2025), and still on-going, this study follows more than 100 independently living seniors who were recruited with either NC or MCI. To determine cognitive status over time, these passive observations are complemented by annual, in-depth neurocognitive assessments and consensus panel diagnoses. In this way, cognitive declines and the progression from a healthy cognitive state (from NC to MCI) can be identified.

To realize the robust machine learning model, we applied Fusion Adaptive Resonance Theory (Fusion ART) (Tan et al. 2019) neural network model to learn the characteristics of the participants' daily routines with MCI and NC. Derived from sensor data and established in the study, the characteristics are based on a set of digital biomarkers comprising four principal aspects of daily life, namely the physical, activity, cognitive, and sleep domains. Fusion ART employs *complement coding*, as an encoding scheme supporting attributes generalization and the masking of missing or unknown values. This encoding method let the neural network robustly learn the prediction model despite the highly missing rate and sparse distribution of data availability. Furthermore, the learned patterns of daily biomarkers are explicitly profiled and clustered in Fusion ART, informing not just about the possibility of the moment the participant has MCI, but also the likelihood that the participant has MCI (or NC). Each cluster can be read out to explicitly show the pattern of daily routines in terms of prediction rules given the status of biomarkers as the antecedents.

Validated by clinical assessment findings, our experiments show that machine learning models, particularly Fusion Adaptive Resonance Theory (Fusion ART) (Tan et al. 2019), are capable of identifying MCI based on these biomarkers, achieving reasonably high $F1$ scores. Extending a previous work with a smaller set of data in a relatively brief period of time (Tan et al. 2024), in this paper, the machine learning algorithms are employed to predict MCI cases based on the labeled biomarkers from the entire study cohorts from the year 2020 to the middle of 2025. Beyond MCI prediction, we also demonstrate how Fusion ART, as a transparent neural network model, can be leveraged to explain its prediction results in terms of classification rules. The rules explicitly indicate which aspects within the four domains of digital biomarkers can lead to MCI cases in general, providing insights to healthcare workers or clinicians about the relationships between MCI and aspects of daily lives.

This work's key technical contributions are the following: (1) a robust MCI detection system employing Fusion ART is developed that can reliably distinguish MCI from NC cases using the encoding scheme to tackle missing values and sparse data distribution; (2) applying the explicit biomarkers profiles as clusters in Fusion ART to produce the prediction scores or the likelihoods of a cognitive status in day-to-day basis; and (3) applying a trained Fusion ART network from the large cohort in the study to extract the rules of MCI prediction over noisy and sparsely available data.

Clinical Protocol and Methodology

The participants in SINEW project were recruited from the community. They are community-dwelling seniors above 65 years who were living alone and functioning independently. They also must be able to communicate and provide written consent in English and/or Mandarin. Once the participant decides to take part in the study, a baseline screening and assessment is conducted. Every year, comprehensive assessments are conducted to every participant. Participants completed comprehensive assessments using validated instruments to measure various aspects of health including functional, psychosocial, mental, and cognitive well-being. These validated assessments were administered by trained research assistants. The instruments used to understand function and social vulnerability included the Lawton's iADL scale (Lawton and Brody 1969), the Friendship Scale (Hawthorne 2006), and the Lubben's Social Network Scale (Lubben 1988). Mental health and sleep was assessed using the Zung Self-Rating Depression Scale (Zung 1965), the 15-item Geriatric Depression Scale (Yesavage et al. 1982), Pittsburgh Sleep Quality Index (PSQI) (Buysse et al. 1989), and the Apathy Inventory (Robert et al. 2002). Lastly, for cognitive health, a series of neurocognitive tests adapted to local populations were conducted and repeated annually.

Type of Sensors	Sensor ID / Location	Value / Interpretation
Motion Sensors	M-01 (Living Room)	No Motion: 0
	M-02 (Kitchen)	Motion: 255
	M-03 (Bedroom)	(updated every 30 sec)
Contact Sensors	D-01 (Main Door)	Closed : 0
	D-02	Opened : 255
	(Medication Box)	Whenever triggered
Pressure Sensor	D-03 (Bed)	Pressure on: 0
		Pressure off: 255
		(whenever triggered)
Beacon Sensors	B-01 (Key)	Near Gateway: Timestamp
		updated every 4 mins
	B-02 (Wallet)	Out of range: No update
Wearable Sensor	Step Count	Number of steps
	Heart Rate (bpm)	Beats per minute

Table 1: A summary of the sensors deployed and their characteristics.

