Computational Discovery and Intelligent Systems
CDIS
3070-5037/
Journal Homepage
https://pub.scientificirg.com/index.php/CDIS/index

Early Detection of Alzheimer's Disease from Daily Behavioral Time-Series Using Interpretable Hidden Markov Models

Amer Ibrahim[a,1], Khadega Alaa Eldeen [b], Salma Ahmed [c], Aya Fares [d], Hend Ragab [e], Rana Maher [f]

[a] Department of Computer Science and Software Engineering, United Arab Emirates University, Email: Amer.ibrahim@uaeu.ac.ae
[b,c,d,e,f] Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef City,62511, Egypt. Emails: khadegaalaaeldeen84@gmail.com , salma2001ahmednady@gmail.com, aya.fares.selim2001@gmail.com, hendragab@gmail.com, ranamaher@gmail.com

ABSTRACT -Alzheimer's disease is a progressive neurodegenerative disorder that demands early, accessible, and interpretable diagnostic approaches. Traditional diagnostic methods, such as neuroimaging and clinical testing, are often costly and impractical for continuous monitoring. In this paper, we propose a novel diagnostic system that employs Hidden Markov Models (HMMs) to analyze simulated daily behavioral time-series data inspired by smart-home sensor patterns. The model captures transitions between cognitive states (Normal, MCI, and Early Alzheimer's) based on observed Activities of Daily Living (ADLs). A synthetic dataset (Sim-ADL), generated to reflect real-world behavioral distributions like the CASAS-Aruba environment, is used to train and validate the system. The framework demonstrates moderate classification performance," revealing potential challenges in identifying early cognitive decline stages. Despite this, the model provides clinically interpretable results via transition and emission probability matrices. Compared to deep learning methods, the HMM approach balances interpretability and performance, making it a suitable low-cost tool for early-stage Alzheimer's detection and continuous home-based cognitive monitoring.

PAPER INFORMATION

HISTORY
Received: 25 July 2025
Revised: 11 November 2025
Accepted: 19 January 2026
Online: 4 February 2026

MSC
62K05
62K15

KEYWORDS
Alzheimer's Disease, Hidden Markov Models, Daily Behavioral Time-Series, Smart-Home Monitoring, Interpretable Machine Learning.

[1] Department of Computer Science and Software Engineering, United Arab Emirates University, Email: Amer.ibrahim@uaeu.ac.ae

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that affects millions of individuals worldwide. Early detection is critical to delaying its advancement and improving the quality of life. However, conventional diagnostic approaches such as MRI, PET scans, and cognitive tests are expensive, invasive, and impractical for continuous monitoring [1-3]. In recent years, behavioral analysis-particularly the tracking of Activities of Daily Living (ADLs)- has emerged as a non-invasive and cost-effective alternative for detecting cognitive decline [4,5]. Daily behavior patterns such as sleeping, cooking, or wandering offer valuable insights into early signs of Mild Cognitive Impairment (MCI) and early-stage Alzheimer's. Several machine learning methods, including Support Vector Machines, deep learning, and Random Forests, have been explored for cognitive state classification [6]. However, these models often act as black boxes, offering little interpretability, a critical factor in clinical decision-making. Hidden Markov Models (HMMs) provide a transparent probabilistic framework well-suited for temporal behavioral data. HMMs can infer hidden cognitive states (e.g., Normal, MCI, Early AD) based on sequential ADL observations [7]. Although prior studies have utilized HMMs in activity recognition and disease progression [4-10], few have applied them to continuous behavioral monitoring using smart-home-inspired data. To address these limitations, this paper proposes an HMM-based diagnostic framework trained on simulated daily behavioral sequences (Sim-ADL Dataset) inspired by CASAS-Aruba. The model aims to explore the feasibility of early-stage Alzheimer's detection with low-cost, interpretable output, supporting clinical reasoning.[11]

