Journal of Computer Science

Research Article

Ensemble Feature Selection for Elderly Health Condition Classification

Rattanawadee Panthong

School of Information and Communication Technology, University of Phayao, Thailand

Article history
Received: 19-10-2025
Revised: 22-04-2026
Accepted: 05-08-2026

Email: rattanawadee.pa@up.ac.th

Abstract: Feature selection is a pivotal technique for constructing classification models with high-dimensional data. It facilitates the identification and elimination of redundant or irrelevant features, thereby reducing model complexity and enhancing predictive accuracy. The success of a classification model largely depends on the quality of the selected features. Accordingly, feature selection is performed during data preparation, prior to data processing, to support the classification of large-scale datasets. This study proposes an ensemble feature selection strategy based on embedded methods to identify the most informative feature subset for classifying elderly health conditions. The selected features include clinical indicators from Activities of Daily Living (ADL) and Comprehensive Geriatric Assessment (CGA), enabling a holistic evaluation of functional status. The proposed approach achieves an impressive classification accuracy of 98.11%. The resulting optimal feature subset provides a valuable foundation for planning and advancing elderly healthcare and rehabilitation.

Keywords: Ensemble, Feature Selection, Classification, Health Condition, Elderly

Introduction

The classification of elderly health conditions is a critical task in healthcare, as it supports early assessment, monitoring, and personalized care planning. With the increasing aging population, healthcare systems require accurate and efficient methods to analyze complex and high-dimensional data derived from Activities of Daily Living (ADL) and Comprehensive Geriatric Assessment (CGA) (Patrizio et al., 2021). However, such datasets often suffer from high dimensionality and class imbalance, which can negatively affect classification performance and model reliability (Pudjihartono et al., 2022; Theng and Bhoyar, 2024).

The primary challenges in applying machine learning to classify data types stem from datasets with a very large number of features, which significantly increases model complexity and computation time. Furthermore, when features lack meaningful interrelations, the classification performance is suboptimal. Effective classification typically relies on features that are interdependent or semantically related. Consequently, an integrated feature classification approach is proposed as an alternative. This method addresses the issue of high-dimensional data by managing feature interrelationships anFId reducing dimensionality, ultimately enhancing model efficiency and predictive accuracy. Feature classification, also known as supervised feature selection or supervised classification, aims to identify the features or samples that contribute most significantly to predictive performance. (El-Mageed et al., 2025).

Selecting appropriate features is critical for improving classification accuracy, especially in large and high-dimensional datasets. High-dimensional data can negatively impact predictive performance, increase computational costs, and lead to overfitting commonly referred to as the “curse of dimensionality” (Pudjihartono et al., 2022; El-Mageed et al., 2023; Theng and Bhoyar, 2024; Tukhtaev et al., 2024; Gaye et al., 2024). This issue is particularly prevalent in healthcare datasets, which often contain numerous features, making feature selection essential for reducing dimensionality, eliminating irrelevant or redundant variables, and enhancing learning efficiency.

In addition to challenges related to high-dimensional data and model efficiency, structured analytical frameworks have been shown to improve decision-making in complex domains. For example, the NIST framework organizes multidimensional information into structured categories through its core functions, including Identify, Protect, Detect, Respond, and Recover, thereby enhancing analytical accuracy and interpretability. Analytical accuracy and interpretability. This concept is

consistent with the study by systematically grouping elderly indicators from Activities of Daily Living (ADL) and Comprehensive Geriatric Assessment (CGA) into clearly defined health condition categories. This framework-based approach to organizing data according to a predefined structure enhances feature management efficiency and improves the performance of high-dimensional data classification in healthcare applications.

Feature selection is a process of identifying the most informative variables associated with the target label, serving as a crucial preprocessing step before model training (Chen et al., 2020; Kuzudisli et al., 2023; Hu et al., 2023). By selecting relevant features, this process simplifies model structure, reduces computational complexity, accelerates model development, and improves prediction accuracy. The performance of a classification model strongly depends on the quality of the selected features, which are organized into subsets of relevant variables to optimize the learning process and enhance predictive efficiency (Hoque et al., 2018; Uçar, 2020; Wang et al., 2021).