Sensor Network and Data Collection

Data collection in the SINEW project is conducted through an Internet of Things (IoT) architecture which consists of a combination of wireless devices, an intelligent gateway, and a cloud server that serves as the back-end of the home-based sensor network systems. Each SINEW sensor network system is designed to collect the raw sensor readings data from the home of an individual participant and transmit the data to the cloud-based database via in-house gateway device. The raw sensor readings are then further processed to extract the digital biomarker features. Currently, all the machine learning algorithms for MCI prediction run on the main server in a centralized manner.

Table 1 contains a list of the wireless sensors deployed, consisting of three motion sensors for living room, kitchen, and bedroom, two contact sensors for main door and medication box, one pressure-based bed sensor, two beacon sensors attached to key and wallet, and one wearable device. As illustrated in Figure 1, in every home, wireless contact and motion sensing devices are strategically positioned in various locations, encompassing the main door, medication box, living room, kitchen, and bedroom. In addition, the wearable device is to be worn by the participant at all times and the beacons are tagged to his/her key and wallet. Due to con-

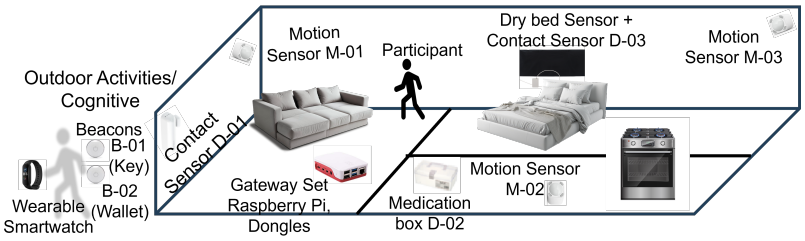

Figure 1: An illustration of sensors' placement in the participants' homes.

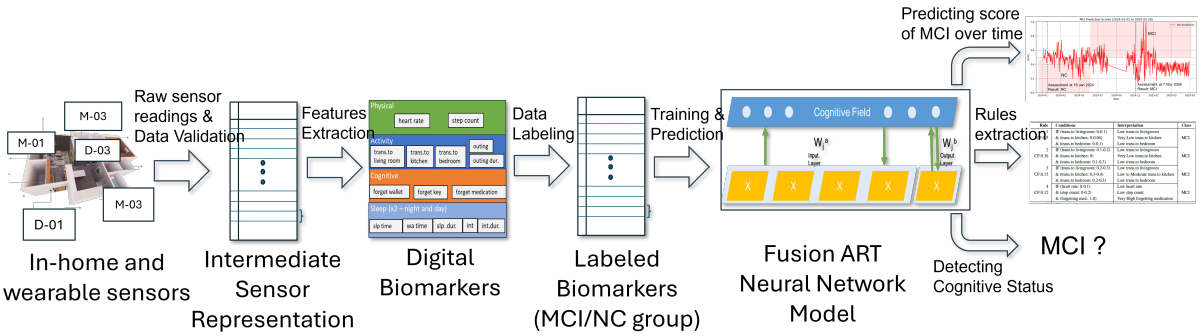

Figure 2: The overall process flow of the system to detect MCI, predict MCI scores, and extract classification rules.

cerns of the participants' privacy, no sensor devices capturing image, audio, or video data are included in the study.

The home gateway device is based on a single-board computer (e.g Raspberry Pi) and uses a combination of portable internet modem-WiFi hub dongles with a wireless receiver for sensor system. The sensors are set up in a way that they are able to periodically sense the environmental cues and send the information to the gateway device.

Digital Biomarker Extraction

The data sent by the in-home sensors are raw sensor data in the sense that each entry in the sensor database indicates a single reading from a sensor at a particular time in one participant's home. As such, high level biomarker features need to be extracted from the raw sensor readings to enable further analysis and prediction.

