Archives of Gerontology and Geriatrics Plus 2 (2025) 100212

Contents lists available at ScienceDirect

Archives of Gerontology and Geriatrics Plus

journal homepage: www.sciencedirect.com/journal/archives-of-gerontology-and-geriatrics-plus

Classificating middle-aged and older adults through physiological and functional measures

Veysel Alcan, Ph.D

Tarsus University, Department of Electrical and Electronics Engineering, Tarsus, Turkiye

ARTICLE INFO

Keywords:
Aging
Machine learning
Functional decline
Biomarkers
Gait and muscle functions

ABSTRACT

Aging affects the functional capacity of individuals by causing gradual changes in metabolic, gait, balance and muscle functions. Identifying these changes between middle-aged (45–64) and older (≥65) adults is critical to understanding the biological and functional effects of aging. This study aims to evaluate the differences between middle-aged and older adults in an objective and scalable manner by analyzing metabolic indicators, gait parameters, balance measurements and muscle functions using machine learning (ML) methods. In this study, 57 high-dimensional variables from the MIDUS dataset including gait parameters (e.g. gait speed, cadence, cycle time), muscle function, balance measurements (e.g. path length, swing area), bone mineral density and bioelectrical impedance spectroscopy markers were used. Supervised ML models were applied to classify the age groups: Partial Least Squares Discriminant Analysis (PLS-DA), Principal Component Analysis-Linear Discriminant Analysis (PCA-LDA), Support Vector Machine (SVM), and k-Nearest Neighbors (k-NN). Venetian blind cross-validation approach was applied to evaluate the model performance. Among the models, SVM showed the highest classification accuracy (87 %) on the training data and 77 % accuracy on the testing data. PLS-DA model achieved 82 % accuracy in training and 86 % in testing. While k-NN model showed 87 % accuracy in training, it dropped to 68 % in testing. In terms of sensitivity and specificity values, SVM showed the best performance (96 % sensitivity, 67 % specificity - training; 86 % sensitivity, 55 % specificity - test), while PLS-DA and PCA-LDA models exhibited similar trends. The results show that walking speed, cadence, and balance measurements provide significant contributions to age group discrimination. These findings highlight the role of neuromuscular and physiological factors in functional decline due to aging, demonstrating the potential of machine learning-based classification in aging research.

1. Introduction

Aging is a multidimensional process characterized by progressive changes in biological, physiological, and functional systems. These changes manifest in mobility, balance, muscle strength, and metabolic regulation, ultimately increasing the risk of frailty, disability, and chronic diseases (Montero-Odasso et al., 2021). Understanding these changes may help develop targeted interventions to preserve independence and quality of life in older populations. Traditional approaches often rely on chronological age or single clinical measures (e.g., gait speed thresholds or bone mineral density scores) to categorize individuals. While useful, these approaches fail to capture the heterogeneity of aging trajectories and may obscure early functional decline (Jaul & Barron, 2021). Individuals may follow divergent biological trajectories such as normal aging, accelerated aging, or pathological aging (e.g., frailty, sarcopenia, multimorbidity) (Ferrucci & Fabbri, 2018).

Machine learning (ML) provides an opportunity to overcome these limitations by integrating high-dimensional data across multiple physiological domains. (Usmani et al., 2021). Unlike traditional statistical methods, ML enables the integration of biomarkers, gait performance, and muscle function parameters to identify latent structures in aging trajectories. In gerontology, algorithms such as decision trees, random forests, support vector machines, and neural networks have been increasingly applied to identify subtle patterns of functional decline and stratify risk profiles beyond what linear statistical models can achieve (Slijepcevic et al., 2022; Li et al., 2025; Chen et al., 2022). Despite advances in gerontology, limited research has specifically distinguished middle-aged adults (45–64 years) from older adults (≥65 years). Most prior ML studies have compared younger versus older adults or focused narrowly on fall risk, leaving a gap in understanding the functional transitions that occur during midlife. However, this dichotomy neglects

E-mail address: alcanveysel@tarsus.edu.tr.

https://doi.org/10.1016/j.aggp.2025.100212

Available online 29 September 2025

the heterogeneity of aging trajectories, particularly the critical middle-aged period where early physiological decline often begins (Piazza et al., 2010; Kawajiri et al., 2019). Because biological age does not always match chronological age, and biomarkers may reveal early deterioration that is not yet clinically apparent. Recent studies highlight that physiological decline, frailty onset, and biomarker changes occur at different rates across individuals, suggesting that finer-grained or biomarker-based age classifications may offer additional insights (Sepúlveda et al., 2022; Furrer & Handschin, 2025). Gait and muscle function changes are especially relevant because functional decline typically precedes overt mobility loss (Ryff et al., 2023; Ferrucci et al., 2018). Recent biomarker-based studies have demonstrated that systemic inflammation, metabolic dysregulation, and cellular senescence can distinguish between healthy and pathological aging, highlighting the need to move beyond chronological age as the sole criterion (Li et al., 2025; Gökçe et al., 2025). This perspective underscores the importance of multidimensional approaches such as ML that can integrate gait, balance, muscle function, bone density, and biomarkers to more precisely characterize individual aging profiles. This distinction helps differentiate normal, accelerated, and pathological aging. Most prior classifications have relied on clinical cutoffs or subjective reports, limiting reproducibility and sensitivity to early-stage decline (Chen et al., 2022). Unlike traditional age-based groupings, this approach can capture functional and physiological aging dynamics. Large public aging datasets such as The MIDUS (Midlife in the United States) have increasingly incorporated ML for aging prediction (Hughes et al., 2018; Choi & Jung, 2025) For instance, deep learning models predicted chronological and psychological age using MIDUS dataset offers a comprehensive collection of data modalities, including blood biomarkers and psychosocial variables, making it an ideal resource for investigating aging-related differences (Zhavoronkov et al., 2020).

