

ORIGINAL ARTICLE

Lung Functions and Self-Reported Physical Activity Level for the Classification of Physical Frailty in Older Adults Using a Machine Learning Approach

Nur Wahidah Iskandar Amer Bukhorri ¹, Siti Ruzita Mahmud ^{1,2*}, Maheza Irna Mohamad Salim ², Devinder Kaur Ajit Singh ³

¹ Physiotherapy Laboratory, Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia

² Advanced Diagnostics and Progressive Human Care (DIAGNOSTICS) Research Group, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia

³ Center for Healthy Ageing and Wellness, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia

ABSTRACT

Introduction: Assessing physical frailty often requires multiple performance-based measures, which may be impractical for older adults with health impairments or in time-constrained settings. Lung function is a key predictor of frailty progression, and physical activity is strongly associated with lung function. However, whether these two factors alone can accurately classify frailty using machine learning (ML) remains unclear. This study evaluates the classification of frailty status using lung function and self-reported physical activity levels, comparing Neural Networks (NN) and Support Vector Machines (SVM). **Methods:** Data from 37 older adults (66.24 ± 4.34 years) were collected completely at community centers. Physical activity was assessed using the International Physical Activity Questionnaire (IPAQ), and lung function was measured via spirometry (Spirobank II, MIR, Italy). The dataset was split into training and testing with a ratio of 70:30. Within the training set, 10-fold cross-validation was applied to optimize model performance before final testing. **Results:** The NN model outperformed SVM, achieving 90.9% testing accuracy with high Area Under the ROC Curve (AUC) scores: frail (1.00), pre-frail (0.99), and robust (1.00). The Linear SVM model showed slightly lower discriminatory power. **Conclusion:** Lung function and physical activity levels effectively predict frailty status. The superior performance of the NN model supports its potential for early frailty detection. Future research should explore larger datasets and additional physiological factors to enhance predictive accuracy and clinical utility.

Malaysian Journal of Medicine and Health Sciences (2025) 21(s2): 108–116. doi:10.47836/mjmhs.21.s2.16

Keywords: Pre-frailty detection, Frailty classification, Machine learning, Neural network, Physical activity level

Corresponding Author:

Siti Ruzita Mahmud, PhD

Email: sruzita@biomedical.utm.my

Tel: +6013-7705502

INTRODUCTION

The global rise in life expectancy has led to an ageing population with Malaysia projected to have a significant elderly demographic by 2040 (1). The ageing process often brings about frailty, a condition marked by reduced resilience and increased vulnerability to health risks (2, 3). Frailty or functional decline is a geriatric syndrome that is dynamic, multidimensional, fluctuates over time, and is common in older persons where poor physical function is shown to be one of the modifiable risk factors for frailty (2, 4) especially in Malaysian older adults (5).

Besides, frailty is also associated with adverse outcomes and increased healthcare costs, making it a pressing public health concern, as Malaysia has the highest prevalence of frailty compared to other Asian countries (6). Thus, early pre-frailty identification in older adults is crucial as it may require preventive management that is different from the rehabilitation of those who are frail (7, 8).

Frailty is a complicated health state characterised by a decrease in biological functions and increased vulnerability to severe health effects. This condition is often seen in people who are at least 55 years old and older (9) and is associated with a reduced ability to cope with stressors, leading to a higher risk of disability, falls, hospitalisation, and mortality. However, frailty is a dynamic process in older adults, whereby the

frailty level has the potential to alter over time. Frailty is influenced by multiple factors, including physical decline characterized by muscle weakness, weight loss, exhaustion, and slow walking speed, as outlined in Fried's Frailty Phenotype (10). Chronic diseases such as diabetes, cardiovascular conditions, COPD, and arthritis further contribute to frailty risk (11) while cognitive frailty, which includes both physical frailty and cognitive impairment, increases the risk of adverse health outcomes (12). Socioeconomic challenges, including lower education levels, restricted healthcare access, and poverty, are associated with higher frailty prevalence (13). Besides, unhealthy lifestyle habits, such as smoking, alcohol consumption, and physical inactivity, significantly contribute to frailty development (4). Various assessment tools are used to determine frailty status, including Fried's Frailty Phenotype (FFP) (10), the Clinical Frailty Scale (CFS) (14), and the Frailty Index (FI) (15). The FI requires at least 30 health variables, making it complex and time consuming to collect. While the CFS is simple and widely used in clinical settings, its subjective assessment relies on clinical judgment. Among these tools, FFP is considered more practical for determining frailty status as it consists of only five criteria. However, its primary focus is on performance-based measures involving the musculoskeletal system that could be difficult to conduct in large populations or clinical settings with limited time (6).