Sensor Data and Digital Biomarkers

A record of the sensor data consists of four attributes, including *participant's id* as the unique identifier of the participant, *timestamp* indicating the date and time of the sensor reading, *sensor id* as the identifier of the sensor, and *sensor value* as the reading of the sensor. The *timestamp* attribute is the date and time at which the home gateway receives the sensor reading message before submitting to the cloud server. Digital biomarkers are designed to capture the high level information relating to the daily routines and activities of each participant. In this study, the biomarkers are used to capture different aspects of daily living that include **physical**, **activity**, **cognitive**, and **sleep**. Specifically, a daily record consisting of a total of 15 biomarker features organized according to the four basic aspects of daily living as follows:

- **Physical Biomarkers** track bodily movement produced by skeletal muscles which can be monitored using wearable sensors (Rawtaer et al. 2020; Taraldsen et al. 2012). Two features are captured in this study, namely the *heart rate* (beat per minute) and the *step count*.
- **Activity Biomarkers** track the person's activity-related behavior in relation to movement to spatial zones/locations in daily lives. Five biomarkers are captured for zonal movement, including the number of *transitions to bedroom*, *transitions to kitchen*, *transition to living room*, *outing* (number of times going outside), and the *outing duration* (total time spent outside) in a day.
- **Cognitive Biomarkers** monitors the participant's memory performance which may be indicative of cognitive decline in daily activities. There are three biomarker features related to forgetfulness, namely the number of times *forgetting wallet* (situations wherein the participant forgets to bring his/her wallet when going outdoor), *forgetting keys* (the participant forgets to bring the door key), and *forgetting medication* (forgetting to take medication according to the prescribed frequency).
- **Sleep Biomarkers** capture the statistics of sleep patterns of the participant (Rawtaer et al. 2020; Chen et al. 2019). In this paper, only night time sleep consisting of five features namely *sleep time*, *wake time*, *sleep duration*, number of *sleep interruption*, and *sleep interruption duration*.

Daily Biomarker Extraction and Labeling

To extract the daily digital biomarkers from the raw sensor data, a data validation process is conducted to create an intermediate sensor representation simplifying further extrac-

tion of features for the digital biomarkers as shown in the beginning of the system flow (Figure 2). The validation step consists of *data purging* operation to filter away invalid raw sensor data (due to a power down of the gateway or participant unavailability) and *features pre-processing* to sequentially aggregate the valid sensor data into intervals of δ minutes (currently we use $\delta = 5$). *Biomarker feature extraction* then follows to extract the biomarkers based on predefined pattern matching rules to detect the participant's state of behavior.

To enable the training of machine learning models with the digital biomarkers, each record needs to be labeled with the cognitive status of the participant so that the machine learning algorithms can associate the daily data with the corresponding labels. As illustrated in Figure 2, the extracted digital biomarkers go through the *data labeling* process to produce *labeled biomarkers* for training and prediction. In this case, each record is assigned with a label based on the result of the closest annual clinical assessment. Specifically, each record dated within the six month period before and after the date of a clinical assessment is labeled based on its result (MCI or NC) or the data will be excluded from the training otherwise. However, if the date of the record lies in between two consecutive clinical assessments with the same results (both indicating MCI or NC), the record is still included in the training although it is too distant from both assessment dates.

MCI Prediction Model

The biomarker-based predictive model is implemented using Fusion ART employing a multi-channel architecture, comprising a cognitive field F_2^c and a fixed number of input fields F_1^{ck}.

In this paper, we present a two-channel Fusion ART architecture, comprising a cognitive field F_2^c and two input fields F_1^{c1} and F_1^{c2} (Figure 3). The first input field F_1^{c1} encodes the input feature vector, representing digital biomarkers, while the second input field F_1^{c2} encodes the corresponding cognitive class labels (MCI or NC).

Each input vector $X^a = (a_1, a_2, ..., a_N)$ consists of normalized values within the range [0,1] representing biomarker and activity features. The class label vector $X^b = (b_1, b_2, ..., b_M)$ is similarly encoded. To improve learning stability and reduce code proliferation, complement coding is applied, generating input vectors of the form $I = (X^a, X^{ac})$, where $a_i^c = 1 - a_i$. For features with missing values, the input is encoded as $(1, 1)$ to preserve existing weights during learning (Teh, Rawtaer, and Tan 2022).

During training, Fusion ART evaluates the similarity between input vectors and stored category nodes through a bottom-up activation process and fuzzy matching operations. A competitive selection process selects the most activated node, which undergoes template learning based on the fuzzy AND operation. Learning dynamics are controlled by model parameters including the learning rate (β), vigilance threshold (ρ), and channel-specific contribution weights (γ). The architecture dynamically adds new nodes to represent new input patterns, making the learning *instance-based* and *incremental*.