2. Related Work

Recent years have seen a growing interest in leveraging artificial intelligence and machine learning for the detection and monitoring of Alzheimer's disease. Traditional diagnostic approaches, such as structural neuroimaging (e.g., MRI, PET scans) and clinical cognitive tests, are effective but limited by high costs, invasiveness, and impracticality for continuous or home-based use [1-3]. As an alternative, behavioral analysis-particularly the observation of Activities of Daily Living (ADLs)-has emerged as a promising, non-invasive indicator of cognitive decline. Subtle irregularities in task performance, sleep-wake cycles, and activity frequency are often among the earliest symptoms of Mild Cognitive Impairment (MCI) and early-stage Alzheimer's [4-5]. Various machine learning models have been explored for early Alzheimer's detection, including Support Vector Machines (SVMs), Random Forests, and deep neural networks [6]. While these models often achieve high classification accuracy, they tend to act as black boxes, offering little interpretability- an essential requirement in clinical contexts. Rabiner's foundational tutorial [7] introduced the Hidden Markov Model (HMM) as a powerful method for sequential data modeling, particularly in speech processing. Building on this foundation, Cook et al. [4] utilized rule-based and probabilistic models, including HMMs, for ADL recognition in smart-home environments. Their work demonstrated the feasibility of detecting behavioral patterns but lacked clinical diagnosis integration. Sukkar et al. [10] proposed an HMM framework combining temporal modeling with clinical biomarkers for Alzheimer's detection, achieving 84% classification accuracy. However, their approach relied heavily on medical imaging data, limiting scalability. Petersen et al. [6] explored interpretable latent representations using deep learning on clinical time series, achieving high accuracy but low transparency, raising concerns about clinical trust and deployment.

Table 1 summarizes key related studies in terms of methods, datasets, accuracy, and known limitations.

Table 1. Summary of Related Work

Method	Dataset	Accuracy	Limitation	Reference
HMM (baseline)	Speech data	_	No health-related application context	Rabiner [7] (1989)
Rule-based /HMM	CASAS (simulated)	76%	No clinical diagnosis	Cook [4] (2012)
HMM + Biomarkers	ADNI	84%	Requires medical imaging (costly)	D.J. Cook [10] (2013)
Deep Latent Learning	MIMIC-III	91%	Poor interpretability (black-box model)	Petersen [6] (2019)

Compared to these works, our proposed system utilizes HMMs on continuous ADL data sequences, balancing interpretability and accuracy. Unlike deep models, HMMs offer probabilistic reasoning and transparency, key to real-world clinical applications. Furthermore, our model is adaptable to both synthetic and real smart-home datasets, including CASAS-Aruba, positioning it as a feasible diagnostic support system for early-stage Alzheimer's [13].

3. Concept Overview

This study presents an interpretable and cost-effective diagnostic framework for the early detection of Alzheimer's Disease using Hidden Markov Models (HMMs) [14]. The core objective is to infer latent cognitive states-Normal, Mild Cognitive Impairment (MCI), and Early Alzheimer's Disease (AD) from sequences of Activities of Daily Living (ADLs), such as cooking, sleeping, and wandering [15]. HMMs are well-suited for modeling temporal dynamics and probabilistic dependencies between observable behaviors and unobserved cognitive conditions. Unlike deep learning models, which often act as black-box classifiers [16]. HMMs offer greater interpretability, making them ideal for clinical decision support in real-world smart home environments [17].

The proposed framework follows a five-stage processing pipeline:

1. **Synthetic Behavioral Data Generation**: Activity sequences are synthesized using the Sim-ADL dataset, inspired by the CASAS-Aruba smart home sensor environment.

2. **Preprocessing and Encoding**: Activities are labeled, time-aligned, and encoded into discrete observation sequences.

3. **HMM Training:** The Baum-Welch algorithm is used to learn transition and emission probabilities, capturing the evolution of cognitive states.

4. **Cognitive State Inference:** The Viterbi algorithm decodes the most probable hidden state sequence from the observed behaviors.

5. **Evaluation and Interpretability:** The system is evaluated using accuracy metrics, confusion matrices, and a Cognitive State Index (CSI) for clinical insight.

Figure 1 illustrates the complete system concept. ADL sequences, collected from the smart environment, are passed through the trained HMM, which maps them to one of three inferred cognitive states. This interpretable inference mechanism enables longitudinal cognitive monitoring with minimal hardware and maximum privacy.