On the other hand, lung function is increasingly recognized as a key predictor of frailty development and progression in older adults (16, 17) and physical activity levels have a protective effect on lung function. However, it remains unclear whether lung function and physical activity alone can predict frailty status as effectively as the composite FFP score. A study used machine learning (ML) models to identify pre-frailty in older adults based on lung function along with more than 30 other variables (18). However, incorporating such a large number of variables requires extensive hours of health assessments, which may be impractical in settings where older adults have health impairments and low participation rates, leading to incomplete data collection. In such cases, missing data is inevitable. Traditional statistical factor analysis typically requires complete datasets or advanced imputation techniques to handle missing values. In contrast, ML are inherently more robust to missing data and can process incomplete datasets without compromising accuracy. Additionally, ML excels in analyzing complex data, identifying non-linear relationships and handling high-dimensional datasets (18, 33), which may not be possible with traditional statistical factor analysis.

Individuals with compromised lung function are susceptible to frailty (17). Although the existing method of frailty classification did consider lung function as one of the variables included in the machine learning models (18), it was measured based on test scores

that did not explicitly describe how these scores were measured. Specifically, lung function can be measured using a digital spirometer, making it easier to integrate into machine learning models. In addition to lung function and physical activity levels, a previous study incorporated other 68 variables for frailty classification (18), requiring extensive hours of health assessment sessions. The large number of health variables involved in identifying pre-frailty in older adults may not be easily adaptable for clinical or community rehabilitation use if the assessment is to be conducted periodically.

Higher lung function, as measured by spirometry, has been associated with increased physical activity levels (4), yet it remains unclear whether these variables alone can effectively classify different frailty levels in older adults. Additionally, previous research attempted to model frailty risk using machine learning based on real-time skeletal movements during functional assessment tests, without incorporating lung function. However, this method has limitations, as movement based data may contain noise due to errors in joint position tracking and the accuracy of data captured using Kinect sensors decreases with distance, further compromising measurement reliability (19). Given that pre-frailty predictors are largely mutable and amenable to population interventions (20), frailty assessments involving numerous performance-based measurements may at times impractical for older adults with certain health impairments in clinical or community rehabilitation settings.

Therefore, this present study investigates whether lung function and self-reported physical activity levels alone are sufficient to classify frailty levels in older adults using a machine learning (ML) approach. While previous studies have shown that Neural Networks (NN) and Support Vector Machines (SVM) achieve classification accuracies exceeding 90%, outperforming other classifiers across various sample sizes (19) this present study also aims to determine the most effective model for early frailty detection by comparing the classification accuracy of NN and SVM based on lung function and physical activity levels.

METHODS

This study employed a supervised machine learning (ML) approach to classify frailty status in older adults based on lung function and physical activity levels. Data collection and model development were conducted in 3 phases, including participant recruitment, data acquisition and preprocessing and ML model training and evaluation.

Participants

A total of 38 older adults (mean age: 66.24 ± 4.34 years) were recruited from Senior Citizen Activity Centers in Johor Bahru and Pasir Gudang, Malaysia. Inclusion criteria required participants to be aged 60 years or

older, have a cognitive score of ≥ 11 on the Intervention for Dementia in Elderly Africans (IDEA) test, and be able to walk independently. Participants with acute illnesses or comorbidities that physically debilitated them were excluded. Ethical approval was obtained from the UTM Medical Research and Ethics Committee (UTMREC-23-22), and informed consent was collected from all participants before they participated in the present study.

Data Collection

Apart from the demographic and health factors of the participants, other key variables that were collected in a field study include the physical activity level, lung function parameters, and frailty as shown in Table 1. The first two variables were the predictors while the last was the target variable. Physical activity levels were assessed using the International Physical Activity Questionnaire (IPAQ)–Short Form, a validated tool for measuring physical activity in adults (9). Participants were asked to recall their physical activities over the past seven days, including time spent in vigorous activities, moderate activities, walking, and sitting. Responses were converted into Metabolic Equivalent of Task (MET) scores, and participants were classified into low, moderate, or high physical activity levels according to IPAQ scoring guidelines.

Lung function was evaluated using a portable digital spirometer (Spirobank II, MIR, Italy), following the American Thoracic Society (ATS) guidelines. It is a handheld device that measures individuals' ability to exhale air from their lungs and provides an indicator of airway ventilation. Each participant performed two spirometry trials, with the best trial selected for analysis to ensure accuracy. The procedure was conducted where participants were seated, fitted with a nose clip, and instructed to form a tight seal around the mouthpiece as demonstrated in Fig. 1. Then, they were asked to take a deep breath in and then exhale forcefully as quickly

and completely as possible. The spirometry measures include the total volume of air forcibly exhaled after a deep inhalation known as Forced Vital capacity (FVC) and the volume of air exhaled in the first second of a