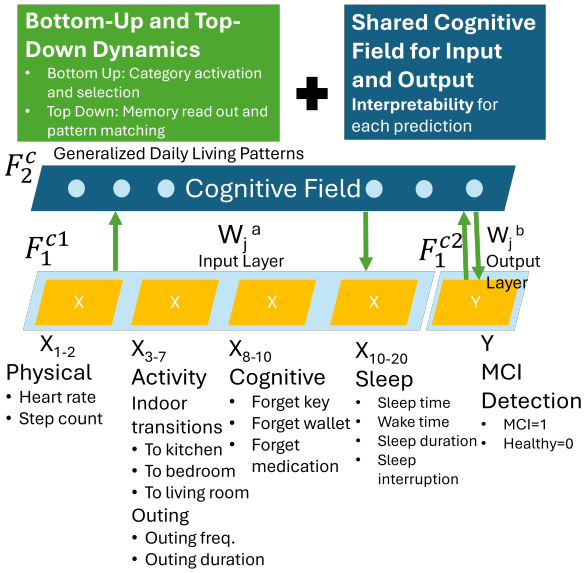

Figure 3: Fusion ART Neural Network Architecture

Prediction can be conducted through the bottom-up activation and competitive selection to select the most activated node before reading out the output layer F_1^{c2} as the predicted label. Alternatively, a prediction score can be obtained for a label (MCI or NC) proportional to the number of top activated nodes with the label weighted according to their order of level of activation. To support interpretability, each category node in F_2^c is assigned a *confidence factor* (CF) computed from its usage frequency and predictive accuracy. Nodes with higher confidence factors are translated into symbolic IF-THEN rules, representing behavioral patterns that distinguish between MCI and NC classes. This rule-based representation enables transparent reasoning of the model's predictions based on real-world digital biomarkers and ADL data.

Data Validation Experiments

We have conducted extensive empirical experiments wherein predictive models are built using various machine learning techniques based on the digital biomarker data collected. From the sensor data collected over a period of five years from January 2020 to May 2025, we obtained a total of 65,075 daily biomarker records from 117 participants. After filtering away those records without cognitive status label (no cognitive assessment or baseline result available), we obtained a data set containing 59,214 daily biomarker records from 81 participants.

Experiments with Monthly Averaged Data

We aggregated the daily biomarkers data by averaging them into monthly biomarkers records for the experiment. We obtained a total of 2631 monthly averaged biomarker records. As shown in Table 2, the monthly feature data obtained was still relatively sparse and noisy with 81% missing rates for

biomarkers		averaged missing rate (%)	
aspect	feature	weekly	monthly
activity	trans.to living room	13	10
	trans. to kitchen	13	10
	trans. to bedroom	13	10
	outing	19	16
	out.duration	29	16
cognitive	forget med.	48	30
	forget keys	59	55
	forget wallet	48	45
sleep	sleep duration	76	67
	sleep interrupt.	76	67
	sleep time	76	67
	wake time	76	67
	sleep interr.duration	76	67
physical	heart rate	83	81
	step count	83	81

Table 2: Average distribution of weekly and monthly biomarker records with missing features

physical aspects and 67% for sleep features. 10% of motion transition features were also missing.

For detection of MCI cases, five distinct machine learning models were used in our experiments for building predictive models, namely Transformer (Vaswani et al. 2017), Long-Short-Term Memory (LSTM) (Hochreiter and Schmidhuber 1997), Support Vector Machine (SVM) (Corrina Cortes 1995), Random Forest (RF) (Ho 1995), and Fusion ART (Tan et al. 2019). These models are chosen as they have been used and showed satisfactory results in previous studies on MCI prediction (Ghoraani et al. 2021; Ahamed, Shahrestani, and Cheung 2020; Bogdanovic, Eftimov, and Simjanoska 2022; Teh, Rawtaer, and Tan 2022). Each of these machine learning models was trained and tested in a supervised learning fashion for predicting MCI cases based on the monthly-averaged biomarker features. For handling missing data, the KNN imputation method (Troyanskaya et al. 2001) was used for all machine learning models, except Fusion ART, which has an in-built mechanism for encoding missing feature values using complement coding (Teh, Rawtaer, and Tan 2022).