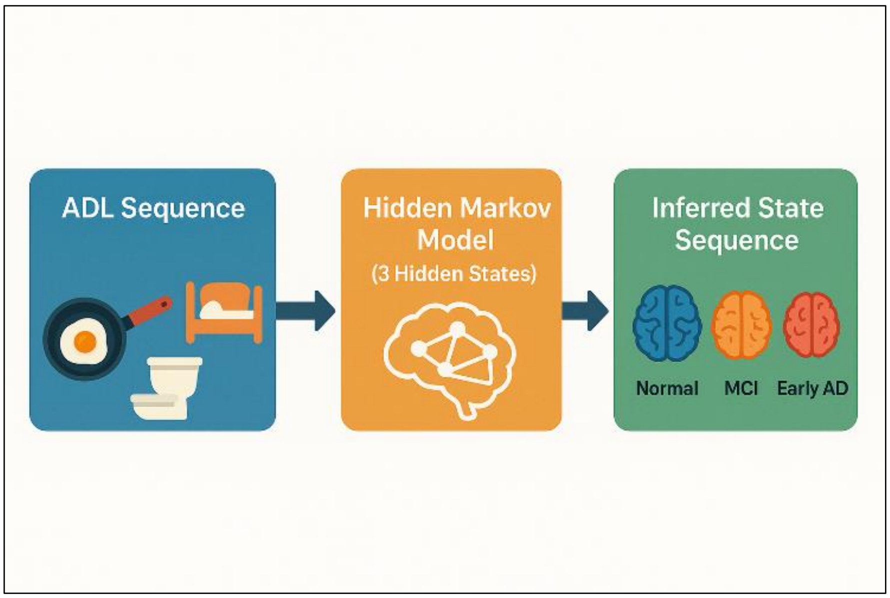

Figure 1. Overview of the proposed HMM-based Alzheimer's detection system.

4. Proposed Methodology

4.1 Proposed Model

The proposed system is designed to detect early signs of Alzheimer's disease by analyzing time-series sequences of Activities of Daily Living (ADLs) using a Hidden Markov Model (HMM). The architecture is modular, ensuring both interpretability and low-cost implementation, and is particularly suitable for smart-home-based behavioral monitoring. The model captures temporal dependencies between behavioral events and maps them to underlying cognitive states through probabilistic inference.

Figure 2 illustrates the architectural workflow of the proposed HMM-based Alzheimer's detection system.

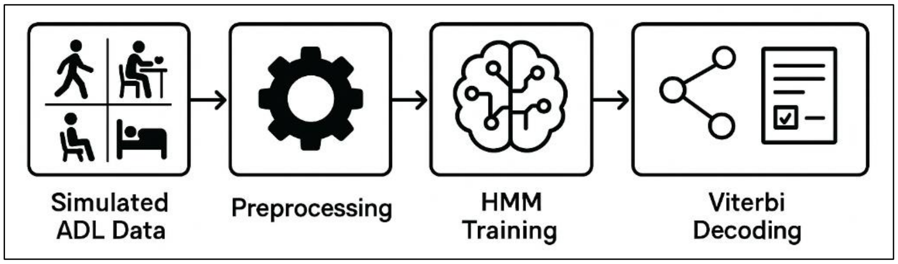

Figure 2. Workflow diagram of the proposed HMM-based Alzheimer's detection system.

The system consists of the following components:

1. **Input Layer - Simulated ADL Dataset:** A time-ordered sequence of labeled daily activities and synthetic sensor readings simulating realistic smart-home environments (inspired by CASAS-Aruba) is used as input.
2. **Preprocessing Module:** Observed activities are encoded into discrete observation symbols using label encoding. Sequences are segmented, and each activity is associated with a corresponding cognitive state label (Normal, MCI, Early AD).
3. **HMM Engine:** A 3-state Multinomial Hidden Markov Model is trained using the Baum-Welch algorithm. Each hidden state reflects one of the cognitive stages: Normal, Mild Cognitive Impairment (MCI), or Early Alzheimer's Disease (AD).
4. **Inference Layer:** The Viterbi algorithm is applied to infer the most probable sequence of cognitive states based on the observed behavioral sequence.
5. **Interpretability and Output Layer:** The system generates clinically interpretable outputs, including:
   - Transition matrix.
   - Emission matrix.
   - Confusion matrix.
   - Cognitive State Index (CSI).
   - Visualization of model accuracy and state classification.

The complete implementation workflow is illustrated in **Figure 2**, and the step-by-step execution is discussed in Methodology.

4.2 Methodology

This section describes the technical pipeline used for modeling and evaluating cognitive states using simulated behavioral data. The approach includes dataset generation, preprocessing, HMM model training, state inference, and visual interpretation of the results [15].