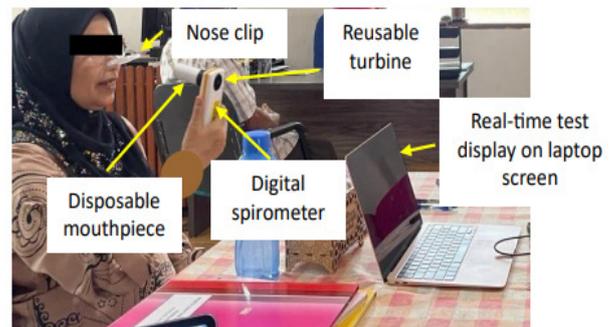


Fig. 1: Spirometry test which was performed by one of the participants forced breath known as Forced Expiratory Volume (FEV_1). A brief rest period was given between trials to prevent fatigue.

Frailty classification was determined using Fried's Frailty Phenotype (FFP), a validated tool for assessing frailty in older adults (10). This assessment consisted of five criteria. Firstly, is the slowness, measured via a gait speed test over a 3m distance. Secondly is the weakness, assessed using a handgrip strength test with a dynamometer. Thirdly is exhaustion, evaluated through self reported fatigue based on the Center for Epidemiologic Studies Depression Scale (CES-D). Fourthly, is unintentional weight loss, defined as a loss of $\geq 5\%$ of body weight within the past year and the last criterion is low physical activity, determined using IPAQ-derived MET scores. Participants meeting three or more criteria were classified as frail, those with one or two criteria were considered pre-frail, and those meeting none were classified as robust (non-frail). The data was then imported into the MATLAB workspace. Next, the Classification Learner App was launched with the previously loaded dataset selected to begin the classification process. Furthermore, the input variables

Table 1: Details of participants characteristics and health assessments features.

Demographics and health factors	Fried's Frailty Phenotype	Lung function (FEV_1 % Predicted)	Physical activity level based on IPAQ (MET)
Age	Slowness based on 3m distance walk (<16.8s)	Normal ($\geq 80\%$)	Low (< 600 MET-minutes/week)
Gender	Weaknesses based on hand grip strength <13.6kg on right and left hands	Mild (70-79%)	Moderate (600 to 2,999 MET-minutes/week)
Marital status		Moderate (60-69%)	
IDEA Cognitive Screen	Exhaustion in the last month	Moderate-Severe (50-59%)	High ($\geq 3,000$ MET-minutes/week)
Hypertension status	Unintentional weight loss >5% within the past year		
High cholesterol status	Activities performed during the last 7-days		
Diabetes mellitus status			
Arthritis status			
Cardiovascular disease status			
Chronic obstructive pulmonary disease status			
Asthma status			
Tuberculosis			

FEV_1 : Forced Expiratory Volume in 1 Second, MET: Metabolic Equivalent of Task
 1 MET = 3.5 mL O_2 /kg/min (oxygen consumption at rest)

were selected, with lung function and physical activity level served as the predictors or independent variables while the frailty level was assigned as the target or dependent variable.

Data Preprocessing

Missing values are a frequent challenge in clinical datasets, often arising from incomplete patient records or inconsistent data collection methods. In this study, 37 out of 38 participants completed data collection. The dataset was split into 70% training and 30% testing. Within the training set, 10-fold cross-validation was applied to optimize model performance before final testing (22). Both Support Vector Machine (SVM) and Neural Network (NN) classifier models were trained using this approach to ensure robust generalization and performance evaluation.

Machine Learning Models Training and Classification

An overview of the ML approach used in this study is illustrated in a flowchart (Fig. 2). Two supervised machine learning (ML) models, Neural Networks (NN) and Support Vector Machines (SVM), were selected due to their high classification accuracy in previous study (18). Rather than employing data-driven feature selection methods, features were manually chosen based on their presumed relevance and interpretability, informed by domain expertise and existing literature.

To classify frailty levels using the MATLAB Classification Learner App, the Support Vector Machine (SVM) model was trained by first selecting an appropriate SVM model type from the available options, including Linear SVM, Quadratic SVM, Cubic SVM, and Gaussian SVM with fine, medium, or coarse settings. Once the model type had been selected, hyperparameters were adjusted to optimise performance, with the regularization parameter

(C) being set to control model complexity and the kernel scale being tuned for Gaussian SVM. After the model was configured, the "Train" button was clicked to fit the SVM model to the dataset. Once training was completed, model performance was evaluated using classification accuracy, the confusion matrix, and the ROC curve to assess its effectiveness.