The Maeka Health Promotion Hospital in Phayao District, Phayao Province, Thailand conducts an annual survey on the health conditions of the elderly to identify health issues, monitor healthcare services, and assess Activities of Daily Living (ADL). The collected data are analyzed to evaluate the health status of the elderly and to inform strategies for health promotion and care. Effective classification of these data is essential for accurately assessing health conditions. However, the survey data are typically high-dimensional and imbalanced across classes, posing challenges for conventional analytical methods.

Despite the abundance of feature selection studies, few have addressed the specific challenges of geriatric multidimensional data, which are characterized by high sparsity, class imbalance, and complex inter-variable correlations. Existing literature often relies on single-criterion selection, which fails to provide a stable feature subset for clinical assessment. This study addresses this gap by proposing a novel synergistic framework that integrates the Fast Correlation-Based Filter (FCBF) with the Optimized Weights Forward (OWF) wrapper. This dual-layered approach is uniquely designed to first eliminate global redundancy and then iteratively optimize feature weights based on predictive interaction, ensuring a subset that is not only accurate but also clinically interpretable and robust a crucial requirement for geriatric healthcare systems. The primary contributions of this work are twofold: First, the development of the EFS-OWF ensemble framework that enhances predictive stability, and second, the validation of selected features that provide meaningful clinical insights for personalized elderly care.

The primary objective of this study is to address high-dimensional data and to identify the key features essential for classifying elderly health conditions and their activities of daily living. High dimensionality and multiple classes can negatively affect machine learning performance, reducing the accuracy and reliability of predictions. Moreover, the complexity of such datasets increases the difficulty of data processing. Feature selection is therefore a critical step in developing effective classification models for elderly health assessment. Ensemble Feature Selection, which combines multiple feature selection methods, leverages the strengths of each approach to identify the most informative subset of features, enhance correlation detection, and reduce the risk of selecting unstable subsets. While individual methods may yield locally optimal subsets, ensemble approaches provide more stable and reliable outputs across the entire dataset.

Background and Related Work

Comprehensive Geriatric Assessment

The Ministry of Public Health classifies the elderly into three groups based on a comprehensive geriatric assessment of their ability to perform activities of daily living:

(1) Bed-ridden patients: Elderly individuals who are unable to care for or assist themselves in daily activities, often considered disabled
(2) House-ridden patients: Elderly individuals who are able to care for and manage themselves independently within the home
(3) Socially active patients: Elderly individuals who are not only self-reliant but also capable of contributing to their family, community, and society

Comprehensive Geriatric Assessment refers to a multidimensional evaluation of an elderly individual's health, encompassing three key aspects.

Physical Health and Functional Assessment

This assessment involves obtaining the elderly person's medical history, including past and current illnesses, as well as family medical history. It also includes the evaluation of common symptoms prevalent among older adults, alongside an assessment of their nutritional status and Body Mass Index (BMI). Evaluating physical capacity helps identify limitations in performing daily activities and social interactions. Moreover, this assessment is critical for monitoring and predicting the progression of ailments, particularly functional decline, which adversely affects disease prognosis. This is achieved by evaluating the individual's ability to perform activities throughout the day and their competence in Activities of Daily Living (ADLs), such as getting out of bed, brushing teeth, walking, and using assistive devices in daily routines (Patrizio et al., 2021).

Psychological Health and Cognitive Assessment

Psychological and cognitive assessments are conducted for elderly individuals experiencing depression, which may lead to dementia. It is essential to identify and manage associated risk factors, including underlying diseases, comorbidities, lifestyle, and nutritional status. Consequently, tailored care plans and activities should be implemented to mitigate the risk of dementia progression (Patrizio et al., 2021).

Social and Environmental Assessment

Social and environmental assessment encompasses factors such as living arrangements, family relationships, caregiver support, participation in community activities, and availability of social support resources.