We applied stratified 10-fold cross validation, wherein the data set was split into ten folds, each with roughly the same distribution of positive and negative cases. Each machine learning model was then trained on the data set (nine folds) and tested on the remaining one fold. The performance of each learning model was evaluated in the weighted averages of the commonly used precision, recall and F_1 measure for detecting MCI and NC cases.

Table 3 shows the predictive performance of the various machine learning models, including Transformer, LSTM, SVM, RF, and Fusion ART, in identifying MCI cases based on the monthly averaged biomarker features. Fusion ART has produced the highest level of performance across all four measures of precision, recall, specificity, and $F1$ scores. As a distant second, RF performed markedly better than the others. The low performing models may be due to the over-generalization in monthly-averaged data with high missing rates.

	Precision	Recall	Specificity	$F1$ Score
Transformer	0.66±0.07	0.64±0.05	0.60±0.04	0.65±0.04
LSTM	0.71±0.04	0.69±0.04	0.65±0.04	0.70±0.04
SVM	0.63±0.08	0.59±0.05	0.62±0.57	0.60±0.05
RF	0.78±0.08	0.75±0.03	0.78±0.03	0.78±0.03
Fusion ART	**0.83±0.03**	**0.83±0.03**	**0.86± 0.02**	**0.83±0.03**

Table 3: Performance of machine learning models in identifying MCI cases based on monthly averaged biomarker features.

	Precision	Recall	Specificity	F_1 Score
Transformer	0.68±0.07	0.66±0.03	0.60±0.02	0.66±0.02
LSTM	0.74±0.01	0.71±0.01	0.65±0.02	0.71±0.01
SVM	0.64±0.07	0.58±0.02	0.61±0.57	0.60±0.01
RF	0.83±0.01	0.81±0.01	0.74±0.03	0.83±0.01
Fusion ART	**0.85±0.01**	**0.85±0.01**	**0.87±0.01**	**0.85±0.01**

Table 4: Performance of machine learning models in identifying MCI cases based on weekly averaged biomarker features.

Experiments with Weekly Averaged Data

We also aggregated the daily biomarkers data by averaging them into weekly biomarkers records for the experiment. A total of 10,033 weekly biomarker records were obtained with 3487 MCI cases and 5585 NC cases. Besides providing more data records for training and testing the machine learning models, the weekly averaging method should also produce a more balanced class distribution of MCI and NC cases. However, as shown in Table 2, the weekly average data has higher missing rates for every features compared to the monthly averages. It has a missing rate of 83% for physical aspects and 76% for sleep features.

We repeated the experiments following the same methodology as in the monthly averaged data. In Table 4, the predictive performance of all machine learning models improved with the weekly biomarker data set. This may be due to more samples of data with more specific representation of behavior, although more values are missing. In particular, fusion ART achieved the best $F1$ score of 0.85. It is shown also

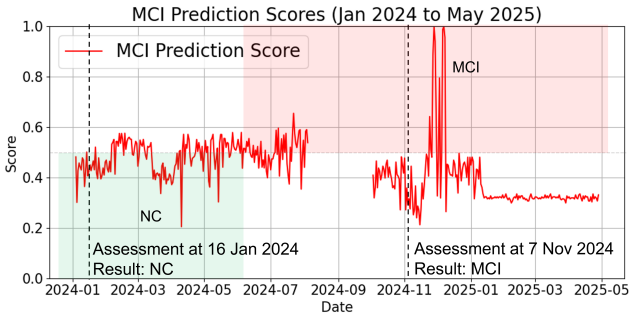

Figure 4: Daily MCI prediction scores of a participant from Jan 2024 to May 2025 across top 5% of activated codes predicting MCI in weekly Fusion ART model, weighted according to their inverse square of ranks of activation values.

that Fusion ART can still perform well with higher missing rates of data. Among the other four models, Transformer, LSTM, SVM produced a similar level of performance while RF becomes the second best model with $F1$ score of 0.83.

