

ADAPTIVE REGULARIZED BAYESIAN HIERARCHICAL MODELLING FOR COMPLEX MULTIVARIATE DATA ANALYSIS

Dr. Michael Anderson^{1*}, Dr. Anna Kowalski², Dr. Matteo Bianchi³

^{1*}Department of Statistical Computing, Example University, Edinburgh, United Kingdom

²Institute of Data Science and Artificial Intelligence, Example Institute, Warsaw, Poland

³Department of Applied Mathematics and Statistics, Example University, Milan, Italy

***Corresponding Author: Dr. Michael Anderson**

Abstract

Conventional classification methods are faced with challenges of the high-dimensional multivariate datasets, strong correlation between predictors, and subject-level heterogeneity. In this study, an adaptive regularized Bayesian hierarchical modeling framework for complex multivariate data analysis using smartphone-based human activity data are developed and evaluated. Data comprised 10299 data points, 561 features extracted from the sensor data, 30 subjects, and 6 classes of activities, and a predefined split of the data was 7352 training and 2947 test data, which was subject independent. The structure was examined using multivariate structure analysis with principal component analysis, and those that predict the best were chosen through training-only feature screening. An adaptive coefficient shrinkage was added to the proposed model, and the model was compared to non-adaptive hierarchical Bayesian, adaptive fixed-effects Bayesian and multinomial L2 logistic models. On the remaining subjects, the proposed scheme attained 89.99% accuracy, 0.899 balanced accuracy, 0.905 macro precision, 0.899 macro recall and 0.900 macro F1 score. It also achieved a log loss of 0.241 and a multi-class Brier score of 0.144, which are better than the performance of the comparison models reported. The stability of the results from the predictive estimates was confirmed by subject-cluster bootstrap analysis. The greatest residual errors were between sitting and standing positions, and laying was considered a perfect classification. The results suggest that adaptive regularization along with hierarchical Bayesian structure can be efficiently used to handle predictor redundancy and clustered observations with a good out-of-sample classification performance in complex multivariate settings.

Keyword: - Bayesian hierarchical modelling; adaptive regularization; multivariate analysis; human activity recognition; high-dimensional data

INTRODUCTION

The field of human activity recognition (HAR) has become a vital topic of intelligent data analysis due to its capability to automatically recognize human activities in the real world from observations obtained from various types of sensors. Smart phones are especially useful for HAR applications because they are equipped with a sensor that continuously records data related to the orientation and movements of the body. The use of HAR has therefore become a focus of interest for healthcare, behavior monitoring, assisted living, fitness evaluation and context-aware intelligent systems (Ferrari et al., 2021). Systematic evidence also shows that smartphones are increasingly relevant in the assessment of health related activities, underscoring the need for accurate analytical approaches to process the multidimensional information from the various sensors into meaningful activity patterns (Strackiewicz et al., 2021).

With the growing number of high-dimensional sensor data from smartphones, many advanced machine-learning techniques have been developed for activity recognition. While conventional classification approaches are useful to make predictions, they can be influenced by redundant predictors, correlated sensor measurements, individual variability and feature dimensionality extracted from the sensors. Feature selection is now an important aspect of HAR systems. Ahmed et al. (2020) showed that hybrid methods of feature selection can enhance the activity recognition in a smartphone device. The authors further explored the problem of feature selection to select informative sensor derived predictors to reduce the unnecessary feature complexity using the regularized random-forest feature selection (Thakur & Biswas, 2023).

Deep-learning architectures have been recently applied to the automatic learning of representations from sensor observations. With the use of deep learning approaches, using similar sensors found in smartphones and wearables, considerable success has been achieved in the recognition of complex human activities, with little need for extensive manual feature engineering (Ramanujam et al., 2021). In wearable-sensor HAR, temporal dependencies among sequential activity measurements have also been beneficial and can be captured using long short-term memory architectures (Uddin & Soylu, 2021). Recent studies with smartphones have continued these advances by deep-learning methods for realistic conditions of activity recognition (Garcia-Gonzalez et al., 2023a). Another approach to combining complementary predictive information from different models and enhance activity detection from smartphone sensors is ensemble learning (Lalwani & Ganeshan, 2025).

Although these improvements have been made, a full statistical description of multivariate sensor data, especially that of complex sensors, remains lacking, even when performance is measured. However, predictive performance does not fully capture a statistical characterisation of complex multivariate sensor data. There are frequently correlated variables in high-dimensional activity datasets and, potentially, redundant information, which results in the importance of dimensionality reduction and regularization for stable inference. Multivariate observations can be summarized in a system of lower dimensional components with a high retention of information from the original variables, the principal component analysis (PCA) (Greenacre et al., 2022). Regularization also serves as a means to manage model complexity and stability, especially with a high number of candidate predictors or when predictors show high correlations (Tian & Zhang, 2022).