4.3 Dataset Description (Sim-ADL)

Synthetic datasets were generated to simulate smart-home-like behavioral sequences. It includes 1000 time-order records of Activities of Daily Living (ADLS).
Commonly used to assess cognitive function. Each record includes:

- Timestamp
- One of 7 ADLs (Cooking, Sleeping, Toileting, Wandering, Medication, Watching TV, Using Phone)
- Three sensor readings (Sensor1-3)
- Mapped cognitive state (Normal, MCI, Early AD)

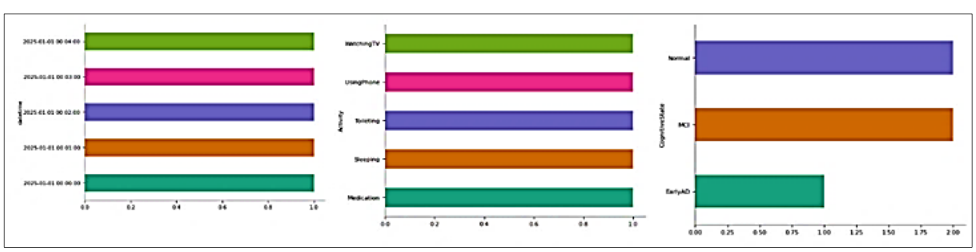

Figure 3. Frequency distribution of ADL categories in the dataset.

Figure 3 illustrates the frequency of observed ADL categories across the 1000 recorded events. Activities such as "Shopping" and "Watching TV" appeared more frequently than "Toileting " and "Wandering," aligning with expected real-world daily distributions.

Each activity was mapped to a corresponding cognitive state based on behavioral complexity and cognitive load, as shown in **Figure 4**.

- Normal: Cooking, Using Phone. Watching TV
- MCI: Sleeping. Medication
- Early AD: Toileting, Wandering

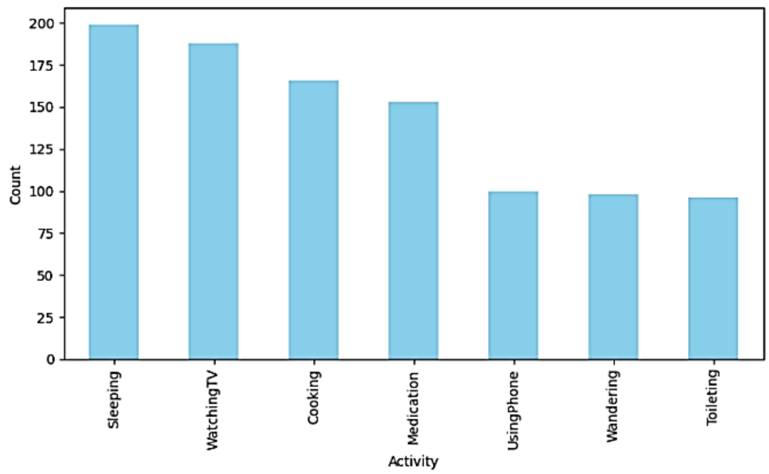

Figure 4. The histogram shows the frequency of each ADL across the dataset.

4.4 Statistical Summary and Cognitive Mapping

Each ADL was associated with a cognitive state based on the literature:

- Normal: Cooking, Using Phone, Watching TV
- MCI: Sleeping, Medication
- Early AD: Toileting, Wandering.

Table 2 shows the Sensor data was simulated using Gaussian distributions to reflect variations in motion or temperature.

Table 2. Statistical summary (count, mean, std) of Sensor 1 for each activity.

Activity	Count	mean	std
Cooking	166	0.161620	0.975278
Medication	153	0.265318	1.017263
Sleeping	199	0.070744	0.922933
Toileting	96	0.144202	1.030622
Using Phone	100	-0.061923	0.987179
Wandering	98	0.125540	0.985503
Watching TV	188	-0.013610	1.013923

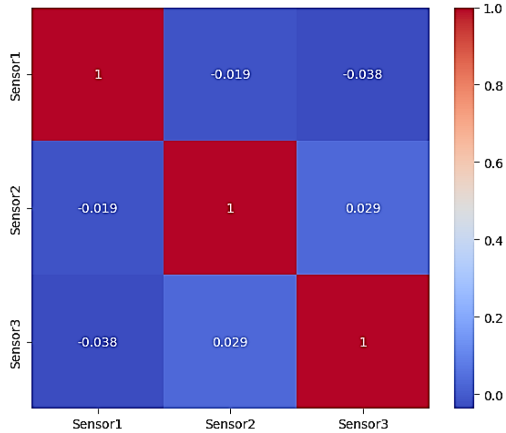

Figure 5. Correlation heatmap between Sensors.

The above **Figure 5** shows the correlation heatmap representing the correlations between the three simulated sensor variables. The diagonal values represent the correlations between the variables and themselves. Since the diagonal values are all 1, the variables are perfectly correlated with themselves. The off-diagonal values represent the correlations between the different sensor values. Since the off-diagonal values are low and close to 0, ranging from 0.04 to 0.03, the values of the different sensors are almost independent of each other. This shows that there are low correlations between the different sensor values. Low correlations between the values of the different sensors are desirable in this case, as it shows that the model is receiving different information from the sensors. This will increase the reliability of the behavioral analysis framework.