Prediction Score and Rules Explaining

Besides the predictive performance, we also looked at the potentials of the prediction by Fusion ART model to handle continuous stream of data over time. We trained the model using the weekly biomarkers data, leaving aside the records from a single participant (namely participant A) with five years of participation and observed transitions between MCI and NC, to reach 0.85 of weighted average of $F1$ score. The learned model was then applied to produce daily MCI prediction scores of participant A from year 2021 to the middle of 2025. The prediction score is based on aggregating top 5% of nodes activated in Fusion ART when predicting MCI from daily biomarkers, weighted to the square of the order of the activation levels. Figure 4 shows the MCI prediction scores of participant A, overlaid with the assigned labels following the date of clinical assessments. For clarity, the range shown is limited to between January 2024 to May 2025. It is shown that the scores are fluctuating but still consistently aligned with the clinical assessments. Interestingly, it can identify sharp peaks in the MCI period closely to the assessment date with MCI, before dropping towards the end of the range in 2025, suggesting a possible behavioral transition towards NC. This suggested the potential of our model not only for MCI detection, but also for capturing behavioral shifts over time.

Besides the scores, the learned nodes and weights can be translated into IF-THEN symbolic classification rules for further analysis. Table 5 shows top eight rules, extracted from Fusion ART based on confidence factors, for both MCI and NC. They indicate that MCI patients tend to be less active in the aspects of physical and activity compared to NC. In addition, the rules indicate that they have a very high tendency to forget their medication.

Rule	Conditions	Interpretation	Class
1 CF:0.5	IF (trans.to liv.room: 0-0.1) & (trans.to kitchen: 0-0.06) & (trans.to bedroom: 0-0.1)	Low trans.to liv.room Very Low trans.to kitchen Low trans.to bedroom	MCI
2 CF:0.36	IF (trans.to liv.room: 0.1-0.2) & (trans.to kitchen: 0) & (trans.to bedroom: 0.1-0.3)	Low trans.to liv.room Very Low trans.to kitchen Low trans.to bedroom	MCI
3 CF:0.15	IF (trans.to liv.room: 0.2-0.3) & (trans.to kitchen: 0.3-0.4) & (trans.to bedroom: 0.2-0.3)	Low trans.to liv.room Low to Mod. trans.to kitchen Low trans.to bedroom	MCI
4 CF:0.12	IF (heart rate: 0-0.1) & (step count: 0-0.2) & (forgetting med.: 1.0)	Low heart rate Low step count Very High forgetting med.	MCI
5 CF:0.76	IF (trans.to liv.room: 0.1-0.2) & (trans.to kitchen: 0) & (trans.to bedroom: 0.1-0.2) & (forgetting med.: 0.5)	Low trans.to liv.room Very Low trans.to kitchen Low trans.to bedroom Mod. forgetting med.	NC
6 CF:0.73	IF (trans.to liv.room: 0-0.05) & (trans.to kitchen: 0.7-0.78) & (trans.to bedroom: 0.7-0.78)	Very Low trans.to liv.room High trans.to kitchen High trans.to bedroom	NC
7 CF:0.59	IF (trans.to liv.room: 0) & (trans.to kitchen: 0.25-0.28) & (trans.to bedroom: 0.2-0.27) & (forgetting med.: 0.5)	Very Low trans.to liv.room Low trans.to kitchen Low trans.to bedroom Mod. forgetting med.	NC
8 CF:0.59	IF (outing: 0.2-0.29) & (outing duration: 0.26-0.3) & (forgetting wallet: 0.16-0.29) & (forgetting med.: 0.5)	Low outing Low outing duration Low forgetting wallet Mod. forgetting med.	NC

Table 5: Top 8 rules for MCI and NC prediction selected based on Confidence Factor (CF).

Discussion

SINEW project has faced various challenges. Some key issues and the way to tackle them are discussed below.

System limitations: Although most sensor devices in SINEW are typically low cost, running them every day in numerous households continuously for long-term while collecting data in real-time requires high-maintenance to minimize lapses in data capture and missing data. A dedicated system to continuously monitor the devices' conditions and to manage both incidental or regular hardware maintenance are required. In addition, the current sensors are designed to capture the data from a single person only and may produce incorrect readings with the presence of another person or pets. In case of the challenge of multiple occupancy, we have configured an additional non-intrusive wearable sensors to be put on by the participant at home, for the next batch of the project, to enable continuous motion and location tracking, minimizing interferences from other occupants.