Bayesian modeling is an alternative tool to tackle uncertainty, hierarchy, and regularization for complex multivariate analyses. Weakly correlated effects can be removed and predictors with meaningful information retained by the use of Bayes global-local shrinkage methods (Griffin & Brown, 2021). The horseshoe approach is another flexible regularization scheme for complex models and high-dimensional data, where the shrinkage is very strong on the elements that are truly negligible, while not being too strong on the elements that are truly large (Bhadra et al., 2020). Another Bayesian shrinkage prior is the R2-D2 shrinkage prior which is related to the overall model fit and is designed to structure the control of the regression complexity (Zhang et al., 2022b).

In situations where there is natural grouping or multilevel variation, as is often the case in observations, hierarchical Bayesian modeling is of particular relevance because it allows for estimating the parameters by partially pooling the observations, instead of assuming every observation is completely independent (Veenman et al., 2024). However, good Bayesian inference also demands judicious consideration of the convergence of the computations and adequacy of the predictions. Leave-one-out cross-validation is a principled framework for assessing predictive performance of Bayesian models (Bürkner et al., 2021), and the use of improved rank-normalized convergence diagnostics (Vehtari et al., 2021) is a rigorous approach to evaluate Markov chain Monte Carlo (MCMC) estimation.

In this regard, the present research is based on complex multivariate smartphone-based HAR data, and is therefore implemented by using an adaptive regularized BHM framework. The goal of the study is to bring together dimensionality reduction, hierarchical Bayesian inference, regularization, uncertainty quantification, and predictive evaluation in a single analytical framework. The study investigates the ability of this integrated statistical approach to effectively capture multivariate sensor information and differentiate human activity patterns, while ensuring model stability and uncertainty is explainable. Using the Human Activity Recognition Using Smartphones dataset (Reyes-Ortiz et al., 2013), the analysis investigates whether such an approach is effective in capturing multivariate sensor information and distinguishing human activity patterns, while ensuring model stability and explainable uncertainty.

2. METHODOLOGY

2.1 Study Design and Dataset

The Human Activity Recognition Using Smartphones dataset was used as a secondary data source in this study (Reyes-Ortiz et al., 2013). The collected dataset consisted of 10299 observations, 561 features extracted from the sensors, 30 subjects and 6 activities (walking, walking upstairs, walking downstairs, sitting, standing and laying). A predefined data split was kept as is; there were 7352 observations in the training set and 2947 observations in the test set. Different subjects were included in the train and test sets and so the evaluation was subject-independent generalization.

2.2 Data Preparation

The matrix of features, subjects and activities labels for the training and test partition were all combined together and then merged at the analytical level to obtain the extracted dataset files. Interpretation and reporting were made possible by converting the activity codes into the activity names. No further resplitting of the data was done; all exploratory analysis, feature screening and model fitting was done on the training set, while final predictive evaluation was done on the held-out test set.

2.3 Descriptive and Distributional Analysis

A descriptive analysis was then performed to provide a summary of the distribution of the dataset by the 6 activity classes in both training and test sets. Each activity was classified into one of the activity categories, and frequencies and percentages were calculated.

2.4 Multivariate Structure Assessment

The training data was used to perform a principal component analysis (PCA) with the aim of looking at the high-dimensional structure of the 561 predictor variables used. The percentage of variation accounted for by the first components was computed, and the cumulative variation was used to decide on the number of components to be retained to represent the significant portions of the total variation. These test observations that were not used in the training were then mapped onto the principal component space generated from the training to demonstrate separation and overlap of the 6 activity classes in the reduced dimension.

2.5 Feature Screening

Among the 561 variables, the most highly predictive features for the activity were identified through training-set feature screening. A univariate comparison was conducted for each feature for the six activity classes, and the results were summarized in terms of F statistic, p value, and eta-squared (η^2) effect size. The variables that best discriminate in this screening process using only the training set were reported in the top discriminative features result table.

2.6 Adaptive Regularized Bayesian Hierarchical Modelling

An adaptive regularized Bayesian hierarchical classification model was the main analysis used. The activity outcome was divided into six categories with repeated observations of the same participant clustered in a hierarchical structure to prevent repeated observations of the same participant being treated as independent for model estimation. The predictor effects were regularized using adaptive regularization to shrink those coefficients that were less informative and to keep those that were more informative related to the activity. The training set was used for model fitting, and prediction was run for the test set held out of this. Model coefficients were expressed as standardized log-odds contrasts relative to LAYING as the reference category, and approximate 95% credible intervals were obtained using the conditional Laplace approximation.