Meanwhile, to train the Neural Network (NN) model for classifying frailty levels using the MATLAB Classification Learner App, a shallow feedforward neural network was selected from the available options in the app. A NN classifier was then chosen to proceed with the training process. Next, the NN parameters were configured, where the number of neurons per layer was selected (e.g., 10, 50, or 100), the activation function for hidden layers was set to ReLU, and the backpropagation algorithm was chosen for training. The network with 10 neurons per hidden layer represented Narrow NN, 50 neurons per hidden layer represented medium NN while the 100 neurons per hidden layer represented Wide NN. After the parameters had been configured, the "Train" button was clicked to fit the model to the dataset. Once training was completed, the test data was evaluated using the predict() function in MATLAB, where frailty levels were classified based on lung function and physical activity levels. The predicted labels were then compared with the true labels to assess the model's performance. In true labels of frailty, individuals were labelled as Frail (Frailty Positive) if they met three or more of the five FFP criteria. Those who met one or two criteria of FFP were classified as Pre-Frail. Finally, individuals who did not meet any of the criteria were categorized as Non-Frail (Frailty Negative).

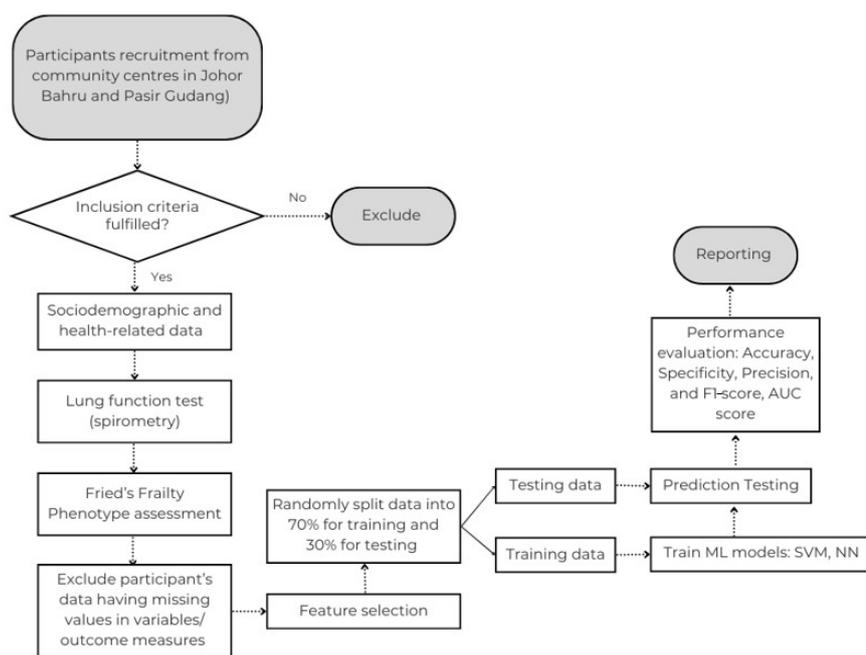


Fig. 2: Flowchart describing the overview of the present study approach.

Evaluation of Models Performance

Model performance was evaluated using the confusion matrix in the MATLAB Classification Learner Toolbox where key performance metrics, including precision, recall (sensitivity), and the F-1 score, were computed as in the following equations:

$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1-score} = (2 \times \text{Recall} \times \text{Precision}) / (\text{Recall} + \text{Precision})$$

TN denotes True Negative, the number of instances correctly classified as negative (e.g., correctly identifying non-frail individuals) while TP denotes True Positive, the number of instances correctly classified as positive (e.g., correctly identifying frail individuals). On the other hand, FN represents False Negative, the number of instances incorrectly classified as negative (e.g., frail individuals wrongly labelled as non-frail) while FP represent False Positive, the number of instances incorrectly classified as positive (e.g., non-frail individuals wrongly labelled as frail).

The confusion matrix was used to assess the performance of the models employed and it is often regarded as the standard for highlighting not only the accuracy of the classification model but also the types of errors made (23). A comparison of model accuracy and performance metrics was conducted in the present study using the summary provided by the app for all trained models. Additionally, Area Under the Curve (AUC) is examined as an additional performance measure because it evaluates the model's ability to distinguish between classes, providing a more comprehensive assessment than accuracy alone. It is especially useful for imbalanced datasets. Higher AUC values indicate better discriminative power, helping to compare models beyond their accuracy scores.

RESULTS

For this study, the input variables were lung function status and physical activity level while the outcome measure was frailty status (FFP) all 9 features were used for FFP classification (frail vs. pre-frail and not-frail). The features consist of 5 categories of lung functions: normal, mild, moderate, moderate-severe and severe lung function with 3 features of low, medium and high physical activity. Frailty classification performed best when using Narrow NN with 90.9% testing accuracy and training accuracy of 92.3% with a 1.4% difference. For the SVM model, the Linear SVM performed the best with also 90.9% testing accuracy even though the training accuracy was 5.3% higher than its training accuracy. A high training accuracy in both classification

models indicates that the model has learned to fit the training data well and can classify the frailty correctly. However, the Narrow NN possessed a lesser difference between the testing and validation accuracy compared to Linear SVM. A comparison of the model's accuracy is presented in Table II.