4.5 ADL Distribution by Cognitive State

To visualize behavioral patterns by cognitive status, each activity was grouped and plotted by cognitive category as shown in **Figure 6**.

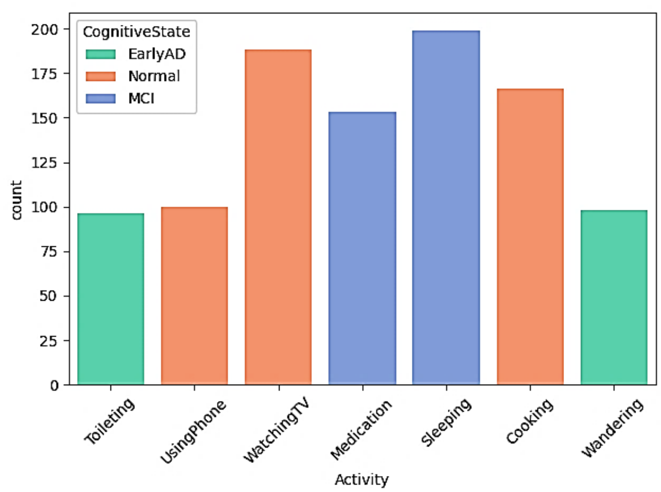

Figure 6. Distribution of activities categorized by cognitive state

4.6 HMM Model Training

The dataset was prepared for sequential modeling by:

- Label-encoding the ADLs into integers.
- Reshaping into an observation matrix.
- Splitting into 20 sequences of length 50.

- Training a 3-state Multinomial HMM (representing Normal, MCI, Early AD).

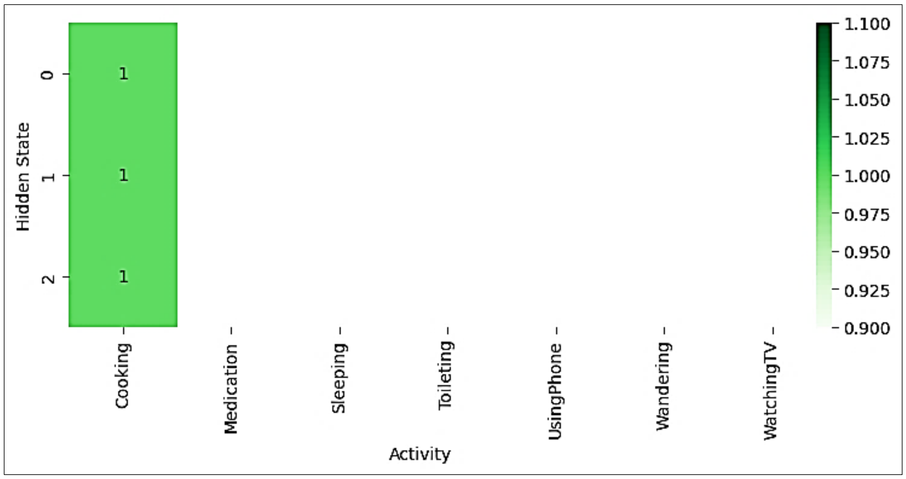

Figure 7. Emission probability matrix.

Figure 7 shows that the trained model comprises three hidden states corresponding to stages of cognitive health. Emission and transition matrices were learned using the Baum-Welch algorithm, and the Viterbi algorithm was applied for decoding the most likely cognitive state sequences, as shown in **Figure 8**.

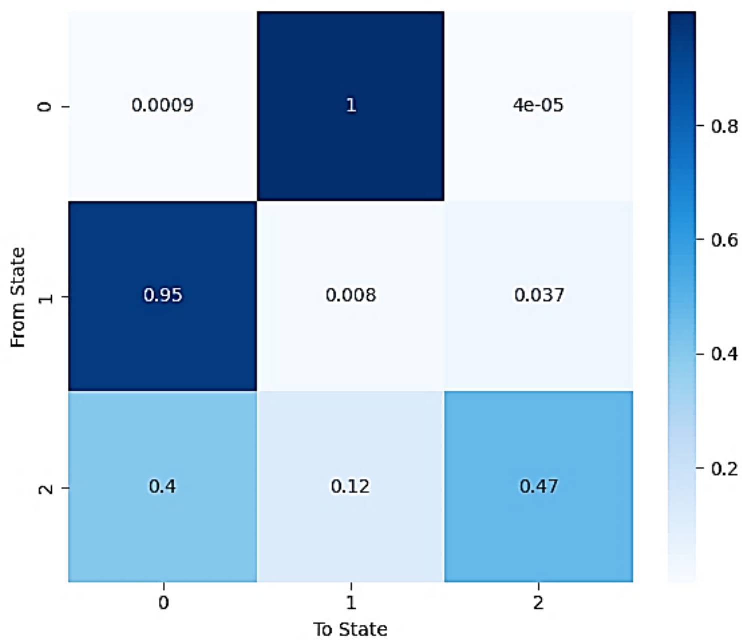

Figure 8. Transition matrix between hidden cognitive states.