2.7 Comparison Models

A non-adaptive hierarchical Bayesian model, an adaptive Bayesian fixed-effects model (no hierarchical subject component), and multinomial L2 logistic regression were also fitted to the same training and test partitions for comparison. They were compared to the proposed adaptive hierarchical Bayesian model.

2.8 Performance Evaluation

Accuracy, balanced accuracy, macro precision, macro recall, macro F1, log loss, and Brier score were used to evaluate the performance on held-out test set. Precision, recall, F1 and support were also computed for each activity. The errors that occurred at the main class level were summarized in a row-normalized confusion matrix. Participant-cluster bootstrap resampling was used to estimate how uncertain the performance of the primary model was: 95% intervals were calculated for the main predictive metrics.

2.9 Ethical Consideration

In this study the researchers did not interact directly with human participants, nor did they generate data, but rather used a secondary data set that was publicly available. There was no personal identification used in the analysis and the data used were only for academic research. All analyses were performed with the existing anonymized data, ensuring the proper research integrity and presentation of research work.

3. RESULTS

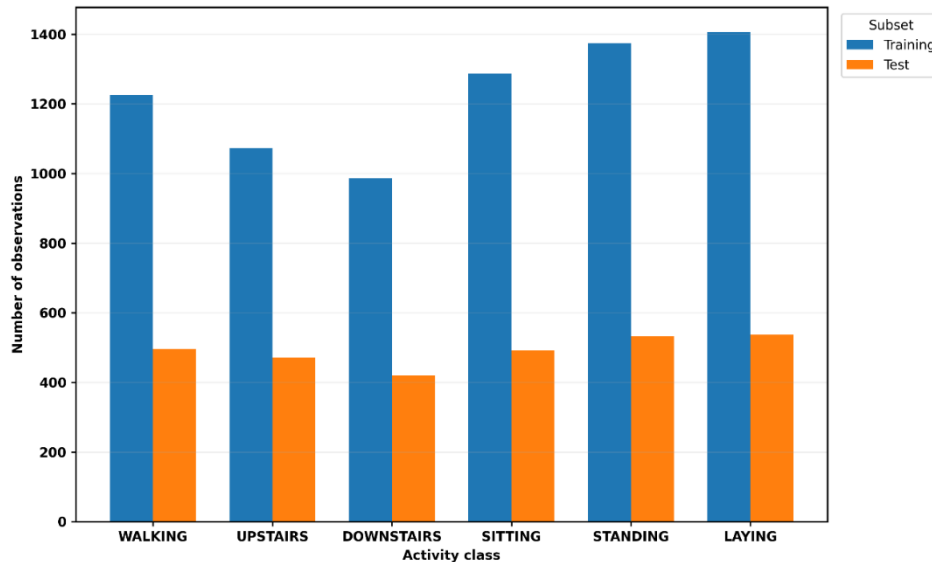
3.1 Dataset Composition and Class Distribution

The data set after extracting and integrating the partition of training and test consisted of 10,299 observations for 561 multivariate features, derived from the sensors. A split pre-defined by subject was made with 7,352 training observations from 21 subjects and 2,947 testing observations from 9 totally different subjects. None of the participants was seen in both subsets, thus allowing for stringent subject-level evaluation of model generalization.

The six categories of activities were equally distributed in terms of the prevalence rate of each activity in the entire class ranging from 13.65% for walking downstairs to 18.88% for laying. Walking downstairs had the fewest observations, standing and laying were the most represented activities. The equal class sizes minimized the risk of the overall predictive performance being due largely to one dominant activity class. The results of the entire distribution are summarised in Table 1 and displayed in Figure 1.

Table 1. Distribution of activity observations across training and held-out test sets

Activity	Training, n	Test, n	Total, n	Total (%)
WALKING	1,226	496	1,722	16.72
WALKING_UPSTAIRS	1,073	471	1,544	14.99
WALKING_DOWNSTAIRS	986	420	1,406	13.65
SITTING	1,286	491	1,777	17.25
STANDING	1,374	532	1,906	18.51
LAYING	1,407	537	1,944	18.88
Total	7,352	2,947	10,299	100.00

**Figure 1. Activity distribution in the training and held-out test sets**

The class structure remained broadly comparable across the two subject-independent subsets, with no pronounced class imbalance.