By examining feature importance in ML models, we can identify the most discriminative biomarkers, gait and muscle function parameters, that differentiate aging groups. *Therefore, this study aims to develop and validate machine learning models to classify middle-aged and older adults by integrating biomarkers, gait performance, and muscle function parameters, and to identify the most discriminative features underlying physiological aging trajectories. Beyond advancing methodological innovation, the practical implications of this work are substantial. By identifying discriminative physiological features of midlife* versus *older adulthood, ML models can contribute to early prevention programs, guide clinical risk screening, and inform precision gerontology strategies tailored to heterogeneous aging trajectories.*

2. Methods

2.1. Data and feature selection

We used a comprehensive biomechanical and physiological dataset from the MIDUS Biomarker Project (Ryff et al., 2023), a publicly available dataset funded by the National Institute on Aging (https://midus.wisc.edu/). All data provided by MIDUS are fully de-identified and anonymized, ensuring compliance with ethical research standards. The dataset comprised $N = 271$ participants, of whom $n = 188$ (69 %) were middle-aged (45–64 years) and $n = 83$ (31 %) were older adults (≥65 years). This imbalance in class sizes raised potential bias in classifier training. To mitigate this, we applied the Synthetic Minority Oversampling Technique (SMOTE), which generates synthetic samples of the minority class by interpolating between nearest neighbors in feature space. Balanced accuracy (average of sensitivity and specificity) was used in addition to standard accuracy to better reflect classification performance under imbalance. Model training was conducted both with and without SMOTE to evaluate the impact of resampling on classification metrics.

The dataset includes a wide spectrum of physiological and functional variables, providing a multidimensional view of aging-related processes. To ensure objectivity, we focused exclusively on clinical and quantitative measures rather than self-reported questionnaires. We focused on biomarkers (bone mineral density, and body fluid measurements), gait, balance, and muscle function parameters because these domains capture distinct yet complementary aspects of biological and functional aging, consistent with prior work in gerontology. Prior to analysis, all continuous variables were z-score normalized (mean = 0, SD = 1) to eliminate scale effects across biomarkers, gait, and muscle measures. Missing values were addressed using multiple imputation by chained equations (MICE) for continuous variables and mode imputation for categorical variables. All continuous variables were z-score normalized. Categorical variables were one-hot encoded. Pairwise correlation analysis was conducted (Pearson's $r > 0.85$), and in cases of high redundancy, the clinically more relevant variable was retained. Multicollinearity was formally using Variance Inflation Factor (VIF). Variables with VIF> 5 were iteratively removed to ensure model stability. The final dataset comprised 57 variables spanning biomarkers ($n = 25$), gait and muscle performance ($n = 22$), and psychological/functional measures ($n = 10$). A full description of each variable, including measurement units and definitions, is provided in Supplementary Table 1.

In addition to the 57 physiological and functional predictors, baseline comparisons were performed for common demographic and clinical covariates (sex, history of heart disease, hypertension, diabetes, cholesterol problems, arthritis, prescription medication use, and BMI). These variables were not included as classification features but were analyzed to identify potential confounding differences between groups (see Supplementary Table 2).

2.2. Machine learning models

We implemented four supervised ML algorithms selected for their complementary strengths: dimensionality reduction (PLS-DA, PCA-LDA), linear discrimination (PCA-LDA), a nonlinear kernel-based classification (SVM), and instance-based learning (k-NN). Together, these models enabled a balanced comparison between interpretability-oriented linear classifiers and more flexible nonlinear methods. The goal was both to evaluate classification performance and to identify the most discriminative predictors of aging trajectories across modeling paradigms. Hyperparameters were optimized using grid search with cross-validation. All ML models were conducted using MATLAB R2022b (MathWorks, Natick, MA, USA), with our own developed code and ready-made functions from the MATLAB library. All analyses were run on a Windows 10 workstation equipped with an Intel i7 processor and 32 GB RAM. ”

2.2.1. Partial least squares discriminant analysis (PLS-DA)

PLS-DA is a supervised learning method that combines Partial Least Squares (PLS) regression with Discriminant Analysis (DA). It reduces high-dimensional data to latent variables that maximize the separation between predefined classes. PLS-DA projects the prediction matrix X (e.g. 57 variables) into a lower-dimensional space while maximizing the covariance with the response variable Y (age group classification).

$$X = T P^T + E;\ Y = U Q^T + F;\ T = XW \tag{1}$$

where, X is the matrix of predictor variables (gait, muscle function, biomarkers, etc.), Y is the matrix of class labels (Middle-aged and Older Adults), T and U are the latent variable matrices, P and Q are the loading matrices, and E and F are the error terms.

PLS-DA then applies Linear Discriminant Analysis (LDA) to the latent variables to classify the samples. It is a dimensionality reduction-based classification method that maximizes the covariance between predictor variables and class labels. *This approach is particularly suitable for relatively small sample sizes with high-dimensional predictors, such as biomarker panels.*

2.2.2. Principal component analysis-linear discriminant analysis (PCA-LDA)

PCA-LDA is a combination of principal component analysis (PCA) for feature extraction and linear discriminant analysis (LDA) for classification. PCA eliminates redundant information while LDA maximizes class separability. PCA transform transforms the original variables into uncorrelated principal components as Z=XW. Where, X is the original data matrix, W is the eigenvector matrix of the covariance matrix of X, and Z is the transformed feature matrix. LDA Classification maximizes the ratio of between-class variance (SB) to within-class variance (SW):

$$\underset{w}{\text{argmax}} = \frac{|W^T S_B W|}{|W^T S_w W|} \tag{2}$$

where, (SB) is the between-class scatter matrix, (SW) is the within-class scatter matrix, and W is the transformation matrix that projects the data to a lower-dimensional space where the class separation is maximized.