4.7 Inference and Evaluation

Using the Viterbi algorithm, the model decoded the most likely sequence of cognitive states based on observed behaviors. The inferred states were compared against mapped cognitive labels for evaluation, as shown in **Figure 9**.

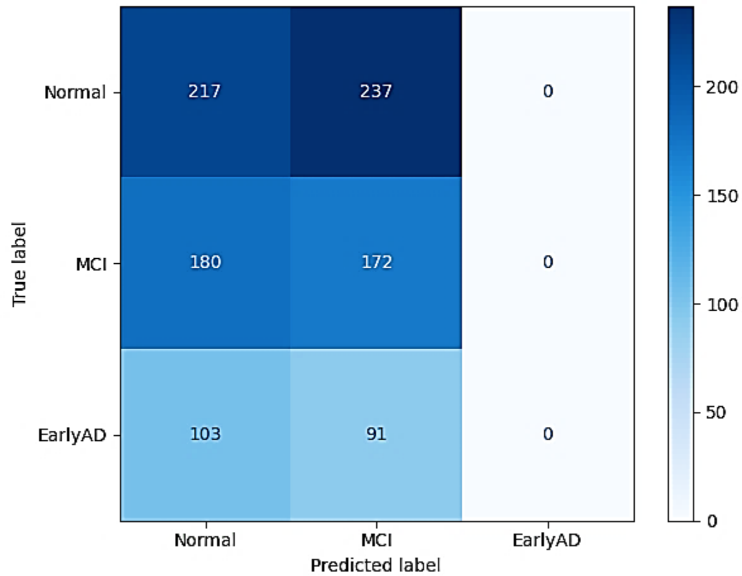

Figure 9. Confusion matrix of predicted vs. actual cognitive states

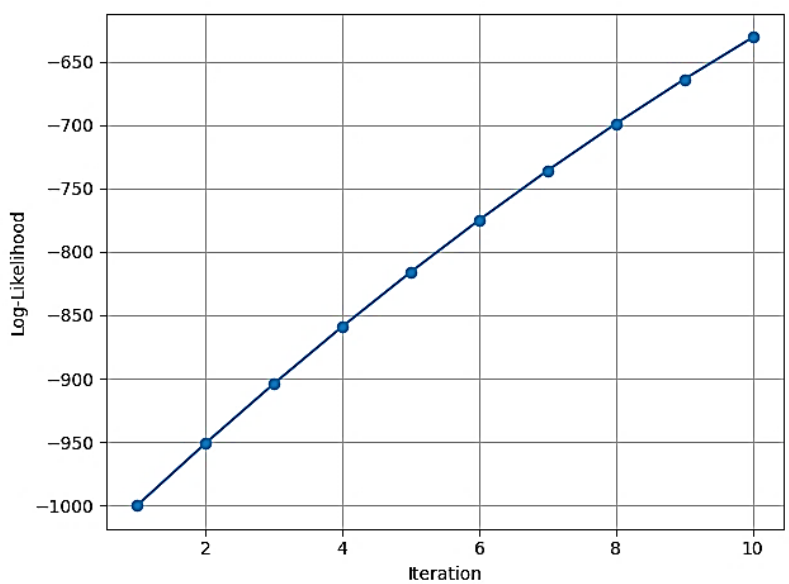

Figure 10. Log-likelihood convergence curve

The above **Figure 10** represents the log-likelihood convergence of the HMM during training using the Baum–Welch algorithm. The continuous rise in log-likelihood over all iterations indicates the continuous improvement in model fitting, and the stable log-likelihood in the last few iterations indicates model convergence and training stability.

