Modeling Human Activity States Using Hidden Markov Models

Group 2
Nicolas Muhigi
Irakoze Amandine

October 31, 2025

1 Background and Motivation

Our group addresses personalized health monitoring in elderly care facilities through human activity recognition. Traditional monitoring relies on intrusive cameras or manual observation, compromising privacy and requiring significant resources. By leveraging smartphone accelerometer and gyroscope sensors with Hidden Markov Models, we developed a non-intrusive system detecting daily activities (standing, walking, jumping, still) to help caregivers identify fall risks, monitor mobility patterns, and detect behavioral anomalies indicating health deterioration, enabling early intervention while respecting patient dignity.

2 Data Collection and Preprocessing Steps

2.1 Data Collection Protocol

We collected motion sensor data using Sensor Logger (v1.47.1) on iPhone X devices. The dataset comprises 50 samples across four activities: standing, walking, jumping, and still. Table 1 summarizes our data collection parameters.

Table 1: Data Collection Summary

Parameter	**Value**
Devices	1 × iPhone X (shared between members)
Sensors	Accelerometer (x,y,z), Gyroscope (x,y,z)
Sampling Rate	99.6 Hz (harmonized median)
Duration per Sample	5-10 seconds
Total Samples	50 (12-14 per activity)
Phone Placement	Waist-level (standing/walking/still); hand-held (jumping)

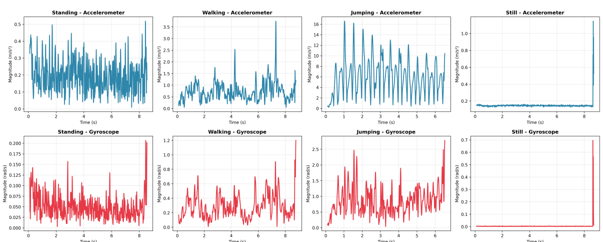

Figure 1: Data Visualization

We used the same iPhone X devices with minimal sampling rate drift (99.8-100.2 Hz). We computed the median frequency (99.6 Hz) for consistency, requiring no resampling due to negligible (¡0.5%) difference.

2.2 Preprocessing and Feature Extraction

Windowing: We applied 3-second sliding windows with 50% overlap (1.5s step), generating 188 feature windows total. This configuration captures 2-4 jump cycles and provides 300 samples per window for stable features.

Features Extracted: We computed 12 features per window:

- **Time-domain (7):** Mean acceleration (x,y,z), standard deviation (x,y,z), signal magnitude area (SMA)
- **Frequency-domain (2):** Dominant frequency via FFT, spectral energy
- **Correlation (3):** Pairwise axis correlations (xy, xz, yz)

Normalization: Features were Z-score normalized ($\mu = 0$, $\sigma = 1$) separately for train/test sets to satisfy Gaussian emission requirements and prevent data leakage.

Train-Test Split: 80/20 split by sample ID yielded 38 training samples (143 windows) and 12 test samples (45 windows).

3 HMM Setup and Implementation Details

3.1 Model Components

Our HMM consists of:

- **Hidden States (Z):** 4 states = {standing, walking, jumping, still}
- **Observations (X):** 12-dimensional feature vectors
- **Transition Matrix (A):** 4×4 where $A_{ij} = P(z_t = j \mid z_{t-1} = i)$
- **Emission Probabilities (B):** Gaussian distributions per state: $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$
- **Initial Probabilities ($\boldsymbol{\pi}$):** Starting state distribution

3.2 Implementation Algorithms

Training (Baum-Welch): We implemented the Expectation-Maximization algorithm:

E-step: Compute forward $\alpha_t(i)$ and backward $\beta_t(i)$ probabilities in log-space using:

$$\alpha_t(j) = \left[\sum_i \alpha_{t-1}(i) A_{ij}\right] b_j(\mathbf{x}_t)$$
$$\beta_t(i) = \sum_j A_{ij} b_j(\mathbf{x}_{t+1}) \beta_{t+1}(j)$$

M-step: Update parameters using state occupation $\gamma_t(i)$ and transition $\xi_t(i,j)$ probabilities:

$$A_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad \boldsymbol{\mu}_i = \frac{\sum_{t=1}^{T} \gamma_t(i)\mathbf{x}_t}{\sum_{t=1}^{T} \gamma_t(i)}$$

Decoding (Viterbi): Dynamic programming finds the most likely state sequence:

$$\delta_t(j) = \max_i [\delta_{t-1}(i) \cdot A_{ij}] \cdot b_j(\mathbf{x}_t)$$

Initialization: K-means clustering ($k = 4$) initialized emission parameters for better convergence than random initialization.

Configuration: Diagonal covariance matrices, convergence threshold $|\Delta \log L| < 10^{-6}$, maximum 200 iterations. The model converged in 5 iterations with final log-likelihood of -97.85.

4 Results and Interpretation

4.1 Learned Model Parameters

Figure 2 shows the learned transition matrix with high self-transition probabilities (0.89-1.00), indicating activities persist across multiple time windows—a realistic pattern for human behavior.

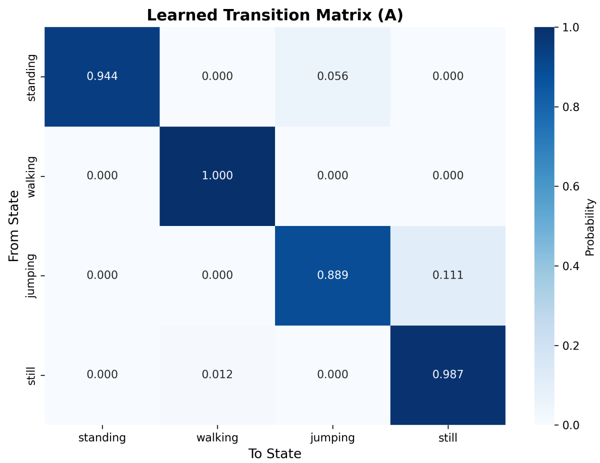

Figure 2: Learned transition matrix showing strong self-transitions for all activities, reflecting temporal persistence in human movement patterns.