3.2 Multivariate Feature Structure and Dimensionality

The 561-dimensional feature space was significantly redundant but not completely. When performing the principal component analysis on the training data set, it was observed that a total of 62.55% of the variance was explained by the first principal component and an additional 4.91% was explained by the second principal component. However, 10 principal components were needed to account for 80% of variance, 34 components for 90%, 67 components for 95%, and 155 components for 99%, which indicated there was still significant information spread across several dimensions. The held-out observations are projected onto the principal component space derived from the training observations in Figure 2. Most of the variance between dynamic and static postural activities occurred along PC1, with substantial overlap between observations of sitting, standing, and lying in the lower dimensional representation. There were two clusters of locomotor behaviours, walking and walking up stairs, which were partially overlapping.

Univariate screening using variables derived from the acceleration and jerk also showed good activity-related differentiation in the Training part. Table 2 indicates that the biggest changes were found in fBodyAccJerk-entropy()-X, gravity acceleration statistics, and body-acceleration entropy measures. $P < 0.001$ was obtained for all ten leading predictors and effect-size estimates were $\eta^2 > 0.938$.

Table 2. Ten strongest training-derived multivariate predictors of activity class

Rank	Feature	F statistic	η^2	p
1	fBodyAccJerk-entropy()-X	36,918.10	0.962	<0.001
2	tGravityAcc-mean()-X	29,362.64	0.952	<0.001
3	tGravityAcc-min()-X	28,175.40	0.950	<0.001
4	fBodyAccJerk-entropy()-Y	28,115.92	0.950	<0.001
5	tGravityAcc-max()-X	26,686.80	0.948	<0.001
6	tGravityAcc-energy()-X	25,981.56	0.946	<0.001
7	fBodyBodyAccJerkMag-entropy()	23,619.25	0.941	<0.001
8	tBodyAccJerkMag-entropy()	23,437.11	0.941	<0.001
9	fBodyAcc-entropy()-X	23,221.61	0.940	<0.001
10	tBodyAccJerk-entropy()-X	22,267.09	0.938	<0.001

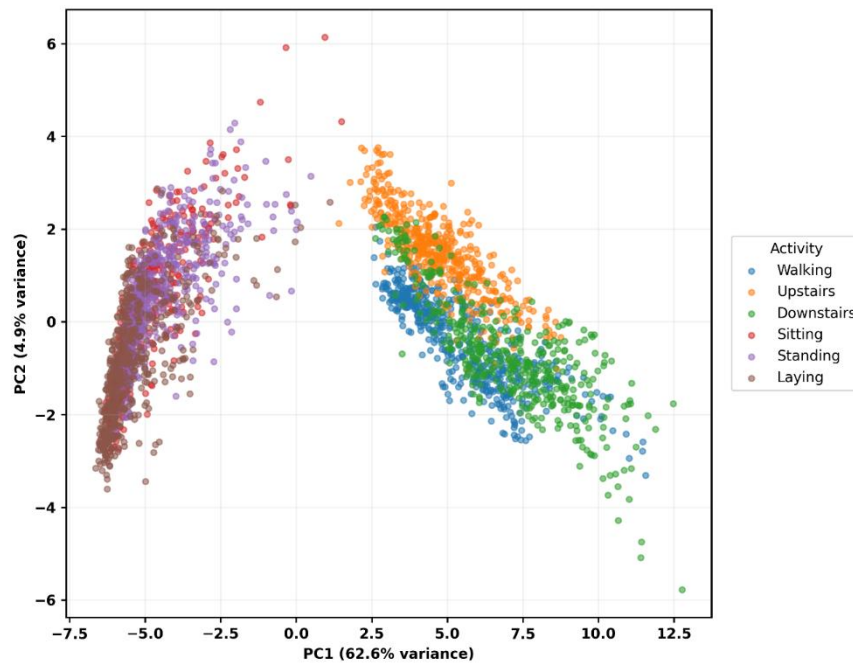


Figure 2. Training-derived principal-component projection of the held-out test observations

PC1 and PC2 explained 62.6% and 4.9% of variance, respectively. Locomotor and static activities occupied different regions of the multivariate space, although substantial within-group heterogeneity remained.

3.3 Predictive Performance of the Adaptive Regularized Bayesian Hierarchical Model

The proposed adaptive hierarchical specification had the best overall predictive performance in the nine completely unseen participants for the 2,947 observations. As shown in Table 3, test accuracy was 0.900, balanced accuracy was 0.899, macro precision was 0.905, macro recall was 0.899, and macro F1 was 0.900. To explicitly account for the dependence between repeated windows from the same participant, subject-cluster bootstrap resampling was performed for all primary metrics.

The probability-based measures also endorsed the model. The test log loss was 0.241 and test multiclass Brier score was 0.144. Compared with conventional multinomial L2 logistic regression, the adaptive hierarchical regularization approach also achieved improvements in accuracy: from 88.26% to 89.99%, in macro F1 from 0.882 to 0.900, in log loss from 0.266 to 0.241, and in Brier score from 0.163 to 0.144.