Human factors: The main challenge in scaling up the system may come from the participants themselves particularly in how they view the system and how they handle the devices installed in their homes. Some participants still concerns about their own privacy even though all kinds of image, video, or audio capturing devices are not used to collect data. Some still feel that certain sensors are intrusive, inconvenient (e.g uncomfortable bed sensors to sleep on), or just forgotten to be properly used (e.g the participant forgets to charge the wearable devices), creating irregularity in the use of the devices to collect the data (Theo 2025). In this case, educating the participants to properly handle the devices and to be reassured no breaching of privacy, is necessary.

Conclusion

A robust in-home cognitive health assessment system to predict MCI cases while providing insightful information for clinical purposes has been presented. As parts of the SINEW project, a large-scale study leveraging ambient intelligence devices and artificial intelligence techniques, the system mitigates the issues and challenges in dealing with noisy and missing values in the data to be learned and processed. Using the Fusion ART neural network architecture, the learned patterns from individuals' daily routines can be explained in terms of rules as the basis for the prediction. Going forward, we can enhance the system to learn a more scalable model from large sensor and biomarker data despite the noisy and highly irregular availability of data. This may include training the model online in day-to-day basis at the gateway devices to predict MCI of different individuals in a more decentralized manner. Integrating the model with the recent generative AI technology, such as large language models (LLMs), can be the next step to improve the model's interpretability, extending the explanation of the prediction model beyond the list of classification rules.

Acknowledgments

This research was supported by the Sengkang General Hospital, Singhealth Foundation and National Medical Research Council Transition Award (TA22jul-0012), the SMU-A*STAR Joint Lab in Social and Human-Centered Computing (Grant No. SAJL-2022-HAS001), and the Lee Kong Chian Professorship awarded to Ah-Hwee Tan by Singapore Management University.

Ethical Statement

This study adhered to ethical guidelines and received approval from SingHealth Centralised Institutional Review Board with ethics approval number CIRB 2019-2026. Informed consent was obtained from all participants, who were assured of anonymity, confidentiality, and the right to withdraw at any time. Data were securely stored and used solely for research purposes.

References

Ahamed, F.; Shahrestani, S.; and Cheung, H. 2020. Internet of things and machine learning for healthy ageing: identifying the early signs of dementia. *Sensors*, 20(21): 6031.

Bogdanovic, B.; Eftimov, T.; and Simjanoska, M. 2022. In-depth Insights into Alzheimer's Disease by using Explainable Machine Learning Approach-Existing Hypothesis Confronted. *Scientific Reports*, 12(6508).

Buysse, D. J.; Reynolds, C. F.; Monk, T. H.; Berman, S. R.; and Kupfer, D. J. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research*, 28: 193–213.

Chen, B.; Tan, H.; Rawtaer, I.; and Tan, H. 2019. Objective Sleep Quality as a Predictor of Mild Cognitive Impairment in Seniors Living Alone. In *2019 IEEE International Conference on Big Data (Big Data)*, 1619–1624. Los Alamitos, CA, USA: IEEE Computer Society.

Corrina Cortes, V. V. 1995. Support-vector networks. *Machine Learning*, 20: 273–297.

Dawadi, P. N.; Cook, D. J.; Schmitter-Edgecombe, M.; and Parsey, C. 2013. Automated assessment of cognitive health using smart home technologies. *Technology and Health Care*, 21: 323–343.

Ghoraani, B.; Boettcher, L. N.; Hssayeni, M. D.; Rosenfeld, A.; Tolea, M. I.; and Galvin, J. E. 2021. Detection of mild cognitive impairment and Alzheimer's disease using dual-task gait assessments and machine learning. *Biomedical signal processing and control*, 64: 102249.

Gielis, K.; Vanden Abeele, M. E.; Verbert, K.; Tournoy, J.; De Vos, M.; and Vanden Abeele, V. 2021. Detecting Mild Cognitive Impairment via Digital Biomarkers of Cognitive Performance Found in Klondike Solitaire: A Machine-Learning Study. *Digital Biomarkers*, 5(1): 44–52.

Hawthorne, G. 2006. Measuring Social Isolation in Older Adults: Development and Initial Validation of the Friendship Scale. *Social Indicators Research*, 77: 521–548.

Ho, T. K. 1995. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, 278–282 vol.1.

Hochreiter, S.; and Schmidhuber, J. 1997. Long Short-term Memory. *Neural computation*, 9: 1735–80.