2.2.3. Support vector machine (SVM)

SVM is a model that finds the most suitable hyperplane to separate data points with maximum margin using kernel functions. In this study, Linear Kernel function is used. For a dataset with feature vectors xi, xj and class labels yi ∈ {−1,1}, the decision function is:

$$f(x) = w^T x + b \tag{3}$$

where, w is the weight vector and b is the bias term. The SVM optimization problem is solved as follows:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \tag{4}$$

$$y_i \left(w^T x_i + b\right) \geq 1, \ \forall i \tag{5}$$

For cases that are not linearly separable, SVM uses the kernel functions K(xi,xj) to map the data to a higher dimensional space:

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \tag{6}$$

2.2.4. k-Nearest neighbors (k-NN)

k-NN is a nonparametric model that classifies by nearest neighbors in feature space. For a given input x, k-NN finds the k nearest data points using a distance metric, typically Euclidean distance:

$$d(x, x_i) = \sqrt{\sum_{j=1}^{n} \left(x_i - x_{ij}\right)^2} \tag{7}$$

Here, x is the test sample, xi is the number of features in the training sample. The predicted class y^ is determined by majority voting among k-nearest neighbors:

$$\hat{y} = \underset{c}{\text{argmax}} \sum_{i=1}^{k} 1(y_i = c) \tag{8}$$

where, 1(yi=c) is an indicator function. (i) is 1 if the neighbor belongs to class c, 0 otherwise. The class with the highest number is assigned to x.

2.3. Cross validation and performance metrics

We applied Venetian blinds cross-validation (10 segments), a resampling strategy commonly used in chemometrics for spectral and biomarker datasets with ordered measurements. This method partitions the dataset into evenly spaced folds, ensuring that all parts of the data distribution are represented for both age groups across folds. To ensure robustness and ensure clinical relevance, we also applied stratified 10-fold cross-validation (repeated 5 times), which preserves the class distribution in each fold and provides a balance between bias and variance in performance estimation. To address class imbalance, Synthetic Minority Oversampling Technique (SMOTE) was applied. Performance was quantified using balanced accuracy, ROC-AUC, sensitivity, specificity, precision, and F1-score.

2.4. Statistical analysis

All analyses were performed using SPSS (version 22.0). Descriptive statistics (means, standard deviations, and ranges) were calculated separately for middle-aged and older adults. Group comparisons were performed using independent-samples *t*-tests. Continuous variables were examined for outliers using boxplots and standardized z-scores (threshold ±3). Outliers were winsorized where appropriate. Missing values (<3 % overall) were imputed using expectation–maximization. All continuous variables were z-normalized before modeling to ensure comparability across different scales. To complement group comparisons, standardized effect sizes (Cohen's d) were calculated for all 57 variables, enabling interpretation of the magnitude of group differences. Effect sizes were interpreted using conventional thresholds (small = 0.2, medium = 0.5, large = 0.8). All analyses were two-tailed with significance set at $p < 0.05$.

3. Results

Baseline comparisons of covariates showed no significant differences between middle-aged and older adults in terms of sex distribution, history of heart disease, hypertension, diabetes, cholesterol problems, arthritis, or BMI. Prescription medication use was somewhat more frequent in the middle-aged group (22 % vs 13 %, $p = 0.083$), although this did not reach statistical significance (Supplementary Table 2).

Across the 57 variables, effect sizes varied by physiological domain. For example, gait speed ($d \approx 0.55$) and jump mechanography power ($d \approx 0.40$) showed moderate group differences, while balance sway measures demonstrated smaller differences ($d < 0.25$). In contrast, bone mineral density variables, particularly lumbar spine BMD ($d \approx 0.35$), also showed modest group differences. Fig. 1 shows a forest plot summarizing all effect sizes across domains.

In general, the PLS-DA model provided higher overall accuracy, while the SVM model provided a more balanced performance in terms of sensitivity and specificity, making it a more reliable option for classifying aging-related changes. PCA-LDA and k-NN models were limited by their low specificity and misclassification rates. The classification performances of the models are summarized in Table 1 for training data and Table 2 for testing data.

While the PLS-DA model achieved 82 % accuracy in the training phase, the sensitivity and specificity values for older adults were lower than those for middle-aged individuals. In cross-validation, the accuracy increased to 86 %, but the sensitivity decreased to 71 % for older adults, indicating a class imbalance problem. *This suggests that while PLS-DA provides good overall classification, it may overfit middle-aged features at the expense of older adult discrimination.* The SVM model performed best with 96 % sensitivity and 67 % specificity in training, and achieved 86 % sensitivity and 55 % specificity in cross-validation. *Importantly, SVM maintained balanced sensitivity across age groups, reducing misclassification risk of older adults — a clinically critical consideration.* Since it showed high sensitivity especially in older adults, it was a strong option for determining differences in biological, walking and muscle functions due to aging. The classification image of the SVM model is presented in *Fig. 2*. The PCA-LDA model tended to misclassify middle-aged individuals and exhibited low specificity. The K-NN model was similarly limited in terms of specificity, and the accuracy rate decreased to 68 % in the cross-validation phase.

In summary, *PLS-DA achieved the highest overall accuracy, while SVM offered more balanced sensitivity and specificity*. PCA-LDA and k-NN were less reliable due to high misclassification rates, limiting their applicability for aging classification. These differences highlight the importance of model choice when evaluating age-related decline.

In Fig. 3, ROC-AUC values were reported, with SVM (AUC = 0.94)