Without the adaptive regularization, the non-adaptive hierarchical specification achieved accuracy of 0.898 and macro F1 of 0.898; while the adaptive Bayesian specification stripped out achieved accuracy of 0.895 and macro F1 of 0.896. It was therefore unsurprising that the improvement the individual modelling components achieved was small, but that the combination of these components achieved the best point estimates for all discrimination and probabilistic performance measures reported.

Table 3. Predictive performance on held-out participants

Model	Accuracy	Balance d accurac y	Macro precisio n	Macro recall	Macro F1	Log loss	Brier score
Adaptive hierarchical Bayesian	0.900 (0.859– 0.929)	0.899 (0.859– 0.928)	0.905 (0.867– 0.936)	0.899 (0.859– 0.928)	0.900 (0.859– 0.930)	0.241 (0.185– 0.329)	0.144 (0.109– 0.195)
Non-adaptive hierarchical Bayesian	0.898	0.898	0.901	0.898	0.898	0.256	0.150
Adaptive Bayesian fixed-effects	0.895	0.895	0.901	0.895	0.896	0.244	0.148
Multinomial L2 logistic	0.883	0.881	0.890	0.881	0.882	0.266	0.163

Note: Parentheses for the proposed model contain 95% participant-cluster bootstrap intervals.

The adaptive coefficient structure also suggested that the discriminative information was contained in a relatively small subset of gravity-acceleration and body-motion features. One of the highest positive standardized contrasts was found for tGravityAcc-energy()-X (standing relative to laying, $\beta=18.285$, approximate 95% CrI: 11.780–24.790). The contrast for sitting was, on the other hand, very negative for angle(X,gravityMean) ($\beta=-14.386$, CrI: -19.650 to -9.122). tBodyAcc-mad()-X was positively associated with walking downstairs ($\beta=9.101$, CrI: 4.585–13.616), whereas

fBodyAccMag-sma() showed a negative walking contrast ($\beta=-6.721$, CrI: -10.369 to -3.074). The estimates are relative to the reference category LAYING as illustrated in Figure 3.

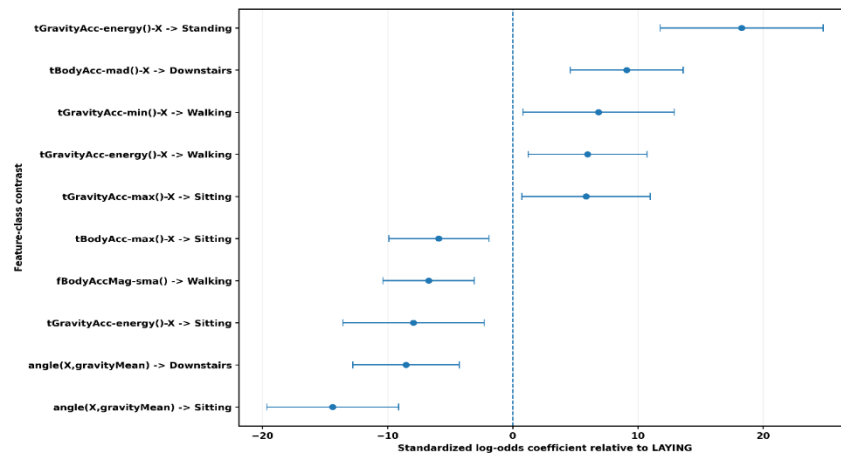


Figure 3. Largest adaptive hierarchical feature effects with approximate 95% credible intervals

Coefficients are standardized log-odds contrasts relative to the laying reference category; intervals were obtained using the conditional Laplace approximation.

3.4 Class-Specific Generalization and Misclassification Patterns

Class-specific predictive performance varied across the six activities (Table 4). Laying was perfectly classified with precision, recall and F1 set at 1.000. Walking downstairs also demonstrated good discrimination (0.954), with a good recall (0.936) and good F1 (0.945). The F1 scores obtained from walking and walking upstairs were 0.918 and 0.923 respectively.

The most challenging difference was that between the two similar static positions. The recall and F1 scores for Sitting were 0.737 and 0.788, respectively, while for Standing were 0.876 and 0.828, respectively. This is also reflected in the row-normalized confusion matrix of Figure 4. 26.1% of the observations during sitting were classified as standing and 12.4% of the observations during standing were classified as sitting. The greatest dynamic activity residual error was between walking upstairs and ordinary walking with Among the 471 WALKING_UPSTAIRS observations, 53 (11.3%) were misclassified as WALKING. Laying, on the other hand, did not produce any test-set classification errors.