These three dimensions of health assessment enable early identification of health issues in the elderly, facilitating effective care management. Furthermore, elderly individuals receive care plans developed by a multidisciplinary team, which may include physicians, dentists, pharmacists, nurses, healthcare staff, and caregivers such as family members or friends. Elderly individuals who are generally healthy and self-sufficient should undergo health assessments at least annually. In contrast, those with chronic illnesses, complex health conditions, or limitations in daily activities require ongoing monitoring and regular check-ups to enable appropriate healthcare planning and health promotion interventions (Patrizio et al., 2021).

Activity of Daily Living (ADL)

ADL refers to an individual's ability to perform essential self-care tasks and routine activities such as bathing, dressing, preparing and eating meals, toileting, mobility, and household chores. ADL primarily focuses on individuals who are disabled, elderly, or chronically ill. The level of independence in performing these activities serves as an important indicator of an elderly person’s degree of self-reliance and helps determine the appropriate level of assistance required. Additionally, assessing ADL performance is crucial for monitoring progress during rehabilitation programs. In elderly populations, a decline in the ability to perform ADLs is commonly observed due to age-related physical changes (Yeo et al., 1995; Collin et al., 1988; Patrizio et al., 2021).

Feature Selection Method

Feature selection is a critical step in data preprocessing that reduces the dimensionality and number of features by selecting a relevant subset from the original dataset. This subset contains features that are important or have significant relationships for model development in classification tasks, thereby improving the accuracy of learning algorithms. The performance of classification models heavily depends on the selected features, especially in datasets with multiple classes and high dimensions, such as heritage, medical, and bioinformatics data.

Datasets with many features often include irrelevant or redundant attributes, which can negatively impact the efficiency and accuracy of machine learning classification (Saeys et al., 2008; Pudjihartono et al., 2022; Theng and Bhoyar, 2024; Mohtasham et al., 2024; Gong et al., 2024; Tukhtaev et al., 2024). In summary, feature selection reduces the dimensionality by choosing important and relevant features from the original data (Yousef, 2021; Mohtasham et al., 2024). Effective classification depends on these selected features, which also reduce model training time, minimize redundancy in the learning process, and eliminate irrelevant features that do not contribute to model performance (Cai et al., 2018; Hu et al., 2023; Gong et al., 2024).

Ensemble Feature Selection

Ensemble feature selection is a method that integrates multiple feature selection techniques by leveraging the strengths of each to identify the most appropriate and relevant subset of features, thereby reducing the risk of selecting unstable feature subsets (Tsymbal et al., 2005; Bolón-Canedo and Alonso-Betanzos, 2019; Spooner, 2023). This approach aims to obtain a stable and robust feature set by combining different selection methods and considering their individual advantages to produce the optimal subset. Generally, ensemble feature selection improves feature quality and minimizes the chance of instability in the selected subset. The process involves generating multiple subsets of the data using bootstrap sampling, followed by applying filter methods to rank features within each subset. Subsequently, the results from these diverse datasets are aggregated into a final feature subset through majority voting. The effectiveness of the resulting feature set depends significantly on the learning algorithm employed (Brahim and Limam, 2018; Wang et al., 2019).

The integrated feature selection method is employed to obtain a stable and robust feature set. This process involves generating multiple data subsets through bootstrap sampling. Subsequently, a filter method is applied to rank features within each subset, and the results from different subsets are aggregated into a final feature subset via majority voting. The effectiveness of the selected feature set largely depends on the learning algorithm used. In summary, ensemble feature selection combines multiple methods to identify the most relevant features while iteratively removing less significant ones, thereby enhancing predictive performance and robustness. This approach often improves model accuracy and mitigates overfitting by leveraging the strengths of diverse algorithms (Mera-Gaona et al., 2021; Joodaki et al., 2021; Kolukisa and Bakir-Gungor, 2023; Akhy et al., 2024).