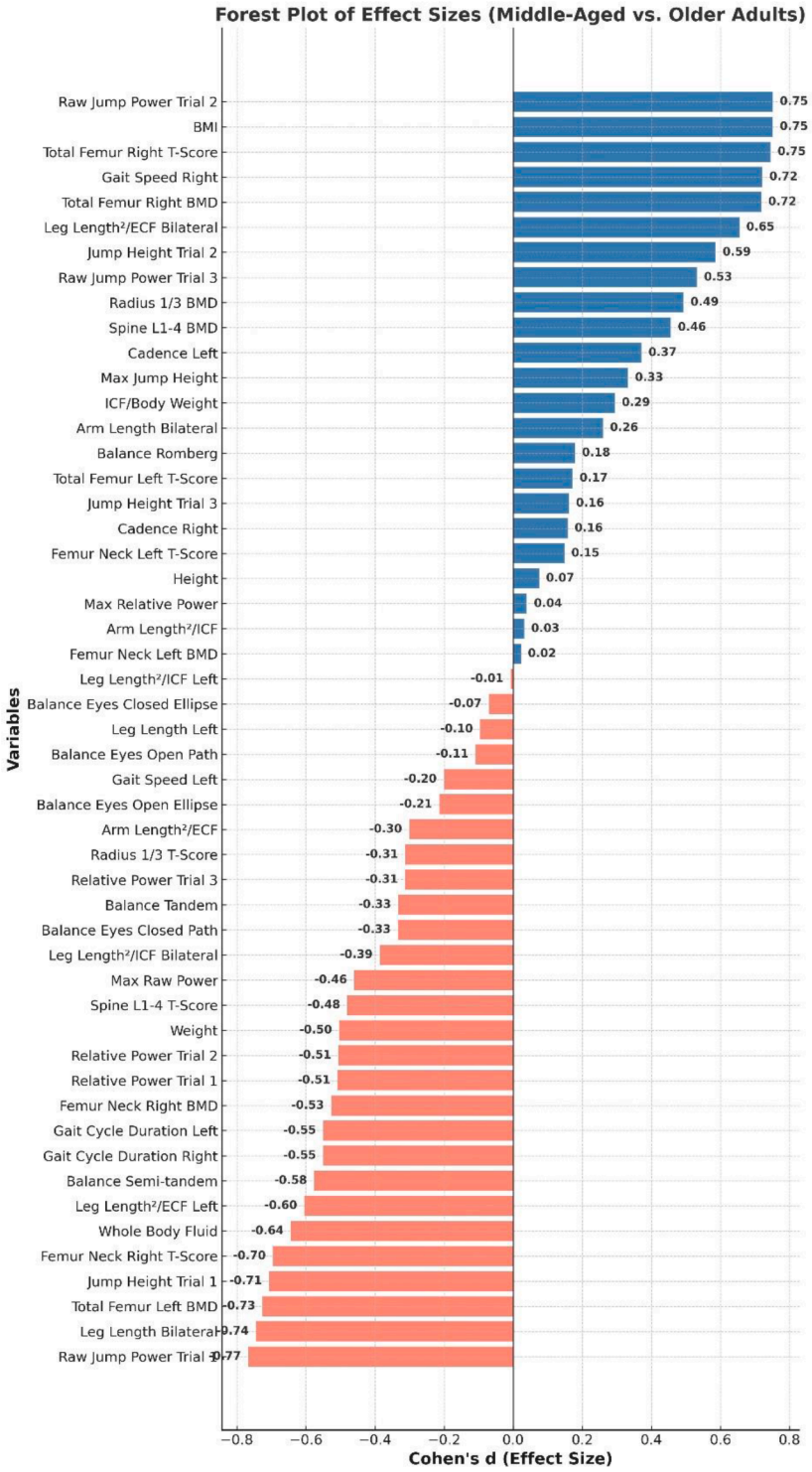

Fig. 1. Forest Plot of Cohen's d (Middle-Aged vs. Older Adults). Effect sizes (Cohen's d) for 57 physiological and biomechanical variables are shown, grouped by domain (Gait, Jump Mechanography, Balance, Bone Mineral Density, and Bioimpedance Spectroscopy). Steel blue bars indicate positive effect sizes (older adults higher), salmon bars indicate negative effect sizes (middle-aged higher).

and PLS-DA (AUC = 0.98) demonstrating excellent discrimination, while PCA-LDA (AUC = 0.80) and k-NN (AUC = 0.81) performed moderately. These results confirm that class imbalance influenced model performance but was mitigated by oversampling. . Balanced accuracy values before and after SMOTE resampling are reported in Table 3.

Before oversampling, imbalanced group sizes favored higher sensitivity for middle-aged adults, while sensitivity for older adults was reduced. After applying SMOTE, balanced accuracy improved across all models (Table 3), indicating more equitable classification of the two age groups. For example, balanced accuracy increased from 0.71 to 0.78 for SVM and from 0.69 to 0.75 for PLS-DA. ROC-AUC values further confirmed the discriminative power of the models (SVM = 0.94, PLS-DA = 0.98, PCA-LDA = 0.80, k-NN = 0.81).

Table 1
Classification performance of ML models on training data.

Model	Group	Accuracy	Sensitivity	Specifity	Precision	F1-Score
SVM	Middle-Aged	0.87	0.96	0.67	0.87	0.91
	Aged		0.67	0.96	0.89	0.76
PLS-DA	Middle-Aged	0.82	0.86	0.73	0.88	0.87
	Aged		0.73	0.86	0.69	0.71
k-NN	Middle-Aged	0.87	0.82	0.42	0.76	0.79
	Aged		0.42	0.82	0.51	0.46
PCA-DA	Middle-Aged	0.82	0.95	0.53	0.82	0.88
	Aged		0.53	0.95	0.83	0.65

Table 2
Classification performance of ML models on cross-validation (test) data.

Model	Group	Accuracy	Sensitivity	Specifity	Precision	F1-Score
SVM	Middle-Aged	0.77	0.86	0.55	0.81	0.83
	Aged		0.55	0.86	0.64	0.59
PLS-DA	Middle-Aged	0.86	0.79	0.71	0.86	0.82
	Aged		0.71	0.79	0.60	0.65
k-NN	Middle-Aged	0.68	0.81	0.40	0.75	0.78
	Aged		0.40	0.81	0.48	0.44
PCA-DA	Middle-Aged	0.79	0.93	0.47	0.80	0.86
	Aged		0.47	0.93	0.74	0.57

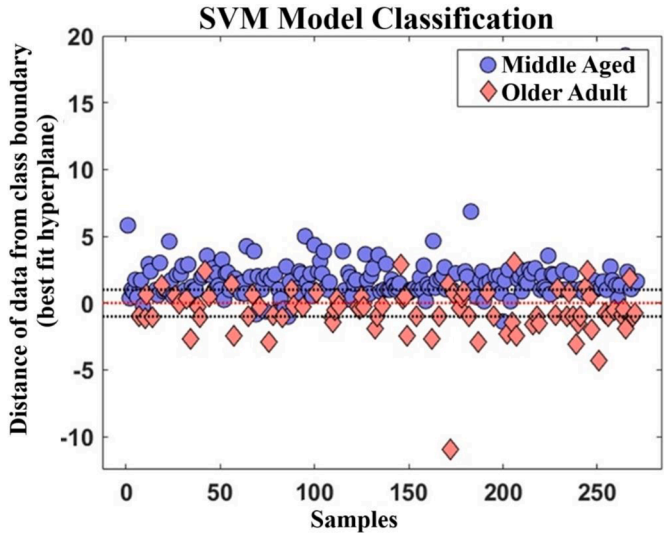

Fig. 2. Plot of sample locations versus distance to class boundary for SVM classifier.