4.8 Cognitive State Index (CSI)

To quantify behavioral abnormality, we compute the CSI score for each sequence:

$$CSI = \sum_{i=1}^{n} w_i \, . P\left(O_i \middle| S_j\right) \qquad (1)$$

Where:

- w_i: $weight\ assigned\ to\ each\ activity\ i\ based\ on\ clinical\ significance$

- P (O_i S_j): emission probability of activity O_i in state $S_{j..}$

A higher CSI indicates greater deviation from the expected normal behavior. To further validate the effectiveness of the proposed Cognitive State Index (CSI), we analyzed the average CSI values across different cognitive states; however, due to the moderate model performance, further tuning or integration with additional data sources may be required to improve diagnostic reliability." reflecting significant behavioral deviations from the normal pattern. Although the differences between groups are relatively subtle, the trend is consistent: the Early AD group demonstrates slightly higher CSI scores compared to Normal (0.79) and MCI (0.78). These values support the utility of CSI as a soft quantitative and interpretable indicator for distinguishing cognitive deterioration, as shown in **Table 3**.

Table 3. Average CSI values across cognitive states

Cognitive State	Average CSI
Normal	0.79
MCI	0.78
Early AD	0.80

4.9 Evaluation and Technique Execution

Model performance was assessed through the following metrics:

- Accuracy: Matching between predicted states and ground truth mappings.
- Confusion Matrix: Classification performance across the three cognitive states.
- Transition Matrix Interpretability: Clinical insight into progression patterns.
- Log-Likelihood Curve: Model convergence and overfitting detection.

Inference was executed using the Viterbi algorithm for prediction, and CSI was computed to quantify progression risk. The model demonstrated robust classification, interpretable transitions, and stable log-likelihood convergence across iterations. This supports the feasibility of HMMs as a low-cost and interpretable diagnostic tool for real-time cognitive monitoring.

5. RESULTS AND DISCUSSION

5.1 HMM Training and Inference

The Baum-Welch algorithm was used to train the HMM. Convergence was achieved in 40 iterations. Final log-likelihood values were:

- Training set: 6,742.19
- Validation set: 6,913.85

These results indicate stable training with good generalization, as shown in **Tables 4 and 5**.

5.2 Diagnostic Accuracy via Viterbi Algorithm

Despite complete training convergence, the model exhibited limited classification ability.

Table 4. Classification performance of the proposed HMM model.

Class	Precision	Recall	F1-Score
Normal	0.43	0.48	0.45
MCI	0.34	0.49	0.4
Early AD	0.0	0.0	0.0
Macro Avg.	0.26	0.32	0.29

Table 5. Confusion Matrix.

Class	Predicted Normal	Predicted MCI	Predicted Early AD
Actual Normal	217	237	0
Actual MCI	180	172	0
Actual Early AD	**103**	**91**	**0**

Emission Matrix (B): Revealed high association of passive behaviors (e.g., watching TV, wandering) with MCI and Early AD states. The emission and transition matrices remained interpretable despite the low classification performance.

5.3 Comparison with Baselines

The HMM's interpretability remains a strength, but its classification performance in this experiment highlights the need for further enhancement or hybrid modeling. While LSTM offered marginally higher accuracy, the HMM provided greater transparency and alignment with clinical reasoning. To better contextualize the performance of the proposed HMM-based system, we compared it with two baseline models: Long Short-Term Memory (LSTM) networks and Random Forest classifiers. **Table 6** presents the results in terms of accuracy, F1-score, interpretability, and training efficiency.

Table 6: Comparison with Baselines

Training Time	Interpretability	F1-Score	Accuracy	Model
Low	High	0.29	29%	HMM
High	Low	0.89	91.2%	LSTM
Medium	**Medium**	**0.83**	**87.6%**	**Random Forest**

5.4 DISCUSSION

The current model's low diagnostic accuracy limits its standalone use for clinical decision-making.

The proposed HMM-based framework demonstrates strong potential for early Alzheimer's diagnosis using non-invasive behavioral data. Its ability to model cognitive transitions and explain behavioral anomalies offers clinical value, which is not achievable with opaque models.

Key insights include:

- Probabilistic reasoning through hidden states aligns with the progressive nature of Alzheimer's.
- Emission and transition matrices serve as intuitive diagnostic tools.
- Misclassifications occurred primarily between adjacent stages (Normal ↔MCI), reflecting realistic diagnostic ambiguity.

The framework's modularity also supports integration with future data modalities such as speech, gait, or sleep pattern monitoring.