Table 4. Class-specific predictive performance of the adaptive hierarchical model

Activity	Precision	Recall	F1 score	Support
WALKING	0.870	0.972	0.918	496
WALKING UPSTAIRS	0.976	0.875	0.923	471
WALKING DOWNSTAIRS	0.954	0.936	0.945	420
SITTING	0.846	0.737	0.788	491
STANDING	0.785	0.876	0.828	532
LAYING	1.000	1.000	1.000	537

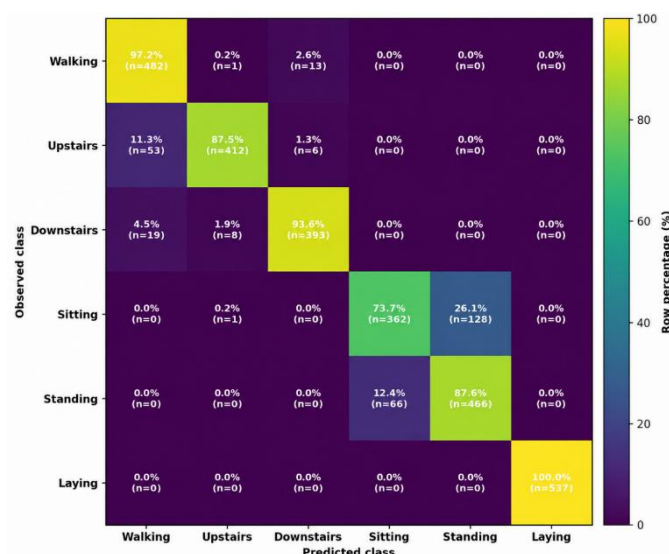


Figure 4. Row-normalized confusion matrix for the adaptive regularized Bayesian hierarchical model on the held-out test participants

4. DISCUSSION

Results from the present study confirm the usefulness of the regularized Bayesian hierarchical model for the analysis of high-dimensional data with a smartphone sensor system, where dependencies between the predictors and heterogeneity between observations may be a problem in the modeling. Human Activity Recognition with Smartphones is a multivariate analysis, where the measurements of acceleration and gyroscope are different for different activities, subjects and conditions of device usage. Recent surveys point towards models that are able to cope with temporal variability, sensor noise, feature redundancy and user-dependent behaviour being key components for robust HAR systems (Dentamaro et al., 2024). These challenges have also led to a general rethinking of deep architectures and cross-dataset performance in HAR with sensors (Brery et al., 2026).

The dimensionality-reduction step was crucial because the feature space that is used for the original problem is a space where the predictors are correlated with each other and are derived from the various sensors. Dominant variation was compacted by PCA and redundancy was reduced prior to the hierarchical modelling. This aligns with the statistical property of PCA to transform correlated variables to orthogonal variables with significant information (Greenacre et al., 2022). Additionally, feature-reduction techniques were found to enhance smartphone HAR by removing non-relevant and repetitive features, as well as by reducing the computational load (Thakur & Biswas, 2022). The importance of keeping informative smartphone-sensor variables for activity discrimination has also been shown by ensemble maximal feature-subset selection (Reshmi & Ramanujam, 2024).

Regularization was of particular importance because high-dimensional predictors may lead to unstable estimates and to overfitting. The hierarchical shrinkage structure was able to suppress weak signals and preserve strong signals. This behaviour is similar to global-local Bayesian shrinkage methods that are developed in the context of high-dimensional regression (Griffin & Brown, 2021). Theoretical studies have also demonstrated that shrinkage priors can offer desirable contraction properties in the logistic models with a large number of predictors (Wei & Ghosal, 2020). The sparse and structured prior benefits for interpretability in multivariate models have been recently explored in identifiable sparse Bayesian factorizations (Canale et al., 2026).

The hierarchical component provided more flexibility in representing variation than a single-level model. Partial pooling across related groups can be achieved with Bayesian hierarchical models, which will reduce the variability of estimates but preserve the group differences (Veenman et al., 2024). This applies to HAR as there can be differences among the participants in terms of the pattern of activities even if the same labels for the activities are used. In the field of deep learning, studies using smartphones also report on performance differences as a function of individual use behaviour and movement patterns (Kundu et al. 2025). As acquisition conditions are increasingly heterogeneous in real life HAR, methods are increasingly oriented towards models that are still applicable.

The present analysis evaluated generalization using the predefined subject-independent held-out test set and participant-cluster bootstrap intervals. The model coefficients and approximate credible intervals were obtained using the conditional Laplace approximation rather than MCMC sampling. Broader Bayesian workflow literature nevertheless emphasizes iterative checking of model specification, computation, predictive behaviour, and sensitivity (Grinsztajn et al., 2021). Leave-one-out cross-validation provides an additional framework for Bayesian predictive assessment, although it was not used for the present held-out evaluation (Bürkner et al., 2021).