Table II: Comparison of the accuracy in classification models' performance.

Classification Model	Configurations	Accuracy	
		Validation (%)	Testing (%)
Neural Network (NN)	Narrow	92.3	90.9
	Medium	92.3	90.9
	Wide	92.3	90.9
	Bilayered	92.3	81.8
Support Vector Machine (SVM)	Trilayered	88.5	81.8
	Linear	96.2	90.9
	Quadratic	96.2	90.9
	Cubic	96.2	90.9
	Fine Gaussian	84.6	90.9
	Medium Gaussian	84.6	90.9
	Coarse Gaussian	84.6	81.8
	Efficient Linear	96.2	90.9

The confusion matrix for both NN and SVM models revealed that both classifiers correctly identified most frail and non-frail cases (Fig. 3). However, the classification of pre-frail individuals was slightly lower in accuracy, which may be attributed to the smaller sample size in this category. Performance metrics such as precision, recall, specificity, and F1-score were computed to further evaluate model effectiveness. The detailed performance metrics are summarized in Table III.

Table III: Performance metrics evaluation of SVM and NN models.

Class	Performance metrics				
	Accuracy (%)	Sensitivity/Recall (%)	Specificity (%)	Precision (%)	F1 -score
Frail	92.5	100	77.3	100	1.0
Pre-frail	77.5	75	86	100	0.86
Robust	92.5	100	91.4	100	1.0

Additionally, AUC scores were analyzed to assess the discriminatory power of the models. The NN model demonstrated a near-perfect AUC score (1.00) for frail and robust classifications, whereas the pre-frail class had a slightly lower AUC (0.99), indicating minor misclassification. The SVM model performed comparably but showed slightly lower discrimination for pre-frail cases. The full AUC comparison is presented in Table IV.

When the difference in accuracy between training and

testing is small, the model performs well and is not significantly overfitting (24), where a good balance between fitting the training data and generalizing to unseen data has been achieved. Following Fig. 3, both models projected the same confusion matrix, the value calculation of the precision, specificity, sensitivity, accuracy and F1-score of the class of both models are thus the same as shown in Table III. Both models are similar in generalizing to new, unseen data similarly, indicating that their performance is consistent across different datasets.

Overall, both classifiers demonstrated a good ability to correctly categorize cases. The accuracy metrics showed that the classifiers achieved high accuracy for frail (92.5%) and robust (i.e. not frail) (92.5%) classes, while the accuracy for the pre-frail class was slightly lower (77.5%) and it may probably be due to the subjects

Table IV: Classification performance based on area under the curve (AUC) score between NN and SVM models

Classes	AUC	
	NN	SVM
Frail	1	1
Pre-frail	0.99	0.98
Robust	1	0.98

in the pre-frail class is lesser than the other classes that cause an imbalance dataset. As such, the F1-score which is referred to as the harmonic mean of precision and recall, provides a single metric that balances both concerns. The frail and robust (non-frail) classes both demonstrated perfect precision and recall (F1-score = 1.0) which means there are no false positives and no false negatives and that two classes. However, an F1-score of 0.86 for the pre-frail class indicates strong but not perfect precision and recall (25). It may indicate that the model is performing well but still makes some errors, either by predicting false positives or missing some true positives in identifying pre-frail older persons.

Based on Table IV, the differences in discriminatory power between the NN and SVM classifier models were compared for their area under the curve (AUC) values. A higher AUC signifies better performance in distinguishing between classes, making it a valuable tool for model assessment, especially in imbalanced datasets. From the results shown in Table IV, the NN classifier model seems to perfectly distinguish between frail and non-frail (robust) classes (i.e. better than the SVM) but not for determining the pre-frail class, with AUC equal to 0.99. Although it is almost perfect, a few errors in distinguishing between classes persist with slightly less discriminatory power.

DISCUSSION

This is the first study to use collected data from community-based Senior Citizen Activity Centers to create a frailty case definition using machine learning based on lung function and physical activity levels. The results of this study indicate that machine learning methods (ML) can classify physical frailty when using variables of lung function and self-reported physical activity level. The NN model with 9 features achieved an excellent AUC score for all classes in FFP, except the pre-frail class while SVM only demonstrated an excellent AUC score in the frail class of FFP. The 2 ML classifier models demonstrated similar performance with 77.5% accuracy, 75% sensitivity, 86% specificity, 100% precision and an F-1 score of 0.86 for the pre frail class. These values are higher than those reported in a study that used SVM to classify prefrail with an accuracy of 65.9%, sensitivity of 45.8%, specificity of 81.6%, precision of 58.3% and F-1 score of 0.5 (18).