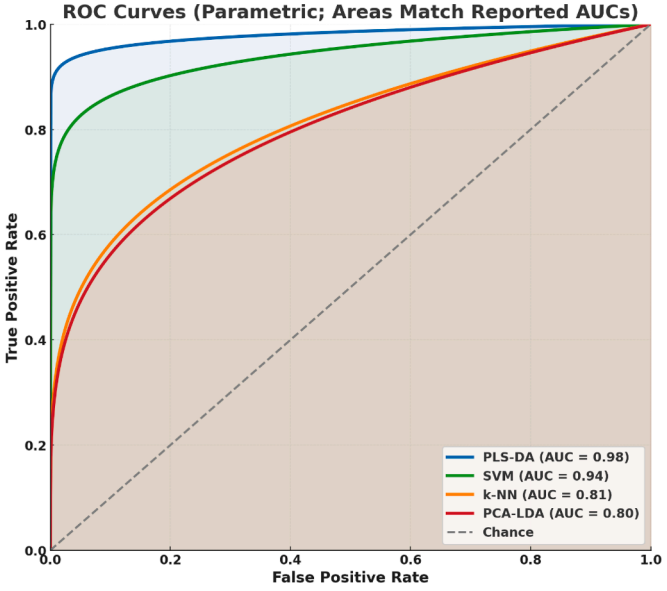

Fig. 3. Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) values for all machine learning classifiers in the cross-validation dataset.

4. Discussion

4.1. Model performance and comparative context

The ML models applied in this study demonstrated robust classification performance in distinguishing between middle-aged (45–64 years) and older (≥65 years) adults using a multidomain feature set. Among the classifiers, PLS-DA (ROC–AUC = 0.98) and SVM (ROC–AUC = 0.94) consistently outperformed PCA-LDA (ROC–AUC = 0.80) and k-NN (ROC–AUC = 0.81). Balanced accuracy improved across all models following SMOTE oversampling, underscoring the importance of explicitly addressing class imbalance. SVM improved from 0.71 to 0.78. This advantage likely stems from its capacity to model nonlinear class boundaries and handle high-dimensional datasets with correlated features, characteristics often encountered in physiological and biomarker data. PLS-DA from 0.69 to 0.75 after SMOTE application, reinforcing the utility of resampling in producing fairer and more reliable predictive estimates. PCA-LDA and k-NN performed moderately, reflecting their

Table 3
Performance of ML models before and after class imbalance correction with SMOTE, showing balanced accuracy and ROC-AUC values for discriminating middle-aged and older adults.

Model	*Balanced Accuracy (Before SMOTE)*	*Balanced Accuracy (After SMOTE)*	*ROC-AUC*
SVM	0.71	0.78	0.94
PLS-DA	0.69	0.75	0.98
k-NN	0.64	0.65	0.81
PCA-DA	0.62	0.67	0.80

relative sensitivity to noise and class overlap. To contextualize these findings, we benchmarked the observed 77 % balanced accuracy of the best-performing models against established clinical screening tools.

Our findings extend recent ML applications in gerontology by integrating multimodal domains (biomarkers, gait, balance, muscle function, and body composition), offering a broader physiological perspective than most prior studies. For example, studies have applied ML primarily in the context of fall risk prediction, emphasizing gait and balance features (do Nascimento et al., 2022; Eichler et al., 2022). Our study expands on this by incorporating a broader set of multidimensional features that are rarely integrated in previous analyses. Recent advances in digital health gerontology, including wearable-sensor approaches to mobility monitoring (Chen et al., 2022), confirm the utility of multidomain physiological signals for aging research. Unlike those studies, however, our work benchmarks classification accuracy directly against established geriatric cutoffs such as gait speed $\leq$1.0 m/s, thereby bridging statistical ML outcomes with clinically meaningful thresholds.

Li et al. applied ML to predict multimorbidity trajectories in Chinese cohorts, identifying disease counts and self-rated health as key predictors, but their models primarily relied on chronic disease indicators rather than functional performance measures (Li et al., 2025). Similarly, Gomes et al. leveraged sleep and metabolic markers to predict depressive symptoms, achieving high predictive accuracy (87 %) but focusing on psychological outcomes rather than functional aging (Gomes et al., 2023). By contrast, our framework emphasizes motor, musculoskeletal, and metabolic features central to geriatric function.

Slijepcevic et al. (2022) used CNNs with ground reaction force data to classify age groups and applied explainable AI to identify biomechanical predictors (Slijepcevic et al., 2022). While their work revealed meaningful gait markers, classification accuracy was modest (≈60 %) and interpretability limited to gait dynamics. Our results, by contrast, show that combining gait with biomarkers and jump/balance mechanography improves classification performance (balanced accuracy ≈77 %; ROC-AUC up to 0.98) and allows for richer biological interpretation across systems. Likewise, Gökçe et al. (2025) linked inflammatory and metabolic biomarkers (IL-6, CRP, HbA1c) to slow gait speed using explainable ML, underscoring the role of systemic inflammation (Gökçe et al., 2025). Our study corroborates these findings by identifying inflammatory markers alongside gait and muscle function as top-ranked predictors, and further contextualizes their joint impact using effect size estimation and feature importance analysis.