In summary, the ensemble feature selection method integrates the advantages of both filter and wrapper techniques to form an iterative model that extracts the most contributory features at each iteration. Feature importance is quantified by assigning scores to each feature, with higher scores indicating greater relevance. This method not only identifies the features that significantly impact model predictions but also distinguishes unimportant features, enabling their removal. Consequently, this process enhances prediction efficiency and reduces feature dimensionality.

The filter selection methods employed in the present study are as follows:

- Gain Ratio (GR) is a feature selection technique that identifies the most relevant features for classification tasks, particularly within decision tree algorithms. It addresses the bias present in other measures such as information gain by normalizing the gain relative to the intrinsic information of the feature. This normalization leads to more balanced and effective feature selection, ultimately enhancing model performance (Prasetiyo et al., 2021).
- Fast Correlation-Based Filter (FCBF) is a feature selection method widely used in machine learning to identify the most pertinent features for a given model, especially when handling high-dimensional datasets where the number of features significantly exceeds the number of observations.

FCBF measures the correlation between features and the target variable often using Pearson correlation or mutual information metrics to detect features strongly associated with the output. The algorithm operates efficiently by initially ranking features based on their correlation with the target and subsequently selecting those with high relevance and low redundancy. It removes features that exhibit high correlation with already selected features to minimize overlap and redundancy. This iterative refinement process continues until an optimal, non-redundant subset of features is obtained, improving both model interpretability and predictive performance (Senliol et al., 2008; Yu and Liu, 2003).

The wrapper selection methods employed in the present study are as follows:

- Optimizing Weight (Forward) Approach is an effective method for selecting the most relevant features while balancing model performance and computational efficiency. This technique employs a Genetic Algorithm to compute attribute weights, where features with higher weights are deemed more relevant. By systematically optimizing these weights based on each feature's contribution, the approach identifies a feature subset that enhances model interpretability and efficiency (Lubis and Chandra, 2023)
- Forward Feature Selection (FWS) is a stepwise, wrapper-based feature selection method used to build predictive models. Starting from an empty set, it iteratively adds the feature that most improves model performance according to a predefined criterion such as maximizing accuracy or minimizing error—by evaluating the feature in combination with already selected features. This approach effectively reduces the search space by sequentially adding features rather than exhaustively evaluating all possible combinations, often yielding a feature subset that significantly enhances model accuracy (Venkatesh and Anuradha, 2019)

Embedded methods typically assign importance weights to features based on their contribution to model performance, enabling the identification of relevant features while reducing redundancy. Common examples include regularization-based methods and tree-based algorithms, where feature importance is inherently derived during training. In the context of ensemble feature selection, embedded methods play a critical role in improving stability and robustness by incorporating model-driven feature evaluation. This is particularly advantageous for high-dimensional and imbalanced datasets, where selecting informative and non-redundant features is essential for achieving reliable classification performance. Consequently, embedded approaches serve as valuable components in hybrid and ensemble-based feature selection frameworks (Pudjihartono et al., 2022; Claude et al., 2024). Therefore, this study integrates these approaches within an ensemble framework.

Related Research

Several researchers have proposed ensemble feature selection approaches for healthcare and disease-related datasets. For instance, Natarajan et al. (2025) introduced an ensemble feature selection technique to enhance diabetes prediction on imbalanced datasets. The method aims to improve model performance by refining feature selection, eliminating irrelevant data, and ensuring robustness. It employs multiple filter-based techniques ANOVA F-score, Fisher Score, Variance Threshold, and Point-Biserial Correlation to identify relevant variables. These are further refined by AdaptDiab, which ensures the selection of only the most relevant and non-redundant features. This is achieved by assigning weights to each feature selection method using mutual information and dynamically aggregating features across methods. The proposed approach demonstrated 100% stability across multiple runs, as measured by the Jaccard Index,

confirming its reliability and consistency.