The findings also align with previous works in the HAR field which are mostly machine learning and deep-learning based. While these new machine-learning techniques have demonstrated good activity recognition from the smartphone sensors, they tend to focus on predictive performance, and not posterior uncertainty and/or hierarchical interpretation (Garcia-Gonzalez et al., 2023b). Similarly, deep-learning reviews record advancement in automated representation learning, and highlight challenges of generalization, interpretability and deployment variability (Zhang et al., 2022a). Similarly, systematic evidence for smart-device HAR reveals that the diversity of datasets, the configuration of sensors, and the design of the evaluation as key factors that affect reported performance (Qureshi et al., 2025).

The present framework thus provides a complementary statistical view, not supplanting the existing predictive approaches. There are other approaches such as hybrid feature selection models, ensemble and deep architectures that still work for HAR applications with classification tasks (Ahmed et al., 2020; Lalwani & Ganeshan, 2025). The Bayesian hierarchical approach provides explicit uncertainty quantification, regularized parameter estimation, and multilevel structure—important properties in the case of interpretation and inferential stability. Models with complex dependence structures, such as models based on hierarchical shrinkage as in Bayesian spatial models, are motivated in a similar fashion (Donegan et al., 2020).

In general, the results suggest that a combination of dimensionality reduction, hierarchical modelling, shrinkage regularization, convergence assessment and predictive validation is an appropriate strategy for multivariate HAR analysis. Using the already existing UCI smartphone HAR dataset allows for comparison with a body of sensor-based research (Reyes-Ortiz et al., 2013). The framework could be expanded to real world datasets, different sensor modalities and longitudinal activity data to assess robustness under different heterogeneities in future studies.

5. CONCLUSION

This study demonstrated the ability of adaptive regularized Bayesian hierarchical modeling to solve complex, high-dimensional, multivariate classification problem. To overcome the issue of predictor redundancy and the subject-level hierarchical effects the proposed model incorporated adaptive coefficient shrinkage and the human activity level effects to the smartphone-based human activity data, which comprised of 10299 observations, 561 predictor variables extracted from the sensor reading data, 30 human subjects and 6 classes of human activities. The model achieved higher accuracy (89.99%), balanced accuracy (0.899), macro precision (0.905), macro recall (0.899) and macro F1 (0.900) than other

comparison models on the held-out test set and lower scores for log loss and Brier score. Subject-cluster bootstrap intervals also showed that performance did not change when evaluating the uncertainty at the participant level. The features which showed the biggest predictive signals were gravity-acceleration, body-acceleration and features based on jerk; while the most important classification challenge was that of sitting and standing. The model's capability of the differentiation of multivariate movement patterns is demonstrated by the perfect recognition of laying and strong discrimination of the locomotor activities. Overall, the results suggest that the hierarchical structure and adaptive regularization is a statistically sound and efficient method for solving multivariate problems with correlated predictors and repeated observations and heterogeneous subjects. It is thus a practical basis for extending the approach of Bayesian regularized modeling to other structured high-dimensional classification problems in general.