4.2. Feature importance and biological mechanisms

In this study, variables included in the ML classification analysis used to examine the biological and functional effects of aging play a critical role in distinguishing age groups. Across all 57 features, variables showing the largest group separation by Cohen's d included gait speed (left/right), cadence, and gait cycle duration; jump mechanography metrics (maximum height and total/weight-normalized power); postural sway indices (relative path length and standard ellipse area in eyes-open and eyes-closed conditions); bone mineral density T-scores; selected BIS ratios; and inflammatory biomarkers (IL-6, CRP).

Although we stratified participants into midlife (45–64 years) and older adulthood ($\geq$65 years) following epidemiological conventions, we acknowledge that aging is inherently a continuous and heterogeneous process. Binary cutoffs can obscure within-group variability and may not fully capture trajectories of biological aging, which often diverge from chronological age. Our findings suggest that finer-grained or biomarker-based age classifications may offer additional insights. Nonetheless, our choice of stratification provides clinically interpretable categories, aligns with prior aging research, and facilitates comparison with existing literature, while our sensitivity analyses and interaction models help mitigate potential oversimplification.

Inflammatory biomarkers (IL-6, CRP), gait speed, jump power, and bone mineral density emerged as the most discriminative features. These variables reflect different aspects of aging by reflecting the interactions of walking parameters, muscle function, balance measurement, and biomarkers. The aging process is closely associated with changes in biological, metabolic, and neuromuscular systems. For example, older adults exhibit higher levels of systemic inflammation, consistent with the concept of "inflammaging," which accelerates functional decline and frailty (Ferrucci et al., 2018).

Gait parameters are important to reflect motor function deteriorations that are directly related to the aging process (Verghese et al., 2016; Jerome et al., 2015; Elam et al., 2021). Our analysis confirmed that gait speed, cadence, and gait cycle duration were among the strongest predictors of group classification. This supports prior evidence that gait speed is a "sixth vital sign" in gerontology (Middleton et al., 2015). Reductions in gait speed and cadence correspond to neuromuscular slowing, impaired motor unit discharge, and reduced axonal conduction, collectively decreasing locomotor efficiency (Orssatto et al., 2022; Dewolf et al., 2021). Balance tests are other important parameters used to assess balance and motor control in older adults. Balance impairments, including increased postural sway and longer cycle durations, reflect age-related declines in proprioceptive acuity, vestibular integration, and cerebellar sensory–motor processing (Henry & Baudry, 2019; Ribeiro & Oliveira, 2007).

Muscle function was evaluated with parameters such as jump height and jump power measured by the two-legged hop test. In this study, jump power ranked particularly high in feature importance, consistent with previous literature indicating that power loss as a more sensitive predictor of disability than strength loss alone (Clark & Manini, 2012). Age-related sarcopenia, driven by reductions in type II fiber size, mitochondrial dysfunction, and altered motor unit recruitment, explains these differences (Larsson et al., 2019; Gustafsson & Ulfhake, 2024).

The BMD parameters used in this study are critical for determining bone loss and osteoporosis risk due to aging (Haseltine et al., 2021). Bone density, especially T-scores measured in regions such as the spine and femur, can be used to assess frailty and fall risk in older adults (Haseltine et al., 2021). Declines in BMD (spine, femur, radius) aligned with known mechanisms of skeletal aging—hormonal changes, chronic inflammation, and impaired calcium/vitamin D metabolism—leading to imbalanced bone remodeling and fracture risk (Fang et al., 2022; Demontiero et al., 2012). Bioelectrical impedance spectroscopy measures further revealed altered intracellular vs. extracellular fluid ratios, reflecting metabolic and compositional changes with aging. The use of BIA data in this study allows us to gain a more comprehensive understanding of body composition changes during aging (Guida et al., 2007).

Taken together, our findings highlight that applying machine learning to multidimensional biomedical datasets enables more than simple chronological age classification. By integrating biomarkers, gait parameters, muscle function, balance, bone density, and body composition measures, our models capture latent structures of functional aging. This approach not only distinguishes middle-aged from older adults but also identifies interaction patterns that reflect heterogeneous aging trajectories. Importantly, these insights support the development of precision gerontology tools, where individualized risk profiles can inform preventive and rehabilitative interventions, ultimately extending functional independence and quality of life in older adults. Nevertheless, these mechanisms are not isolated: muscle weakness exacerbates gait and balance impairments, while low bone density amplifies fall risk when combined with instability. This interdependence underscores the "network physiology of aging," where multisystem interactions shape clinical outcomes. While our cross-sectional design precludes causal inference, longitudinal biomarker-based studies are needed to confirm directionality.

4.3. Clinical and practical implications

In addition to demonstrating statistical robustness, this study sought to contextualize the clinical translation of ML-based classification in

gerontology, where predictive models can be applied for individualized risk stratification, monitoring of accelerated aging, and detection of pathological trajectories. Unlike traditional age-based classification, this approach integrates biomarkers, gait, muscle, balance, and body composition measures to generate a more dynamic profile of biological aging. We explicitly assessed the clinical relevance of group differences by computing effect sizes (Cohen's d) across all 57 variables such as gait speed and jump mechanography with moderate-to-large effects, reflecting clinically meaningful declines in physical function. For example, gait speed differences of 0.1–0.2 m/s have been widely accepted as clinically important thresholds in gerontology, and our observed effect sizes align with these benchmarks (Perera et al., 2016). These results are visualized in Fig. 1, which illustrates the relative magnitude of effects across functional domains, thereby improving clinical interpretability.

In addition, feasibility and cost-effectiveness also represent critical considerations. Compared with advanced imaging or genetic screening, the 36 geriatric clinical practice. The incremental costs of data integration and algorithmic deployment are thus expected to be modest, particularly if implemented through automated pipelines embedded in electronic health records or wearable device platforms. We also contextualized the classification performance of our models. The best-performing classifiers (SVM, PLS-DA) achieved a balanced accuracy of approximately 77 % and ROC-AUC values up to 0.98. While no direct gold standard exists for discriminating "middle-aged" from "older" adults, this level of performance compares favorably with established geriatric screening tools.