Kwon, L.-N.; Yang, D.-H.; Hwang, M.-G.; Lim, S.-J.; Kim, Y.-K.; Kim, J.-G.; Cho, K.-H.; Chun, H.-W.; and Park, K.-W. 2021. Automated Classification of Normal Control and Early-Stage Dementia Based on Activities of Daily Living (ADL) Data Acquired from Smart Home Environment. *International journal of environmental research and public health*, 18(24): 13235.

Lawton, M. P.; and Brody, E. M. 1969. Instrumental activities of daily living (IADL). *The Gerontologist*, 9: 179–186.

Lee, B.; Ahn, C. R.; Mohan, P.; Chaspari, T.; and Lee, H.-S. 2020. Evaluating routine variability of daily activities in smart homes with image complexity measures. *Journal of Computing in Civil Engineering*, 34(6): 04020042.

Lee, B.; Mohan, P.; Chaspari, T.; and Ryan Ahn, C. 2023. Assessing Daily Activity Routines Using an Unsupervised Approach in a Smart Home Environment. *Journal of Computing in Civil Engineering*, 37(1): 04022050.

Lubben, J. E. 1988. Assessing social networks among elderly populations. *Family & Community Health*, 11: 45–52.

Narasimhan, R.; Muthukumaran, G.; Dey, K.; and Ramakrishnan, A. 2022. Recurrent Neural Network based Prediction of Transition to Mild Cognitive Impairment Using Unobtrusive Sensor Data. In *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, 1–6.

Quek, L. J. V.; Heikkonen, M. R.; and Lau, Y. 2023. Use of artificial intelligence techniques for detection of mild cognitive impairment: A systematic scoping review. *Journal of Clinical Nursing*, (September 2022): 1–11.

Rawtaer, I.; Mahendran, R.; Kua, E. H.; Tan, H. P.; Tan, H. X.; Lee, T.-S.; and Ng, T. P. 2020. Early Detection of Mild Cognitive Impairment With In-Home Sensors to Monitor Behavior Patterns in Community-Dwelling Senior Citizens in Singapore: Cross-Sectional Feasibility Study. *J Med Internet Res*, 22(5): e16854.

Robert, P. H.; Clairet, S.; Benoit, M.; Koutaich, J.; Bertogliati, C.; Tible, O.; Caci, H.; Borg, M.; and Bedoucha, P. B. P. 2002. The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. *International Journal of Geriatric Psychiatry*, 17: 1099–1105.

Tan, A.-H.; Subagdja, B.; Wang, D.; and Meng, L. 2019. Self-organizing neural networks for universal learning and multimodal memory encoding. *Neural Networks*, 120: 58–73.

Tan, A.-H.; Ying, W.-Y.; Subagdja, B.; Huang, A.; D, S.; Tay, T. C.-I.; and Rawtaer, I. 2024. Predicting Mild Cognitive Impairment through Ambient Sensing and Artificial Intelligence. In *2024 IEEE Conference on Artificial Intelligence (CAI)*, 1098–1104.

Taraldsen, K.; Chastin, S. F.; Riphagen, I. I.; Vereijken, B.; and Helbostad, J. L. 2012. Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: A systematic literature review of current knowledge and applications. *Maturitas*, 71(1): 13–19.

Teh, S.-K.; Rawtaer, I.; and Tan, A.-H. 2022. Predictive self-organizing neural networks for in-home detection of Mild Cognitive Impairment. *Expert Systems with Applications*, 205: 117538.

Theo, J. 2025. Smart sensors used in study to detect cognitive decline in seniors who live alone. *The Strait Times*, Monday 7 July 2025: 15.

Troyanskaya, O.; Cantor, M.; Sherlock, G.; Brown, P.; Hastie, T.; Tibshirani, R.; Botstein, D.; and Altman, R. B. 2001. Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17(6): 520–525.

Vaswani, A.; Shazeer, N. M.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is All you Need. In *Neural Information Processing Systems*.

Yesavage, J. A.; Brink, T.; Rose, T. L.; Lum, O.; Huang, V.; Adey, M.; and Leirer, V. O. 1982. Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17: 37–49.

Zung, W. W. K. 1965. A self-rating depression scale. *Archives of General Psychiatry*, 12: 63–70.