Extending this line of research, Claude et al. (2024) explored Hybrid Ensemble Feature Selection (HEFS) strategies for identifying transcriptomic biomarkers in various cancers. HEFS proved to be a robust approach for discovering stable and generalizable biomarkers by integrating multiple feature selection techniques, including filter, wrapper, and embedded methods. The study evaluated four HEFS scenarios across several cancer types, including colorectal, kidney, lung, and endometrial cancers. It combined Differential Gene Expression (DEG) and variance-based filters with two resampling strategies Random Stratified (R-S) and Distribution-Balanced Stratified (DB-S). The results demonstrated that HEFS provided an effective balance between diversity and stability in feature selection, making it well-suited for biomarker discovery in transcriptomic datasets.

In line with this, Ensemble Feature Selection (EFS) methods have also been widely adopted in broader medical and bioinformatics applications to improve the robustness and accuracy of predictive models, particularly in high-dimensional and imbalanced datasets. For example, Kolukisa and Bakir-Gungor (2023) proposed a probabilistic EFS method for Coronary Artery Disease (CAD) diagnosis using three public datasets. Their approach combined exhaustive and probabilistic strategies with classifiers such as a Multi-Layer Perceptron (MLP), achieving high diagnostic accuracy (up to 91.78%) and demonstrating scalability and adaptability across diverse clinical datasets. Spooner et al. (2023) introduced data-driven thresholding strategies for EFS in Alzheimer's disease biomarker discovery. Using upper quartile and Kernel Density Estimation (KDE) for feature importance scoring, their methods achieved stable and biologically relevant feature sets across two clinical datasets (MAS and ADNI), with Robust Rank Aggregation (RRA) and thresholding algorithms yielding the best results. In the context of injury prediction, Kar and Sarkar (2022) proposed a hybrid feature reduction technique to enhance medical Decision Support Systems (MDSS), combining information gain and correlation-based methods to reduce dimensionality while preserving relevant features. This approach improved diagnostic efficiency and provided a scalable solution for high-dimensional datasets.

This study presents a novel approach for Oligodendroglioma brain tumor diagnosis by addressing the challenge of imbalanced healthcare data, employing a refined Hybrid Feature Selection method enhanced by Global Optimization, coupled with an Ensemble Classification technique (Huda et al., 2016). Collectively, these studies demonstrate the versatility and effectiveness of ensemble and hybrid feature selection techniques in improving predictive performance, enhancing model stability, and enabling clinically relevant insights across various domains in healthcare and biomedical research.

Ensemble feature selection methods have been widely studied for their ability to improve classification performance and reduce the risk of selecting unstable feature subsets. Previous studies have shown that combining multiple feature selection techniques can enhance robustness and provide more reliable results across different datasets.

Recent studies have increasingly focused on hybrid and ensemble feature selection approaches to further improve classification performance, particularly in high-dimensional and imbalanced datasets. Hybrid frameworks that integrate filter and wrapper methods with optimization techniques have been shown to effectively balance feature relevance and redundancy, thereby improving model generalization (Qaraad et al., 2021). Following this, recent studies have proposed robust ensemble feature selection methods that integrate multiple techniques with group Lasso. These methods aim to enhance stability and predictive performance, demonstrating superior results in both simulation and real-world biomedical applications (Maseno and Wang, 2024). In addition, ensemble feature selection methods that aggregate multiple selection strategies have demonstrated strong performance in enhancing robustness and reducing sensitivity to data variation. More recent approaches have incorporated hybrid optimization and ensemble techniques into feature selection frameworks to further improve performance in complex datasets (Le et al., 2024). In line with these advancements, this study addresses feature selection challenges in high-dimensional clinical data, particularly for survival analysis with censored and limited samples.

Despite these advancements, several limitations remain. Most existing studies primarily focus on improving classification accuracy, with limited attention to the stability and consistency of selected feature subsets. In addition, the integration of filter-based and embedded methods within a unified ensemble framework remains underexplored. Moreover, few studies address the challenges posed by high-dimensional and imbalanced healthcare datasets, in which both relevance and robustness are essential for reliable classification. These limitations highlight the need for more effective and stable feature selection strategies.