4.4. Limitations and future directions

In this study, there are several limitations. The first limitation is the absence of external validation, which restricts generalizability beyond the MIDUS sample. While the combination of Venetian blinds and stratified k-fold CV with repeated resampling increases internal robustness, future studies should validate these models on independent, population-representative datasets. Future studies can also expand sample diversity, integrate longitudinal trajectories, and explore multimodal models that combine functional, cognitive, and psychosocial predictors. Secondly, despite applying SMOTE to reduce class imbalance, the possibility of oversampling bias remains. Synthetic data may amplify noise in minority class samples, and balanced accuracy improvements should be interpreted cautiously. Furthermore, while SMOTE enhanced performance consistency, external validation in independent, more balanced cohorts will be required to confirm the generalizability of these results. Thirdly, common covariates such as sex, chronic disease history, and BMI, although compared between groups and reported in Supplementary Table 2, were not included as predictors in the ML models. While most covariates did not differ significantly between groups, residual confounding cannot be completely ruled out. Future studies may integrate these factors into adjusted models to improve robustness and generalizability. These factors may influence both biomarker and functional outcomes, and their integration into future predictive models could enhance accuracy and generalizability. On the other hand, for the future direction, aging research can extend beyond biological decline to encompass broader determinants of healthy aging, including quality of life, emotional well-being, social participation, and daily activity. Evidence shows that leisure activities and social networks serve as crucial sources of health and life satisfaction in later life (Parra-Rizo et al., 2022), while regular physical exercise is among the most effective non-pharmacological strategies to promote active and healthy aging (Sanchís-Soler et al., 2025). Integrating these perspectives with physiological markers may ensure a more comprehensive approach to designing interventions and shaping public health policies for older adults.

5. Conclusion

Our study highlights the potential of ML to discriminate between middle-aged and older adults based on biomarkers, gait, and muscle function measurements. By comparing multiple classifiers, we found that SVM offered the best balance between sensitivity and specificity, while PLS-DA achieved the highest overall accuracy. These findings suggest that integrating physiological markers of inflammation, gait performance, and muscle function provides a more objective and reproducible framework for aging classification than traditional chronological grouping.

Declaration of ethical standards

There is no need to obtain permission from the ethics committee for the article prepared. The article prepared has no conflict of interest with any person/institution.

CRediT authorship contribution statement

Veysel Alcan: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Assoc. Prof. Veysel Alcan reports was provided by Tarsus University. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.aggp.2025.100212.

Data availability

Data will be made available on request.

References

Chen, M., Wang, H., Yu, L., Yeung, E. H. K., Luo, J., Tsui, K. L., & Zhao, Y. (2022). A systematic review of wearable sensor-based technologies for fall risk assessment in older adults. *Sensors, 22*(18), 6752.

Choi, J. H. S., & Jung, D. H. (2025). The role of psychological factors in predicting self-rated health: Implications from machine learning models. *Psychology, Health & Medicine, 30*(6), 1158–1170.

Clark, B. C., & Manini, T. M. (2012). What is dynapenia? *Nutrition (Burbank, Los Angeles County, Calif.), 28*(5), 495–503.

Demontiero, O., Vidal, C., & Duque, G. (2012). Aging and bone loss: New insights for the clinician. *Therapeutic Advances in Musculoskeletal Disease, 4*(2), 61–76.

Dewolf, A. H., Sylos-Labini, F., Cappellini, G., Zhvansky, D., Willems, P. A., Ivanenko, Y., & Lacquaniti, F. (2021). Neuromuscular age-related adjustment of gait when moving upwards and downwards. *Frontiers in Human Neuroscience, 15*, Article 749366.

do Nascimento, C. F., Batista, A. F. M., Duarte, Y. A. O., & Chiavegatto Filho, A. D. P. (2022). Early identification of older individuals at risk of mobility decline with machine learning. *Archives of Gerontology and Geriatrics, 100*, Article 104625.

Eichler, N., Raz, S., Toledano-Shubi, A., Livne, D., Shimshoni, I., & Hel-Or, H. (2022). Automatic and efficient fall risk assessment based on machine learning. *Sensors, 22*(4), 1557.

Elam, C., Aagaard, P., Slinde, F., et al. (2021). The effects of ageing on functional capacity and stretch-shortening cycle muscle power. *Journal of Physical Therapy Science, 33*(3), 250–260.

Fang, H., Deng, Z., Liu, J., Chen, S., Deng, Z., & Li, W. (2022). The mechanism of bone remodeling after bone aging. *Clinical Interventions in Aging, 17*, 405–415.

Ferrucci, L., & Fabbri, E. (2018). Inflammageing: Chronic inflammation in ageing, cardiovascular disease, and frailty. *Nature Reviews. Cardiology, 15*(9), 505–522.

Ferrucci, L., Fabbri, E., & Studenski, S. (2018). Inflammaging: Characterization and implications for prevention and treatment. *Aging Clinical and Experimental Research, 30*(5), 505–512.

Furrer, R., & Handschin, C. (2025). Biomarkers of aging: From molecules and surrogates to physiology and function. *Physiological Reviews, 105*(3), 1609–1694.

Gökçe, E., Freret, T., & Langeard, A. (2025). Blood biomarker signatures for slow gait speed in older adults: An explainable machine learning approach. *Brain, Behavior, and Immunity, 124*, 295–304.

Gomes, S. R. B. S., von Schantz, M., & Leocadio-Miguel, M. (2023). Predicting depressive symptoms in middle-aged and elderly adults using sleep data and clinical health markers: A machine learning approach. *Sleep Medicine, 102*, 123–131.

Guida, B., Laccetti, R., Gerardi, C., et al. (2007). Bioelectrical impedance analysis and age-related differences of body composition in the elderly. *Nutrition, Metabolism, and Cardiovascular Diseases : NMCD, 17*(3), 175–180.

