Smartphone-Based Sit-to-Stand Power Assessment for Frailty Risk Screening

Ozair Ismail (06966695)

1. Motivation

Frailty affects approximately 10% of adults over 65 and is the strongest predictor of falls, hospitalization, and loss of independence. The 30-second chair stand test (30s CST) is a clinically validated assessment of lower-limb function, but traditional administration only counts repetitions using a stopwatch. Research demonstrates that sensor-derived parameters — peak acceleration, angular velocity, and movement variability — differentiate frailty levels even when repetition counts are identical [1]. We aim to build a smartphone application that transforms the standard chair stand test into a comprehensive frailty screening tool by combining ML-based repetition detection with clinically validated movement quality indicators.

The input to our system is raw triaxial accelerometer and gyroscope data (6 channels at 50Hz) from a waist-mounted smartphone. We use a Random Forest classifier, a Logistic Regression classifier, and a threshold-based baseline to output a binary prediction per time window: sit-to-stand (1) or other activity (0). Consecutive positive windows are then clustered into discrete rep events, and per-rep features are extracted for clinical assessment.

2. Dataset

We train and evaluate on the UCI HAPT dataset [3], which contains raw inertial signals from 30 participants (ages 19–48) wearing a waist-mounted Samsung Galaxy S II. The dataset includes 12 activity classes; we frame the task as binary classification: sit-to-stand (activity ID 8) vs. everything else. We identified 62 sit-to-stand segments averaging 2.59 seconds.

Preprocessing: Signals are segmented into 2.56-second windows (128 samples at 50Hz) with 50% overlap. Windows are labeled by majority vote with a 50% purity threshold. We initially used 80% purity per standard practice, but found this too aggressive: 16% of sit-to-stand segments (10/62) were too short to ever reach 80% purity, eliminating all positive examples for 3 subjects. The 50% threshold recovered these.

Feature extraction: We compute acceleration and gyroscope magnitude as orientation-independent L2 norms, yielding 8 channels. For each window, 6 statistics per channel (mean, std, min, max, range, energy) produce **48 features per window**. Final dataset: 17,453 windows — 126 sit-to-stand (0.72%), 17,327 other (99.28%). For external validation, we use SisFall [4]: 15 elderly subjects (ages 60–75), 149 sit-to-stand trials, resampled from 200Hz to 50Hz.

3. Method

We compare three approaches: (1) **Threshold Baseline** — predicts sit-to-stand if accel_mag_max and accel_mag_range exceed grid-searched thresholds (2 features, no learning); (2) **Logistic Regression** — linear classifier on all 48 features with class_weight='balanced' and StandardScaler; (3) **Random Forest** — 100 decision trees with class_weight='balanced', capable of learning nonlinear feature interactions. All evaluated using 30-fold LOSO-CV. We report precision, recall, F1, and PR-AUC for the sit-to-stand class at both window and event level.

4. Preliminary Experiments

4.1 Internal Validation (LOSO-CV)

Table 1: Window-level results (mean ± std, 30 folds)

Model	Precision	Recall	F1	PR-AUC
Threshold Baseline	0.011 ± 0.002	0.933 ± 0.135	0.022 ± 0.004	—
Logistic Regression	0.186 ± 0.070	0.933 ± 0.117	0.305 ± 0.096	0.675 ± 0.148
Random Forest	0.778 ± 0.377	0.454 ± 0.283	0.554 ± 0.301	0.742 ± 0.212

Random Forest achieves the best F1 (0.554) and PR-AUC (0.742). Post-processing interacts critically with model type: strict filtering destroys RF's sparse high-confidence predictions. With minimal post-processing:

Table 2: Event-level results (minimal post-processing)

Model	Rep MAE	Event Prec	Event Rec	Event F1
Threshold Baseline	14.73	0.101	0.823	0.180
Logistic Regression	3.23	0.151	0.387	0.217
Random Forest	0.60	0.909	0.645	0.755

RF achieves MAE of 0.60 reps (20/30 subjects exact) and 91% event precision. The pooled confusion matrix (Figure 1) shows only 5 false positives, all SITTING windows. Feature importance analysis reveals gyroscope x-axis energy as the top feature (54% of total Gini importance from gyroscope features).

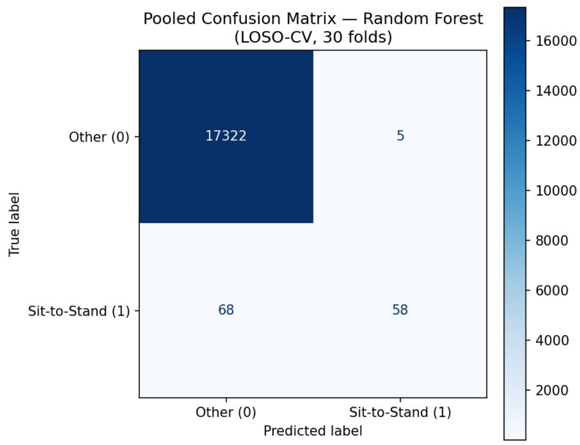

Figure 1: Pooled confusion matrix for Random Forest (LOSO-CV, 30 folds). Only 5 false positives, all sitting windows at the transition boundary.

4.2 External Validation (SisFall Elderly)

Final models trained on all UCI HAPT subjects tested on 149 SisFall elderly trials without retraining. RF completely fails (0% recall) — nonlinear boundaries overfit to training domain. LR generalizes well (92.6% recall). Threshold baseline generalizes best (96.0%). Classic bias-variance tradeoff.

4.3 Feature Ablation

Removing gyroscope features cuts RF event F1 from 0.755 to 0.341 (Δ0.41), confirming gyroscope captures rotational dynamics critical for sit-to-stand detection.

5. Next Steps

For the final report: (1) quality assessment layer — extract six per-rep clinical features (peak dynamic acceleration, relative muscle power via Alcázar equation, time-per-rep, peak gyroscope magnitude, CV, fatigue slope) mapped to Fried frailty dimensions; (2) three-tier output: rep count, power score, and movement quality flags; (3) full discussion of generalization findings and limitations.

Contributions

Solo project. All work performed by Ozair Ismail.

References

[1] Millor, N., et al. (2013). An evaluation of the 30-s chair stand test in older adults. *J. NeuroEng. Rehab.*, 10, 86.
[2] Alcázar, J., et al. (2021). Relative sit-to-stand power. *J. Cachexia, Sarcopenia and Muscle*, 12(4), 1013–1028.
[3] Reyes-Ortiz, J. L., et al. (2015). Transition-aware human activity recognition. *Neurocomputing*, 171, 754–767.
[4] Sucerquia, A., et al. (2017). SisFall: A fall and movement dataset. *Sensors*, 17(1), 198.
[5] Park, C., et al. (2021). Optimal sensor-based frailty phenotype assessment. *IEEE J. Biomed. Health Inform.*, 25(8), 3057–3067.
[6] Galán-Mercant, A. & Cuesta-Vargas, A. I. (2014). Mobile Romberg test assessment. *BMC Research Notes*, 7, 640.

[7] Van Lummel, R. C., et al. (2016). The instrumented sit-to-stand test. *PLoS ONE*, 11(7), e0157968.