The accuracy difference could be affected by the sample sizes where the study had more than 600 participants aged 40 to 75 years old (18) while ours had only 37 participants between 62 to 70 years old, a larger sample size is typically expected to improve the performance of machine learning models by providing more data to learn from. However, the accuracy shown in our study for the frailty and robust (non-frailty) classes was 92.5%, aligned with the findings by Vabalas et al. (2019). Their study reported that the accuracy of NN and SVM had

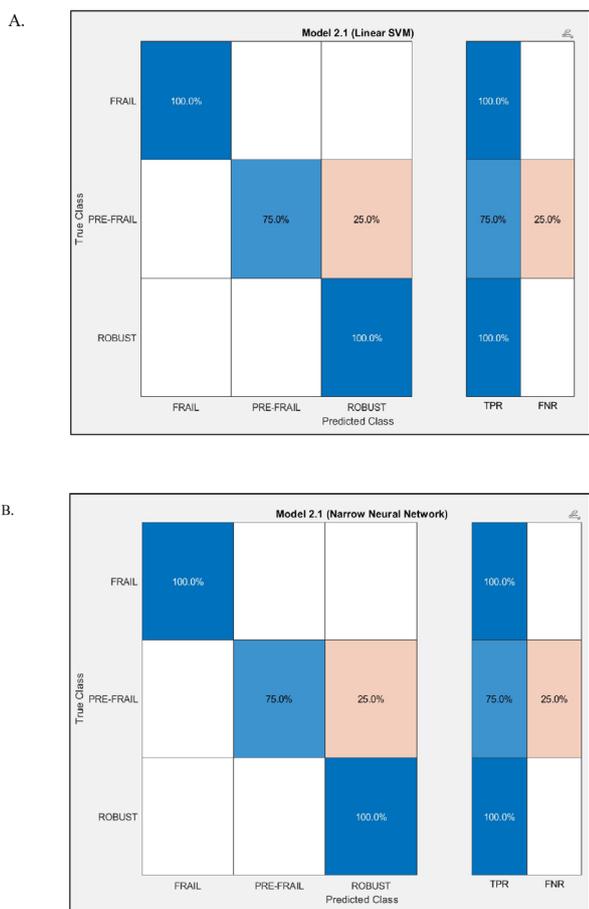


Fig. 3: Confusion Matrix of (a) NN model and (b) SVM model.

more than 90% performance and outperformed all sample sizes and other classifiers (26) with large sample sizes (between 100 to 1000) substantially decreased the accuracy between 60 and 70% (27). The decreasing overall accuracy found in some large sample sizes (26) could be due to irrelevant or redundant features that can dilute the model's ability to learn from relevant features.

Nevertheless, other studies supported that high accuracy of frailty classification can be obtained through large sampled data (28, 29). On the contrary, the lower accuracy observed in pre-frailty than in the other classes is probably due to a challenge in identifying the early manifestations of declining function or pre-frailty in the adult community when there is no widely agreed measurement approach, thus it is mostly detected too late as it already progresses to frailty state (30).

Using the original imbalanced dataset, the NN model in our study showed high predictive accuracy (AUC scores of 0.99) for the pre-frailty classification. Meanwhile, the AUC score of 0.7 for the pre-frail prediction was previously reported (16) despite having 20 input features compared to ours which was limited to 9 features from lung function and physical activity level. It could be due to the right features which enhance the model's performance by focusing on the most relevant information and ignoring noise (31) rather than the large number of irrelevant features. Another difference that was also noticeable is that their study excluded the frail participants due to low numbers that created imbalanced data while ours included frail participants, similar to another study (32) despite having low numbers. There is one study that tried to overcome the imbalanced dataset between classes with synthetic minority over-sampling technique (SMOTE) showed that the balanced dataset created by SMOTE did not result in better performance anyway as compared with the original imbalanced dataset (33). However, lung function was the common variable used in our study and the previous study (18) for the pre-frailty classification. Our study used a 70:30 train and test split ratio with 10-fold cross-validation to evaluate the classifier's validity. Nevertheless, a simulation study showed that K-fold cross validation produced strongly biased performance estimates even with a 1000 sample size or smaller (27).

Furthermore, the area under the receiver-operating characteristic curve (AUC) curve scores suggests that the NN model is marginally better than the SVM model at distinguishing between pre-frail and frail cases. Therefore, the NN model has a slight advantage over the SVM model based on the AUC values when assessing the discriminatory power and capacity to differentiate classes.

