

Analysis of In-Home Movement Patterns for Depression Assessment in Older Adults – A Feasibility Study

Mitchell DENNIS ^b, Deepa PRABHU ^{a,1}, Stephanie BAKER ^b and David SILVERA-TAWIL ^a

^a *Commonwealth Scientific and Industrial Research Organisation*

^b *James Cook University*

ORCID ID: Mitchell Dennis <https://orcid.org/0009-0001-5313-8121>, Deepa Prabhu <https://orcid.org/0000-0002-0257-2200>, Stephanie Baker <https://orcid.org/0000-0003-0467-7791>, David Silvera-Tawil <https://orcid.org/0000-0003-2653-0141>

Abstract. Depression significantly impacts the wellbeing of older Australians, posing considerable challenges to their overall quality of life. This study aimed to detect in-home movement patterns of participants that could be indicative of depressive states. Utilising data collected over a 12-month period via smart home ambient sensors, this feasibility study conducted a comparative analysis using machine learning techniques on features derived from motion sensors, sociodemographic variables, and the Geriatric Depression Scale. Three machine learning models, specifically Extreme Gradient Boost (XGBoost), Random Forest (RF), and Logistic Regression (LR), were implemented. Results showed that the performance of XGBoost was relatively higher compared to RF and LR, with an Area Under the Receiver Operating Characteristic Curve (AUROC) value of 0.67. Feature analysis indicated that bathroom and kitchen movements and the level of home care support were among the top influential features influencing depression assessment. This is consistent with clinical evidence on appetite, hygiene, and overall mobility changes during depression. These findings underscore the feasibility of leveraging in-home movement monitoring as an indicator of health risks among older adults.

Keywords. depression, older adults, machine learning, smart home, motion sensor, health informatics

1. Introduction

Depression is characterised by persistent feelings of sadness and diminished interest in activities, lasting for at least two weeks [1]. Depression is highly prevalent among the geriatric population, with a reported rate of 10%-40% [2]. Conventional depression assessment methods rely on self-reported measures or direct observation by healthcare professionals, which can be biased or challenging to complete [4]. However, emerging evidence suggests that changes in daily living activities, as monitored by smart home technologies, can provide valuable insights into mental health risks such as depression [4, 5]. This study investigates the feasibility of using non-invasive, ambient smart home

¹ Corresponding Author: Deepa Prabhu, Deepa.Prabhu@csiro.au.

sensors to identify changes in in-home movement patterns associated with depressive states among older adults.

Prior research in this space predominantly focused on wearable devices to monitor mental health risks [6-8]. However, this population has several challenges for adopting wearable devices, including non-compliance [9]. As a result, current sensor-based depression assessment methods rely on public datasets of wearable sensor data to assess activity changes as indicators of depressive symptoms [10]. The integration of machine learning (ML) techniques for depression assessment using data from wearable sensors has been explored in previous research, with several models, including Extreme Gradient Boost (XGBoost), Random Forest (RF), Logistic Regression (LR), and Convolutional Neural Networks, often cited for their effectiveness [11, 12]. However, work investigating the use of ambient sensors to assess depression is limited. The use of motion sensors with statistical approaches demonstrates the possibility of assessing depression from in-home movement patterns [4]. This study extends that work by using ML to analyse motion sensor data to identify activity changes associated with depressive episodes.

Smart home systems, such as the Smarter Safer Homes (SSH) platform developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) [13], use interconnected devices and sensors on an Internet of Things (IoT) framework. These systems offer continuous, discreet monitoring of environmental conditions and residents' activities, facilitating the early detection of changes in daily routines that could indicate health risks such as depression. These technologies offer advantages over wearables, including overcoming compliance issues, while meeting the need for continuous, passive data collection in geriatric care contexts. The potential of smart home sensors for discreetly monitoring depression in older people living at home remains largely unexplored.

This paper explored the feasibility of applying machine learning techniques to investigate the correlation between survey data — including Geriatric Depression Scale (GDS) scores and sociodemographic data — and physical activity collected using ambient motion sensors. Specifically, this study aims to assess the feasibility of detecting depression-related in-home movement changes among community-dwelling older adults.