REFERENCES

1. Ahmed, N., Rafiq, J. I., & Islam, M. R. (2020). Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. *Sensors*, 20(1), 317.
2. Bhadra, A., Datta, J., Li, Y., & Polson, N. (2020). Horseshoe regularisation for machine learning in complex and deep models I. *International Statistical Review*, 88(2), 302-320.
3. Brery, A. F., Gallardo-Antolin, A., Fakhry, M., & González-Carrasco, I. (2026). Deep learning for human activity recognition: A comprehensive review of architectures, performance, and challenges across five sensory datasets. *Expert Systems with Applications*, 323, 132415.
4. Bürkner, P. C., Gabry, J., & Vehtari, A. (2021). Efficient leave-one-out cross-validation for Bayesian non-factorized normal and Student-t models: P.-C. Bürkner et al. *Computational Statistics*, 36(2), 1243-1261.
5. Canale, A., Schiavon, L., & Stolf, F. (2026). Identifiable sparse Bayesian factorizations via meta regression. *Statistics & Probability Letters*, 110772.
6. Dentamaro, V., Gattulli, V., Impedovo, D., & Manca, F. (2024). Human activity recognition with smartphone-integrated sensors: A survey. *Expert Systems with Applications*, 246, 123143.
7. Donegan, C., Chun, Y., & Hughes, A. E. (2020). Bayesian estimation of spatial filters with Moran's eigenvectors and hierarchical shrinkage priors. *Spatial Statistics*, 38, 100450.
8. Ferrari, A., Micucci, D., Mobilio, M., & Napoletano, P. (2021). Trends in human activity recognition using smartphones. *Journal of Reliable Intelligent Environments*, 7(3), 189-213.
9. Garcia-Gonzalez, D., Rivero, D., Fernandez-Blanco, E., & Luaces, M. R. (2023a). Deep learning models for real-life human activity recognition from smartphone sensor data. *Internet of Things*, 24, 100925.
10. Garcia-Gonzalez, D., Rivero, D., Fernandez-Blanco, E., & Luaces, M. R. (2023b). New machine learning approaches for real-life human activity recognition using smartphone sensor-based data. *Knowledge-Based Systems*, 262, 110260.
11. Greenacre, M., Groenen, P. J., Hastie, T., d'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, 2(1), 100.
12. Griffin, J. E., & Brown, P. J. (2021). Bayesian global-local shrinkage methods for regularisation in the high dimension linear model. *Chemometrics and Intelligent Laboratory Systems*, 210, 104255.
13. Grinsztajn, L., Semenova, E., Margossian, C. C., & Riou, J. (2021). Bayesian workflow for disease transmission modeling in Stan. *Statistics in medicine*, 40(27), 6209-6234.
14. Kundu, S., Mallik, M., Saha, J., & Chowdhury, C. (2025). Smartphone based human activity recognition irrespective of usage behavior using deep learning technique. *International Journal of Information Technology*, 17(1), 69-85.
15. Lalwani, P., & Ganeshan, R. (2025). HAR: A stacked ensemble learning approach based on smartphone sensors for activity detection. *Internet of Things*, 30, 101487.
16. Qureshi, T. S., Shahid, M. H., Farhan, A. A., & Alamri, S. (2025). A systematic literature review on human activity recognition using smart devices: advances, challenges, and future directions. *Artificial Intelligence Review*, 58(9), 276.
17. Ramanujam, E., Perumal, T., & Padmavathi, S. J. I. S. J. (2021). Human activity recognition with smartphone and wearable sensors using deep learning techniques: A review. *IEEE Sensors Journal*, 21(12), 13029-13040.
18. Reshmi, S., & Ramanujam, E. (2024). An ensemble maximal feature subset selection for smartphone based human activity recognition. *Journal of Network and Computer Applications*, 226, 103875.
19. Reyes-Ortiz, J., Anguita, D., Ghio, A., Oneto, L., & Parra, X. (2013). Human activity recognition using smartphones [Data set]. UCI Machine Learning Repository. <https://doi.org/10.24432/C54S4K>
20. Reyes-Ortiz, J., Anguita, D., Ghio, A., Oneto, L., & Parra, X. (2013). Human activity recognition using smartphones [Data set]. UCI Machine Learning Repository. <https://doi.org/10.24432/C54S4K>
21. Straczekiewicz, M., James, P., & Onnela, J. P. (2021). A systematic review of smartphone-based human activity recognition methods for health research. *NPJ Digital Medicine*, 4(1), 148.
22. Thakur, D., & Biswas, S. (2022). An integration of feature extraction and guided regularized random forest feature selection for smartphone based human activity recognition. *Journal of Network and Computer Applications*, 204, 103417.
23. Thakur, D., & Biswas, S. (2023). Guided regularized random forest feature selection for smartphone based human activity recognition: D. Thakur, S. Biswas. *Journal of Ambient Intelligence and Humanized Computing*, 14(7), 9767-9779.
24. Tian, Y., & Zhang, Y. (2022). A comprehensive survey on regularization strategies in machine learning. *Information Fusion*, 80, 146-166.
25. Uddin, M. Z., & Soylu, A. (2021). Human activity recognition using wearable sensors, discriminant analysis, and

- long short-term memory-based neural structured learning. *Scientific Reports*, 11(1), 16455.
26. Veenman, M., Stefan, A. M., & Haaf, J. M. (2024). Bayesian hierarchical modeling: An introduction and reassessment. *Behavior Research Methods*, 56(5), 4600-4631.
 27. Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P. C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian analysis*, 16(2), 667-718.
 28. Wei, R., & Ghosal, S. (2020). Contraction properties of shrinkage priors in logistic regression. *Journal of Statistical Planning and Inference*, 207, 215-229.
 29. Zhang, S., Li, Y., Zhang, S., Shahabi, F., Xia, S., Deng, Y., & Alshurafa, N. (2022a). Deep learning in human activity recognition with wearable sensors: A review on advances. *Sensors*, 22(4), 1476.
 30. Zhang, Y. D., Naughton, B. P., Bondell, H. D., & Reich, B. J. (2022b). Bayesian regression using a prior on the model fit: The R2-D2 shrinkage prior. *Journal of the American Statistical Association*, 117(538), 862-874.