CONCLUSION

We were able to create a supervised classification NN model based on lung functions and physical activity levels as input variables for the identification of frailty with a 75% sensitivity and 86% specificity using collected data for usage in the Malaysian context. Lung function and physical activity level variables are sufficient to predict the participants' frailty status accurately thus, can be done with fewer participants' efforts and time than the other functional measures. Between NN and SVM classification models, the NN model was the optimal choice for effectively classifying frailty status into 3 categories: frail, pre-frail and robust with 90.9% testing accuracy and the highest AUC value. However, variables related to anthropomorphic, environmental, social, lifestyle and other physiological factors for frailty classification could be further investigated when using similar ML models.

ACKNOWLEDGEMENTS

This work is supported by the Fundamental Research Grant Scheme (PY/2022/03465) from the Malaysia Ministry of Higher Education and in collaboration with the Department of Welfare, Johor Bahru. This paper has been sent to UTM in partial fulfillment of the thesis requirements for the award of the degree of Bachelor of Engineering (Bio-Medical).

REFERENCES

1. Department of Statistics Malaysia. Population Projection (Revised), Malaysia, 2010–2040, Putrajaya, Malaysia. 2016
2. Lam JYJ, Barras M, Scott IA, Long D, Shafiee Hanjani L, Falconer N. Scoping Review of Studies Evaluating Frailty and Its Association with Medication Harm. *Drugs Aging*. 2022;39(5):333-53.doi: 10.1007/s40266-022-00940-3.
3. Shur NF, Creedon L, Skirrow S, Atherton PJ, MacDonald IA, Lund J, et al. Age-related changes in muscle architecture and metabolism in humans: The likely contribution of physical inactivity to age-related functional decline. *Ageing Research Reviews*. 2021;68:101344.doi:https://doi.org/10.1016/j.arr.2021.101344
4. Luzak A, Karrasch S, Thorand B, Nowak D, Holle R, Peters A, et al. Association of physical activity with lung function in lung-healthy German adults: results from the KORA FF4 study. *BMC Pulm Med*. 2017;17(1):215.doi: 10.1186/s12890-017-0562-8.
5. Badrasawi M, Shahar S, Kaur Ajit Singh D. Risk Factors of Frailty Among Multi-Ethnic Malaysian Older Adults. *International Journal of Gerontology*. 2017;11(3):154-60.doi: https://doi.org/10.1016/j.ijge.2016.07.006
6. Embong J, Amir K, Nawawi A, Razali R, Justine M. Prevalence, Risk Factors and Measures of Frailty in