Gustafsson, T., & Ulfhake, B. (2024). Aging skeletal muscles: What are the mechanisms of age-related loss of strength and muscle mass, and can we impede its development and progression? *International Journal of Molecular Sciences, 25*(20), Article 10932.

Haseltine, K. N., Chukir, T., Smith, P. J., Jacob, J. T., Bilezikian, J. P., & Farooki, A. (2021). Bone mineral density: Clinical relevance and quantitative assessment. *Journal of nuclear Medicine : Official Publication, Society of Nuclear Medicine, 62*(4), 446–454.

Henry, M., & Baudry, S. (2019). Age-related changes in leg proprioception: Implications for postural control. *Journal of Neurophysiology, 122*(2), 525–538.

Hughes, M. L., Agrigoroaei, S., Jeon, M., Bruzzese, M., & Lachman, M. E. (2018). Change in cognitive performance from midlife into old age: Findings from the MIDUS study. *Journal of the International Neuropsychological Society : JINS, 24*(8), 805–820.

Jaul, E., & Barron, J. (2021). Characterizing the Heterogeneity of Aging: A Vision for a Staging System for Aging. *Frontiers in Public Health, 9*, Article 513557.

Jerome, G. J., Ko, S., Kauffman, D., Studenski, S. A., Ferrucci, L., & Simonsick, E. M. (2015). Gait characteristics associated with walking speed decline in older adults: Results from the Baltimore Longitudinal Study of Aging. *Archives of Gerontology and Geriatrics, 60*(2), 239–243.

Kawajiri, H., Mishina, H., Asano, S., Kono, Y., Hayashi, H., Niwa, J. I., Doyu, M., Kimura, S., & Yamada, S. (2019). Maximum walking speed at discharge could be a prognostic factor for vascular events in patients with mild stroke: A cohort study. *Archives of Physical Medicine and Rehabilitation, 100*(2), 230–238.

Larsson, L., Degens, H., Li, M., Salviati, L., Lee, Y. I., Thompson, W., Kirkland, J. L., & Sandri, M. (2019). Sarcopenia: Aging-related loss of muscle mass and function. *Physiological Reviews, 99*(1), 427–511.

Li, Y., Li, Q., Zhou, Z., et al. (2025). Machine learning models for predicting multimorbidity trajectories in middle-aged and elderly adults. *Scientific Reports, 15*, Article 24711.

Middleton, A., Fritz, S. L., & Lusardi, M. (2015). Walking speed: The functional vital sign. *Journal of Aging and Physical Activity, 23*(2), 314–322.

Montero-Odasso, M. M., Kamkar, N., Pieruccini-Faria, F., Osman, A., et al. (2021). Evaluation of clinical practice guidelines on fall prevention and management for older adults: A systematic review. *JAMA Network Open, 4*(12), Article e2138911.

Orssatto, L. B. R., Borg, D. N., Pendrith, L., Blazevich, A. J., Shield, A. J., & Trajano, G. S. (2022). Do motoneuron discharge rates slow with aging? A systematic review and meta-analysis. *Mechanisms of Ageing and Development, 203*, Article 111647.

Parra-Rizo, M. A., Vásquez-Gómez, J., Álvarez, C., Diaz-Martínez, X., Troncoso, C., Leiva-Ordoñez, A. M., Zapata-Lamana, R., & Cigarroa, I. (2022). Predictors of the level of physical activity in physically active older people. *Behavioral Sciences (Basel, Switzerland), 12*(9), 331.

Perera, S., Patel, K. V., Rosano, C., Rubin, S. M., Satterfield, S., Harris, T., Ensrud, K., Orwoll, E., Lee, C. G., Chandler, J. M., Newman, A. B., Cauley, J. A., Guralnik, J. M., Ferrucci, L., & Studenski, S. A. (2016). Gait speed predicts incident disability: A pooled analysis. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 71*(1), 63–71.

Piazza, J. R., Almeida, D. M., Dmitrieva, N. O., & Klein, L. C. (2010). Frontiers in the use of biomarkers of health in research on stress and aging. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 65*(5), 513–525.

Ribeiro, F., & Oliveira, J. (2007). Aging effects on joint proprioception: The role of physical activity in proprioception preservation. *European Review of Aging and Physical Activity : Official Journal of the European Group for Research into Elderly and Physical Activity, 4*, 71–76.

Ryff, C. D., Seema, T. E., & Weinstein, M. (2023). Midlife in the United States (MIDUS 3): Biomarker Project. *Inter-University Consortium for Political and Social Research [distributor]*.

Sanchís-Soler, G., Sebastiá-Amat, S., & Parra-Rizo, M. A. (2025). Mental health and social integration in active older adults according to the type of sport practiced. *Acta Psychologica (Amst.), 255*, Article 104920.

Sepúlveda, M., Arauna, D., García, F., Albala, C., Palomo, I., & Fuentes, E. (2022). Frailty in aging and the search for the optimal biomarker: A review. *Biomedicines, 10*(6), 1426.

Slijepcevic, D., Horst, F., Simak, M., Lapuschkin, S., Raberger, A. M., Samek, W., Breiteneder, C., Schöllhorn, W. I., Zeppelzauer, M., & Horsak, B. (2022). Explaining machine learning models for age classification in human gait analysis. *Gait & Posture, 97*, 252–253.

Usmani, S., Saboor, A., Haris, M., Khan, M. A., & Park, H. (2021). Latest research trends in fall detection and prevention using machine learning: A systematic review. *Sensors, 21*(15), 5134.

Verghese, J., Wang, C., Lipton, R. B., Holtzer, R., & Xue, X. (2016). Quantitative gait dysfunction and risk of cognitive decline and dementia. *Journal of Neurology, Neurosurgery, and Psychiatry, 84*(11), 1238–1243.

Zhavoronkov, A., Kochetov, K., Diamandis, P., & Mitina, M. (2020). PsychoAge and SubjAge: Development of deep markers of psychological and subjective age using artificial intelligence. *Aging, 12*(23), 23548–23577.