Assessing the Risk of Falls in Older Adults Living in the Community Using Machine Learning Models with Imputation

Atul Teja Vellampalli[1,2], Yugan Kumar[1,2], Kai Zhe Tan[2], Kok Yang Tan[3,4], Vanessa Koh[3,4], Rahul Malhotra[3,4], Angelique W. M. Chan[3,4], David Matchar[3,4], Navrag B. Singh[2], Sai G.S. Pai[2]

[1]National University of Singapore (NUS)
[2]Singapore-ETH Centre
[3]Duke-NUS Medical School
[4]Centre for Ageing Research & Education
atul.teja@u.nus.edu, yugan.patchandi@sec.ethz.ch, kaizhe.tan@sec.ethz.ch, kokyang.tan@duke-nus.edu.sg, vanessa.kjw@u.duke.nus.edu, rahul.malhotra@duke-nus.edu.sg, angelique.chan@duke-nus.edu.sg, david.matchar@duke-nus.edu.sg, navrag.singh@sec.ethz.ch, sai.pai@sec.ethz.ch

Abstract

Falls are a major cause of injury and loss of independence among older adults, making prevention a critical priority for healthy aging. Early detection of fall risk through screening can enable timely interventions that reduce these adverse outcomes. Traditional clinical methods, such as using the history of falls and simple questionnaire-based screening, provide a quick and low-cost means of assessment but often have poor predictive accuracy and fail in presence of missing information. To support cost-effective screening and intervention, there is a need for tools that can accurately assess fall-risk in presence of missing information with better accuracy than current approaches. In this study, we developed a k-Nearest Neighbors (kNN) model that predicts whether an older adult will experience at least one fall within 12 months after baseline assessment, while simultaneously imputing missing data. Using data from 2,291 community-dwelling older adults in Singapore and 317 features spanning gait, cognition, physical activity, and comorbidities, our model achieved an AUC = 0.62 and F1 = 0.40, a significant improvement over the current clinical standard based solely on fall history (AUC $\approx$ 0.50). This model offers a more cost-effective screening tool for large-scale community deployment and highlights the feasibility of lightweight, imputation-aware models for practical fall-risk screening in aging populations.

Introduction

Mortality rate from falls among U.S. older adults was 78.0 deaths per 100,000 persons in 2021 and has risen since. More than 27% of older adults reported at least one fall in the past year leading to injuries, hospitalizations, reduced mobility, and diminished quality of life. The economic burden is equally significant, with fall-related medical costs in the United States estimated at USD 50 billion annually (Kakara et al. 2023). These rising rates of incidence, mortality, and costs underscore the need for early, personalized assessment of fall risk and tailored interventions.

Traditional methods, such as questionnaires and clinic-based functional tests, are resource-intensive and difficult to scale in community settings (Törnblom et al. 2025). Advances in digital health technologies, such as wearables, have enabled continuous and cost-effective gait monitoring (Lockhart et al. 2021), a key predictor of fall risk, outside clinical environments. Machine learning models leveraging gait, medical history, and lifestyle factors have shown promise for improving prediction accuracy. However, if a few tests are missed then this leads to incomplete feature sets, limiting applicability in real-world settings.

In this study, we developed a k-Nearest Neighbors (kNN) model capable of predicting fall risk while imputing for missing features, thereby addressing practical challenges in community-scale screening.

Methods

With the objective to develop a machine learning model for fall risk prediction, we used data from the Targeted Assessment and Recruitment of Geriatrics for Effective fall prevention Treatment (TARGET) cohort of 2291 older Singaporeans (Lai et al. 2024). Out of the 2291 individuals enrolled in the TARGET cohort, follow-up is still ongoing as of October 2025, and only 659 participants have completed the 12-month follow-up required for generating fall-risk labels. These 659 participants formed the dataset for model development. Each participant was labeled as positive if they reported at least one fall in any of the follow-up calls within the 12-month period, and negative otherwise. Although data collection for the remaining participants is still in progress, interim analyses were undertaken at this stage to derive a preliminary fall-risk score.

A total of 317 features were derived from (a) gait data collected using six inertial measurement unit sensors during a five-minute walk, and (b) questionnaires on cognition, physical activity, comorbidities, and quality of life. Artificial masking of 5% of features in the dataset was used to simulate missing data. Models including kNN, eXtreme Gradient Boosting (XGBoost), tabular foundation models (TabPFN), Random Forests with Multivariate Imputation by Chained Equations (MICE) and Logistic Regression (LR) with MICE were evaluated. A total of 659 participants with follow-up information (labels) were included in model development.

$N_{features}$	**Masking (%)**	**Model family**	**F1**	**AUC**
317	0	kNN	0.39	0.61
317	5	kNN	0.39	0.61
317	5	MICE	0.26	0.48
317	5	TabPFN	0.23	0.47
317	5	XGBoost	0.24	0.45
317	5	LR	0.21	0.52
39	**5**	**kNN**	**0.40**	**0.62**
–	0	FH	0.21	0.52

Table 1: Performance of machine learning models and the fall history (FH) model. The kNN model trained with 5% masked data and 39 features outperforms all other models significantly, including the FH model, which is the current standard for fall risk screening

These models were compared against the baseline fall history model, which is a single self-report question used in standard fall-risk screening as recommended by the World Guidelines for Falls Prevention and Management (Montero-Odasso, van der Velde et al. 2022). Stratified sampling preserved the fall rate (20.48%) across training (60%), validation (20%), and test (20%) splits. Features were normalized and class imbalance was addressed through random undersampling of non-fallers. Block-wise masking was performed using the formulated feature families. See Supplementary Materials – Section A for further details on data collection and model development process. Models were evaluated using a bootstrap-based procedure (see Supplementary Materials for details) from which the best model was picked based on the mean F1-score. Final performance was evaluated on the held-out test set using F1-score, sensitivity, specificity, accuracy, and AUC.