2. Methods

2.1. Data

This study used data collected from the Dementia and Aged Care Services (DACs) trial, a randomised control trial conducted in 2019, which implemented the SSH platform in community settings [13]. This study included the following data collected for one year during the DACs trial: (i) GDS Scores [14], (ii) sociodemographic variables such as gender, Aboriginal and Torres Strait Islander status, living arrangements (e.g., alone or with others), home care package, marital status, accommodation setting (e.g., privately owned home vs rental) and age, and (iii) continuous ambient sensor data (e.g., motion, room temperature and humidity, room light). GDS scores and sociodemographic data were obtained at three trial stages (start of the trial, after six months, and after 12 months).

GDS scores served as the ground truth for participant classification into depressed or non-depressed categories. Motion sensor firings measured the frequency of sensor activations to reflect participants' movement patterns. This study only considered motion

data on 14-day intervals surrounding GDS assessments. The 14-day period was used in line with the World Health Organization (WHO) categorisation of depression as a condition that lasts for at least two weeks [3].

2.2. Participants

The DACS trial collected data from the homes of 98 older adults. To ensure the precision of motion sensor data, this study excluded participants living in multi-occupancy settings ($n = 47$) due to the technology's inability to identify between movements from the people of interest in multi-occupancy scenarios. Additionally, 21 participants were excluded due to the lack of comprehensive motion data coinciding with the GDS assessment periods. Consequently, the analysis was conducted on 30 participants (mean age 85 years, 19 Females, 11 Males).

2.3. Data Pre-processing

Data pre-processing involved computing GDS scores, motion sensor data preparation, and feature matrix generation. GDS scores were calculated from the GDS surveys following the classification system outlined in the GDS (Short Form) [14]. Participants with scores ≤ 5 and >5 were classified as depressed, and not depressed, respectively. The total number of motion sensor firings (i.e., activations) per room per hour were extracted and averaged for each 14-day period. Additionally, the sociodemographic variables described in Section 2.1 were included as input features. Per-feature normalisation utilising the min-max method was applied to the resulting feature matrix. Missing data were handled by including the mean value of each feature by implementing the SimpleImputer function from the SciKit-learn library in Python. A notable class imbalance was observed, with nine depressed and 32 not-depressed instances. Random undersampling was applied to address this imbalance.

2.4. Machine Learning Models and Evaluation

Three ML models, including XGBoost, RF, and LR, were evaluated for depression detection based on prior research highlighting their efficacy with similar data types [6]. Due to the small dataset, a stratified five-fold cross-validation was employed to assess model performance. The ML algorithms' performances were evaluated using confusion matrices, accuracy, recall, specificity, and precision. A SHAP (SHapley Additive exPlanations) ML interpretability technique [16] was also used on the best-performing model to analyse the impact of various features on the model's output.

3. Results

As shown in Table 1, XGBoost demonstrated the highest recall score, compared to RF and LR, in its ability to accurately identify depression cases. The precision values also mirrored this trend. The AUROC scores obtained for each model reflect their ability to discriminate between positive and negative instances of depression. The XGBoost model achieved a 0.67, indicating moderately improved discrimination performance compared to the RF and LR models.

Table 1. Key Metrics of the evaluated machine learning algorithms.

MODEL	KEY METRICS				
	Accuracy	Recall	Specificity	Precision	AUROC Score
XGBoost	61.5%	44.4%	70.6%	44.4%	0.67
RF	60.0%	33.3%	75.0%	42.9%	0.58
LR	60.0%	33.3%	75.0%	42.9%	0.57

Feature importance analysis conducted on XGBoost, revealed underlying patterns indicative of depression of firings per hour and home care package level exhibited significant correlations with depression. Additionally, the Beeswarm plot shown in Figure 1, showed bathroom and kitchen firings as the leading predictive features, aligning with findings from the confusion matrix shown in Table. 2. Moreover, features such as the average total number of revealed features such as increased bedroom use and age as indicators of depression among older adults.

Table 2: Confusion matrix for XGBoost algorithm.