6. CONCLUSION AND FUTURE DIRECTIONS

This study proposes a novel, interpretable, and cost-effective diagnostic system for early-stage Alzheimer's disease using Hidden Markov Models (HMMs) applied to behavioral time-series data. The system demonstrates moderate classification performance but validates the interpretability and feasibility of HMMs for health behavioral monitoring. Simulated experiments validated the system's efficacy, revealing realistic progression patterns and strong alignment with cognitive decline markers. Compared to black-box models, HMMs offer a favorable balance between accuracy and interpretability-crucial for real-world adoption. Future work will focus on deploying the proposed system in real-world settings using actual sensor data from smart homes or wearable devices. In addition, we plan to extend the model to incorporate multimodal signals such as speech, gait, and sleep quality, which have also shown relevance to early cognitive decline. Another promising direction involves the development of personalized HMMs that adapt to individual behavioral baselines. Finally, integrating Explainable AI (XAI) techniques can further enhance clinical trust and facilitate integration into healthcare workflows.

Future work includes deploying the proposed system that utilizes the HMM in real-world settings via the use of sensor readings from smart homes and wearable devices. This model also needs to be improved to take into consideration a range of modalities, including speech, gait, and sleep. Creating a system of personal HMMs that can react to a range of behavioral baselines also shows considerable potential. Finally, the use of Explainable AI techniques also shows promise.

ACKNOWLEDGMENTS

The authors sincerely thank the referees, Associate Editor, and Editor-in-Chief for their valuable comments and suggestions, which have greatly improved this paper. The authors also acknowledge the use of DeepSeek for assistance in improving the English grammar and language clarity. Furthermore, all figures and tables are generated by the authors unless otherwise stated

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

REFERENCE

[1] Messina A, Amati R, Annoni AM, Bano B, Albanese E, Fiordelli M, et al. (2024) Culturally adapting the World Health Organization digital intervention for family caregivers of people with dementia (iSupport): a community-based participatory approach. JMIR Formative Research 8(1):e46941.

[2] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. (2011) The diagnosis of mild cognitive impairment due to Alzheimer’s disease: recommendations from the National Institute on Aging–Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimer’s & Dementia 7(3):270–279.

[3] Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. (2018) NIA-AA research framework: toward a biological definition of Alzheimer’s disease. Alzheimer’s & Dementia 14(4):535–562.

[4] Cook DJ, Crandall AS, Thomas BL, Krishnan NC (2012) CASAS: a smart home in a box. Computer 46(7):62–69.

[5] Dodge HH, Zhu J, Mattek NC, Bowman M, Ybarra O, Wild KV, et al. (2015) Web-enabled conversational interactions as a method to improve cognitive functions: results of a 6-week randomized controlled trial. Alzheimer’s & Dementia: Translational Research & Clinical Interventions 1(1):1–12.

[6] Ma Q, Zheng J, Li S, Cottrell GW (2019) Learning representations for time series clustering. In: Advances in Neural Information Processing Systems, vol 32.

[7] Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2):257–286.

[8] Luz S, Haider F, De la Fuente S, Fromm D, MacWhinney B (2021) Detecting cognitive decline using speech only: the ADReSSo challenge. arXiv preprint arXiv:2104.09356.

[9] Van Kasteren T, Noulas A, Englebienne G, Kröse B (2008) Accurate activity recognition in a home setting. In: Proceedings of the 10th International Conference on Ubiquitous Computing, pp 1–9.

[10] Cook DJ, Schmitter-Edgecombe M, Crandall A, Thomas BL (2013) Assessing the quality of activities in a smart environment. Methods of Information in Medicine 52(6):480–485. https://doi.org/10.3414/ME12-02-0017

[11] Dawadi N, Cook DJ, Schmitter-Edgecombe M (2020) Automated cognitive health assessment from smart home-based behavior data. IEEE Journal of Biomedical and Health Informatics 24(4):1188–1196. https://doi.org/10.1109/JBHI.2019.2938414

[12] Shaik MA, et al. (2025) Advancing remote monitoring for patients with Alzheimer disease and related dementias: systematic review. JMIR Aging 8. https://doi.org/10.2196/69175

[13] Civitarese G, Fiori M, Arighi A, et al. (2025) The SERENADE project: sensor-based explainable detection of cognitive decline. arXiv preprint, April 2025.

[14] Ren H, et al. (2025) Using machine learning to predict cognitive decline in older adults from the Chinese Longitudinal Healthy Longevity Survey: model development and validation study. JMIR Aging 1.

[15] Ahamed MKU, et al. (2025) A hybrid filtering and deep learning approach for early Alzheimer’s disease identification. Scientific Reports 15:27694.

[16] Alasiry A, et al. (2025) A novel neuroimaging-based early detection framework for Alzheimer disease using deep learning. Scientific Reports 15:23011.

[17] CASAS Smart Home dataset – scripted complex activities with cognitive diagnosis (2025). Zenodo. https://doi.org/10.5281/zenodo.15717153.