Results

A kNN model, with k=13, trained with 317 features and 5% masking, achieved an F1 score of 0.39 and an AUC of 0.61 and outperformed other models (see Figure 1a) including the use of self-reported fall history for screening. A SHAP analysis of this kNN model (see Figure 1b) helped us identify the 39 most important features (and families). Using these 39 features, we retrained the kNN model (k=13) with 5% imputation, which achieved an F1 score of 0.40, and an AUC of 0.62 and significantly outperforms fall history and other machine learning models ($p < 0.05$) (Demšar 2006).

Discussion

The 5% masking used matches closely with 4.91% missingness in the TARGET dataset collected from community-dwelling older adults. In this scenario, the kNN model with 39 features outperforms other machine learning models as well as the use of history of falls to predict propensity for future falls. This kNN model includes important predictors of fall risk (Koh et al. 2024), identified using a SHAP analysis, and accounts for missing feature values. Although questionnaire features such as prior falls contributed the most to model output, the gait features still added discriminative

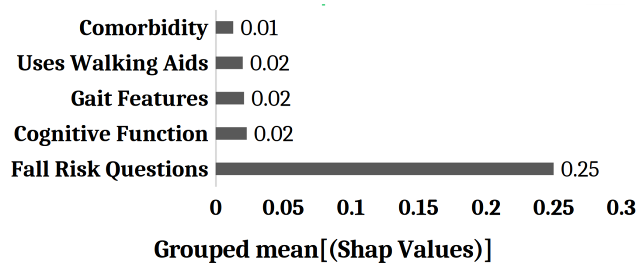

Figure 1: SHAP analysis of the kNN model with 317 features to identify the important feature families that include 39 features used to train the improved model.

power, capturing subtle balance that questionnaires alone may not reflect. The use of machine learning approaches here enabled incorporating information from multiple domains (cognition, gait, fall risk behavior etc.). Current guidelines to incorporate such information generally rely on empirical decision trees (Montero-Odasso, van der Velde et al. 2022).

The kNN model with 5% masking outperforms the same model under no masking condition. This improvement may arise because artificial masking regularizes the model by exposing it to small random perturbations during training, analogous to noise injection or denoising autoencoder behavior, which enhances generalization on incomplete real-world data (Poole et al. 2014). Masking and subsequent imputation improves model performance by improving the quality of distance calculations in kNN as well as reduction in noise through imputation.

The kNN model (with 39 features at 5% imputation) has better sensitivity than the fall history model (0.70 versus 0.19). Higher sensitivity in a screening tool, such as in this kNN model, minimizes false negatives to ensure cost effectiveness. Notably, the k nearest neighbors model achieves an F1 score of 0.40 compared to 0.21 for the fall-history baseline, representing an almost 100 percent relative improvement in balanced accuracy between fallers and non-fallers. Thus, the kNN model developed in this study is a suitable tool for fall-risk assessment in community-centric screening, even in presence of missing values. This model can support recruitment into tailored intervention programs to mitigate fall incidences and adverse outcomes.

Future Work

Although the present analysis was based on an interim dataset with 659 labeled participants, ongoing follow-up in the full cohort will allow validation on a larger sample. Future work will extend this approach to include ensemble models to further improve discrimination.

Acknowledgements

The research was conducted at Future Health Technologies at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. This research is supported by the National Research Foundation Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

References

Demšar, J. 2006. Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, 7: 1–30.

Kakara, R.; et al. 2023. Nonfatal and Fatal Falls Among Adults Aged $\geq$65 Years — United States, 2020–2021. *MMWR Morbidity and Mortality Weekly Report*, 72: 938–943.

Koh, V.; et al. 2024. Performance of Digital Technologies in Assessing Fall Risks Among Older Adults with Cognitive Impairment: A Systematic Review. *Geroscience*, 46: 2951–2975.

Lai, W. X.; et al. 2024. Steps to Avoid Falls in the Elderly – a TECHnology Enhanced Intervention (SAFE-TECH) Study: Randomized Controlled Trial Protocol for a Community-Based, Multi-Component Fall Prevention Program. *BMC Public Health*, 24: 3464.

Lockhart, T. E.; et al. 2021. Prediction of Fall Risk Among Community-Dwelling Older Adults Using a Wearable System. *Scientific Reports*, 11: 20976.

Montero-Odasso, M.; van der Velde, N.; et al. 2022. World Guidelines for Falls Prevention and Management for Older Adults: A Global Initiative. *Age and Ageing*, 51.

Poole, B.; Lahiri, S.; Raghu, M.; Sohl-Dickstein, J.; and Ganguli, S. 2014. Analyzing Noise in Autoencoders and Deep Networks. arXiv:1406.1831.

Törnblom, M.; et al. 2025. Older Persons' Experiences with Wearable Sensor-Based Fall Risk Screening in Free-Living Conditions: A Qualitative Study. *BMC Geriatrics*, 25: 426.