	Predicted	
	Positive (Depressed)	Negative (Not depressed)
True		
Positive (Depressed)	4	5
Negative (Not depressed)	5	12

4. Discussion

The study showed varying levels of performance among the assessed ML models. Notably, XGBoost performed marginally better in terms of accuracy, recall, and precision in predicting depressed cases. XGBoost’s performance was further evidenced by its AUROC value of 0.67. While accuracy scores offered valuable insights, their interpretation is nuanced due to a small imbalanced dataset. The performance scores of RF and LR models were close to that of XGBoost, suggesting their potential utility in depression assessment leveraging ambient smart home sensor technologies with larger datasets. As can be seen from the confusion matrix presented in Table 2, XGBoost model performed better at identifying non-depressed cases in with depressed cases, suggesting the need for further analysis with larger and balanced datasets in future work. These results also align with results previously reported results [6, 11] that recognised XGBoost as one of the best-performing models in similar applications.

Results from feature importance analysis align with previous research, indicating that reductions in motion sensor firings may signify depressive states [4]. These results highlight the significance of average total motion sensor firings per hour and average kitchen motion sensor firings per hour in predicting depression. These findings also resonate with established symptoms of depression, such as changes in activity levels and alterations in eating behaviours, as noted by the WHO [3] and previous research [5].

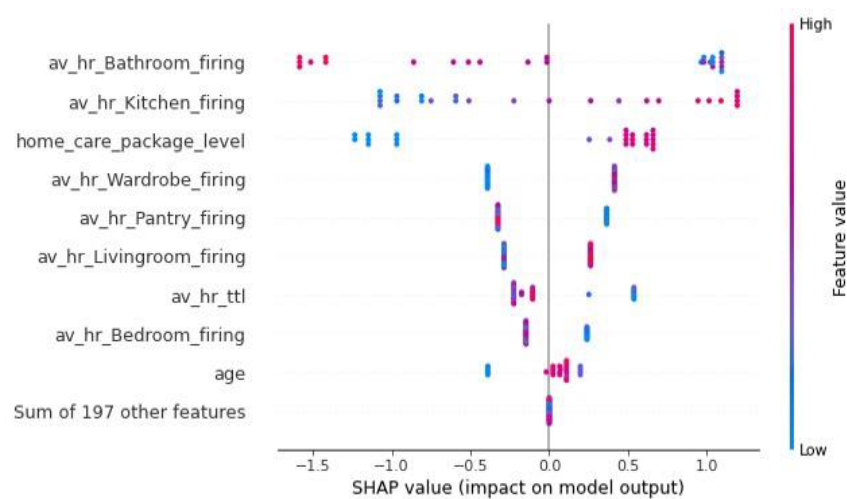


Figure 1. Beeswarm plot for XGBoost algorithm showing feature importance.

Additionally, the level of the home care package (indicative of the amount of care needed) and age, features that contributed to the model, inform the influence of physical status and age-related factors that could contribute to depression. Thus, the alignment between existing research and identified features supports the hypothesis that ML could be used to discern underlying patterns and symptoms of depression from motion sensors.

4.1. Limitations

The constraints of the small and imbalanced dataset impacted the performance of the ML models applied in this study. Additionally, imperfections in data, such as potential disturbances from guests, carers, or pets, impact the interpretation of the data presented in this paper. Moreover, the study’s focus on motion sensor data alone may overlook valuable insights from other sensor data.

5. Conclusions

In conclusion, this feasibility study demonstrated the potential of identifying signs of depression through changes in in-home movement patterns as measured by smart home motion sensors using machine learning. Three machine learning models, XGBoost, LR, and RF, were employed on ambient motion sensor data and sociodemographic variables collected from independent living older adults. XGBoost showed the most promising performance in identifying depression cases. Feature analysis highlighted critical activity indicators (e.g., bathroom, kitchen firings) supporting depression assessment. Future work will focus on leveraging machine learning with larger, more balanced datasets to enhance learning effectiveness and prediction accuracy to help identify key features contributing to depression and to inform targeted interventions.

Acknowledgements

We would like to acknowledge the CSIRO DACS trial team for contributing the data for this study.