- Malaysia: A Scoping Review Malaysian Journal of Medicine and Health Sciences. 2021;17(3):307-18.
7. Rodríguez-Mañas L, García-Sánchez I, Hendry A, Bernabei R, Roller Wirnsberger R, Gabrovec B, et al. Key Messages for a Frailty Prevention and Management Policy in Europe from the Advantage Joint Action Consortium. *The Journal of nutrition, health and aging*. 2018;22(8):892-7. doi: <https://doi.org/10.1007/s12603-018-1064-y>
 8. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. *The Lancet*. 2019;394(10206):1365-75. doi:10.1016/S0140-6736(19)31786-6.
 9. Abbas M, Le Bouquin Jeannis R. A Review of Frailty Analysis in Older Adults: From Clinical Tools Towards Fully Automated Preventive Systems. *IRBM*. 2022;44. doi: 10.1016/j.irbm.2022.11.004.
 10. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. *The Journals of Gerontology: Series A*. 2001;56(3):M146-M57. doi: 10.1093/gerona/56.3.M146.
 11. Figueiredo RG, Pinheiro GP, Arata V, Leal MFM, Santana CVN, Tiraboschi TL, et al. Impact of frailty in elderly patients with moderate to severe asthma. *PLoS One*. 2022;17(7):e0270921. doi: 10.1371/journal.pone.0270921.
 12. Qiu Y, Li G, Wang X, Zheng L, Wang C, Wang C, et al. Prevalence of cognitive frailty among community-dwelling older adults: A systematic review and meta-analysis. *Int J Nurs Stud*. 2022;125:104112. doi: 10.1016/j.ijnurstu.2021.104112.
 13. Norazman CW, Adznam SN, Jamaluddin R. Physical Frailty among Urban-Living Community-Dwelling Older Adults in Malaysia. *Int J Environ Res Public Health*. 2020;17(18). doi: 10.3390/ijerph17186549.
 14. Church S, Rogers E, Rockwood K, Theou O. A scoping review of the Clinical Frailty Scale. *BMC Geriatr*. 2020;20(1):393. doi: 10.1186/s12877-020-01801-7.
 15. Broad A, Carter B, McKelvie S, Hewitt J. The Convergent Validity of the electronic Frailty Index (eFI) with the Clinical Frailty Scale (CFS). *Geriatrics (Basel)*. 2020;5(4). doi:10.3390/geriatrics5040088.
 16. Yang X, Cheng C, Ma W, Jia C. Longitudinal association of lung function with frailty among older adults: the English Longitudinal Study of Ageing. *Eur Geriatr Med*. 2023;14(1):173-80. doi: 10.1007/s41999-022-00732-3.
 17. Weng SC, Hsu CY, Wu MF, Lee WH, Lin SY. The Impact of Frailty Status on Pulmonary Function and Mortality in Older Patients with Chronic Obstructive Pulmonary Disease. *J. Nutr. Health Aging*. 2023;27(11):987-95. doi: 10.1007/s12603023-2017-7.
 18. Sajeew S, Champion S, Maeder A, Gordon S. Machine learning models for identifying pre-frailty in community dwelling older adults. *BMC Geriatr*. 2022;22(1):794. doi: 10.1186/s12877-022-03475-9.
 19. Akbari G, Nikkhoo M, Wang L, Chen CPC, Han DS, Lin YH, et al. Frailty Level Classification of the Community Elderly Using Microsoft Kinect Based Skeleton Pose: A Machine Learning Approach. *Sensors*. 2021;21(12). doi:10.3390/s21124017.
 20. Gordon SJ, Baker N, Kidd M, Maeder A, Grimmer KA. Pre-frailty factors in community dwelling 40-75 year olds: Opportunities for Successful ageing. *BMC Geriatr*. 2020;20(1):96. doi:10.1186/s12877020-1490-7.
 21. Hagstrumer M, Oja P, Sjustrum M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutrition*. 2006;9(6):755-62. doi: 10.1079/PHN2005898
 22. Kaliappan J, Bagepalli AR, Almal S, Mishra R, Hu YC, Srinivasan K. Impact of Cross Validation on Machine Learning Models for Early Detection of Intrauterine Fetal Demise. *Diagnostics (Basel)*. 2023;13(10). doi: 10.3390/diagnostics13101692.
 23. Swaminathan S, Tantri BR. Confusion Matrix-Based Performance Evaluation Metrics. *African Journal of Biomedical Research*. 2024;27:4023-31. doi: 10.53555/AJBR.v27i4S.4345.
 24. Montesinos Lypez OA, Montesinos Lypez A, Crossa J. Overfitting, Model Tuning, and Evaluation of Prediction Performance. In: Montesinos Lypez OA, Montesinos Lypez A, Crossa J. *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Cham: Springer International Publishing; 2022. p. 109-39.
 25. Weng WH, Wagholikar KB, McCray AT, Szolovits P, Chueh HC. Medical subdomain classification of clinical notes using a machine learning-based natural language processing approach. *BMC Medical Informatics and Decision Making*. 2017;17(1):155. doi: 10.1186/s12911-017-0556-8.
 26. Rajput D, Wang W-J, Chen C-C, X X, Y Y. Evaluation decided sample size in machine learning applications. *BMC Bioinformatics*. 2023;24(1):48.
 27. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PLoS One*. 2019;14(11):e0224365. doi:10.1371/journal.pone.0224365.
 28. Anzai E, Ren D, Cazenille L, Aubert-Kato N, Tripette J, Ohta Y. Random forest algorithms to classify frailty and falling history in seniors using plantar pressure measurement insoles: a large scale feasibility study. *BMC Geriatrics*. 2022;22(1):746. doi: 10.1186/s12877-02203425-5.
 29. Bohn L, Drouin SM, McFall GP, Rolfson DB, Andrew MK, Dixon RA. Machine learning analyses identify multi-modal frailty factors that selectively discriminate four cohorts in the Alzheimer's disease spectrum: a COMPASS-ND study. *BMC Geriatrics*. 2023;23(1):837. doi:10.1186/s12877-023-045461.
 30. Beaton K, McEvoy C, Grimmer K. Identifying

indicators of early functional decline in community-dwelling older people: A review. *Geriatrics & Gerontology International*. 2015;15(2):133-40. doi:10.1111/ggi.12379

31. Chen RC, Dewi C, Huang SW, Caraka RE. Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*. 2020;7(1):52.doi:10.1186/s40537020-00327-4.
32. Le Pogam M-A, Seematter-Bagnoud L, Niemi T, Assouline D, Gross N, Trachsel B, et al. Development and validation of a knowledge based score to predict Fried's frailty phenotype across multiple settings using one-year hospital discharge data: The electronic frailty score. *eClinical Medicine*. 2022;44.doi: 10.1016/j.eclinm.2021.101260.
33. Aponte-Hao S, Wong ST, Thandi M, Ronksley P, McBrien K, Lee J, et al. Machine learning for identification of frailty in Canadian primary care practices. *Int J Popul Data Sci*. 2021;6(1):1650.doi: 10.23889/ijpds.v6i1.1650.