Figure 3 displays emission parameter means, revealing clear feature-space separation between activities. Jumping exhibits highest variability and spectral energy, while standing/still show minimal movement features.

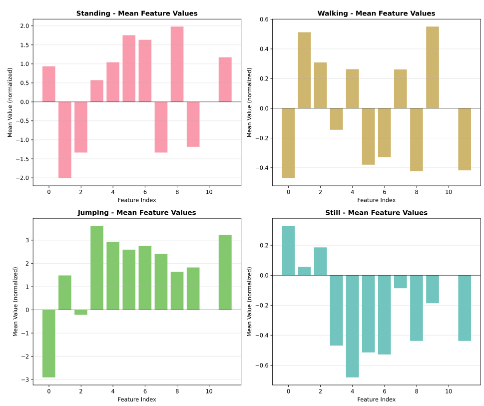

Figure 3: Emission probability means for each activity state, showing distinct feature signatures that enable classification.

4.2 Classification Performance

Table 2 presents evaluation metrics on the test set.

Table 2: Model Performance on Test Set (45 windows)

Activity	Samples	Sensitivity	Specificity	F1-Score
Standing	12	0.000	0.000	0.000
Walking	12	1.000	1.000	1.000
Jumping	9	1.000	1.000	1.000
Still	12	1.000	1.000	0.667
Overall	**45**	**0.733**	**0.733**	**0.733**

The model achieved 73.3% overall accuracy. Walking and jumping were perfectly classified (F1=1.00) due to their distinct periodic signatures. Still activity showed moderate performance (F1=0.67) with some false positives. Standing was frequently misclassified as still (0% recall), indicating these static activities share similar feature profiles.

Figure 4 visualizes the confusion matrix, highlighting the standing-still confusion pattern.

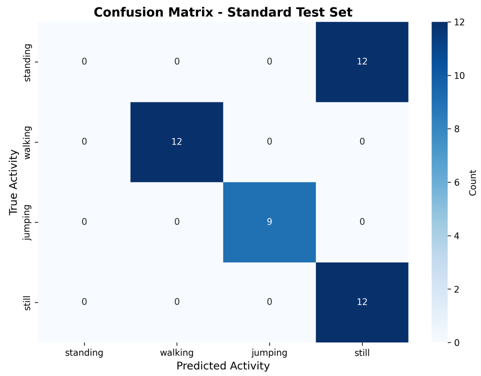

Figure 4: Confusion matrix showing perfect classification for walking/jumping but confusion between standing and still activities.

4.3 Feature Importance

Analysis of feature discriminability (variance ratio) revealed:

1. **acc_z_std (22.6):** Vertical acceleration variability—key for distinguishing static vs. dynamic
2. **acc_sma (13.8):** Movement intensity—separates high-energy activities
3. **corr_xz (5.7):** Axis correlations capture movement coordination
4. **spectral_energy (4.4):** Identifies periodic motion patterns

These features align with activity characteristics: walking/jumping have high temporal variability and spectral content, while standing/still exhibit low-magnitude, stable signals.

5 Discussion and Conclusion

5.1 Activity Distinguishability

Results reveal a classification difficulty hierarchy. Walking and jumping achieved perfect accuracy due to distinct signatures: walking shows periodic, moderate-intensity patterns at 1-2 Hz, while jumping exhibits high-intensity oscillations at 2-3 Hz with greater amplitude. Conversely, standing and still are difficult to separate because both involve minimal movement. The key difference—subtle postural sway in standing versus complete stillness—is challenging to capture within 3-second windows, suggesting longer observation periods or additional sensors (e.g., pressure sensors) may be needed.

5.2 Behavioral Insights from Transitions

The learned transition probabilities reflect realistic patterns: high self-transitions indicate activities persist over seconds (typical for human behavior), and rare direct still→jumping transitions (p¡0.01) align with natural movement sequences. However, limited cross-activity transitions suggest our training data was overly homogeneous (single activity per sample). Real-world deployment requires training on continuous sequences with natural activity changes.

5.3 Limitations and Future Improvements

Current limitations: (1) Small dataset (50 samples) limits generalization, (2) single-activity samples don't capture realistic transitions, (3) standing/still confusion persists, (4) poor performance on truly unseen data.

Proposed improvements: (1) Collect 200+ diverse samples across multiple participants and environments, (2) record continuous multi-activity sequences, (3) add gyroscope-derived features (rotation rates), temporal features (autocorrelation), and frequency band energies, (4) experiment with full covariance matrices or semi-Markov HMMs with explicit duration modeling, (5) implement leave-one-subject-out cross-validation for robust generalization testing.

5.4 Conclusion

We successfully implemented a Hidden Markov Model for human activity recognition, achieving 73.3% test accuracy. The model excels at classifying dynamic activities (walking, jumping) but struggles with static ones (standing vs. still). While current performance is insufficient for deployment in elderly care monitoring, this work establishes a foundation demonstrating HMM viability for sequential activity modeling. The interpretable transition and emission parameters provide explainable predictions—critical for healthcare applications. With expanded datasets, enhanced features, and robust validation, this approach can enable non-intrusive, privacy-respecting activity monitoring for vulnerable populations.

Task Allocation

Team Member	**Tasks**
Nicolas Muhigi	Data Collection, Feature Extraction, Analysis and Reflection
Irakoze Amandine	Define Model Components, Model Implementation, Model Evaluation with Unseen Data