References

- [1] World Health Organization, Mental health and substance use. World Health Organization. Available from: <https://www.who.int/westernpacific/about/how-we-work/programmes/mental-health-and-substance-abuse>
- [2] Gide DN, El-Den S, Lee YLE, Gisev N, Ou K, O'Reilly CL. Community pharmacists' acceptability of pharmacist-delivered depression screening for older adults: a qualitative study. *Int J Clin Pharm*. 2023 Oct; 45(5):1144–1152, doi: 10.1007/s11096-023-01581-1.
- [3] Conijn JM, Emons WHM, Page BF, Sijtsma K, Van der Does W, Carlier IVE, Giltay EJ. Response inconsistency of patient-reported symptoms as a predictor of discrepancy between patient and clinician-reported depression severity. *Assessment*. 2018 Oct;25(7):917-928, doi: 10.1177/10731911166666949.
- [4] Prabhu D, Dennis M, Kholghi M, Lu W, Sandhu M, Packer K, et al., Detecting depression-related movement changes in older adults using smart home motion sensors - a feasibility study. *Annu Int Conf IEEE Eng Med Biol Soc*. 2023 Jul;2023:1-5, doi: 10.1109/EMBC40787.2023.10340431.
- [5] Kurebayashi K, Maeda K, Komuro N, Hirai K, Sekiya H, Ichikawa M. Mental-state estimation model with time-series environmental data regarding cognitive function. *Internet of Things (Netherlands)*. 2023 Jul;22:100730, doi: 10.1016/j.iot.2023.100730.
- [6] Jacobson NC, Feng B. Digital phenotyping of generalized anxiety disorder: using artificial intelligence to accurately predict symptom severity using wearable sensors in daily life. *Transl Psychiatry*. 2022 Aug;12(1):336, doi: 10.1038/s41398-022-02038-1.
- [7] Fukuda S, Matsuda Y, Tani Y, Arakawa Y Yasumoto K. Predicting depression and anxiety mood by wrist-worn sleep sensor. In: *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerComWorkshops)/*; 2020 Mar 23-27; Austin TX. *IEEE Xplore c2020*. p. 1-6, doi: 10.1109/PerComWorkshops48775.2020.9156176.
- [8] Mughal F, Raffé W, Stubbs P, Garcia J. Towards depression monitoring and prevention in older populations using smart wearables: quantitative findings. In: *2022 IEEE 10th International Conference on Serious Games and Applications for Health (SeGAH)*; 2022 10-12; Sydney, Australia. *IEEE Xplore; c2022*. p. 1-8, doi: 10.1109/SEGAH54908.2022.9978305.
- [9] Zhang C, Shahriar H. The adoption, issues, and challenges of wearable healthcare technology for the elderly. In: *Proceedings of the 21st Annual Conference on Information Technology Education (SIGITE '20)*; 2020 Oct 7-9; Virtual event. New York (NY): ACM; c 2020. p. 50–53, doi: 10.1145/3368308.3415454.
- [10] Raihan M, Bairagi AK, Rahman S. A machine learning based study to predict depression with monitoring actigraph watch data. In: *Proceedings of the 12th International Conference on Computing Communication and Networking Technologies*; 2021 Jul 6-8; Kharagpur, India. *IEEE Xplore; c2021*. p. 1-5, doi: 10.1109/ICCCNT51525.2021.9579614.
- [11] Rykov Y, Thach T-Q, Bojic I, Christopoulos G, Car J. Digital biomarkers for depression screening with wearable devices: cross-sectional study with machine learning modelling. *JMIR MhHealth Uhealth*. 2021 Oct;9(10):e24872, doi: 10.2196/24872.
- [12] Ghandeharioun A, Fedor S, Sangermano L, Ionescu D, Alpert J, Dale C, Sontag D, Picard R. Objective assessment of depressive symptoms with machine learning and wearable sensors data. In: *Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*; 2017 Oct 23-26; San Antonio, TX. *IEEE Xplore; 2018*. p. 325-332, doi: 10.1109/ACII.2017.8273620.
- [13] Zhang Q, Varnfield M, Higgins L, Smallbon V, Bomke J, O'Dwyer J, et al. The smarter safer homes solution to support older people living in their own homes through enhanced care models: protocol for a stratified randomized controlled trial", *JMIR Res Protoc*. 2022 Jan;11(1):e31970, doi: 10.2196/31970.
- [14] Sheikh JI, Yesavage JA. Geriatric depression scale - short form. *APA PsycTests*. doi: 10.1037/t01786-000.
- [15] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Guyan I, Von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R. In: *31st Conference on Neural Information Processing Systems (NIPS 2017)*; 2017 Dec 4-9; Long Beach, CA. Red Hook (NY): Curran Associates; c2017. p. 4768-77, doi: 10.48550/arXiv.1705.07874.