Materials and Methods

Dataset

The researcher reviewed relevant literature and collected elderly health condition data from Maeka Health Care Hospital, Maeka Sub-District, Muang District, Phayao Province, Thailand. The dataset comprises general demographic information, health problem data,

physical disability status, elderly behavior patterns, depression evaluation, and Activities of Daily Living (ADL) levels. Data were collected from 1,194 instances in 2022. The original dataset includes 54 features. The target variable consists of three class labels representing ADL levels: Social bound, home bound, and bed bound. Detailed descriptions of each attribute in the health condition dataset are provided in Table S1 of the supplementary material.

Proposed Method

Data Preparation

Data cleaning was performed to ensure accuracy and quality prior to feature selection, following the steps below.

Data Representation

Categorical variables were encoded into numerical formats to facilitate model processing and ensure consistency across the dataset. For example, binary attributes such as gender were encoded into numerical values.

Imputation

Missing values were handled using a combination of statistical and advanced imputation techniques. Initially, mean and mode imputation were applied to numerical and categorical variables, respectively. To better capture complex missing data patterns, missingness indicator variables were introduced, allowing the model to account for potential dependencies between missingness and target outcomes. This approach helps preserve data distribution and reduce bias associated with simple imputation methods. Furthermore, domain knowledge was incorporated to ensure that the imputed values were clinically plausible and consistent with the characteristics of elderly health data.

Management of Missing Values

Records with missing class labels were excluded due to the inability to reliably assess outcomes. Instances with excessive missing feature values were removed, provided that such removal did not exceed 5% of the total dataset. For missing values caused by data collection errors, an "unknown" category was assigned to preserve data completeness. In addition, duplicate records identified by overlapping ID card numbers were carefully reviewed, and only one record was retained based on expert validation to ensure data correctness and consistency for subsequent analysis.

Class Imbalance Handling

After handling missing data, class imbalance was addressed to improve model performance and prevent bias toward majority classes. In survey-based healthcare datasets, imbalance is common, particularly in multi-class scenarios where certain classes are underrepresented. To address this issue, class imbalance was addressed using a hybrid resampling strategy that combines Synthetic Minority Over-sampling Technique (SMOTE) and undersampling. SMOTE was applied to generate synthetic samples for minority classes, while undersampling was used to reduce the dominance of majority classes. This hybrid approach aims to achieve a balanced data distribution while preserving the underlying structure of the dataset.

Integrating imputation with resampling ensures that missing data are handled before class balancing, thereby preventing distortion of feature distributions and improving the reliability of synthetic data generation. This combined strategy enhances the robustness of the learning process by ensuring data completeness and balanced class representation, ultimately leading to improved classification performance.

Ensemble Feature Selection Method

After the dataset has been prepared, an ensemble feature selection method is employed to identify the optimal subset of features for model construction. In this phase, two techniques are applied: Ensemble Feature Selection with Fast Correlation-Based Filter (EFS-FCBF) and Ensemble Feature Selection with Optimized Weights Forward (EFS-OWF).

EFS-FCBF operates as a filter-based method by evaluating feature relevance and redundancy using statistical correlation measures. In contrast, EFS-OWF incorporates an embedded feature selection approach, in which feature importance is evaluated as part of the model training process. Specifically, the OWF method integrates feature selection within the learning algorithm through an optimization-based mechanism, implemented via an evolutionary search strategy. This process assigns and iteratively updates weights for each feature according to its contribution to classification performance. In addition, this mechanism is conceptually aligned with embedded tree-based learning algorithms, where feature importance is derived during training based on splitting criteria. During the optimization process, features that contribute more significantly to improving model performance are assigned higher weights, while less informative or redundant features are gradually eliminated. This enables dynamic, model-driven feature selection, improving both interpretability and computational efficiency.

Feature importance is determined based on the optimized weights, and features are subsequently ranked according to their scores. Both EFS-FCBF and EFS-OWF utilize a selection operator that prioritizes features with high weights and strong correlations, selecting the top-k ranked features.

Table 1: The definition of parameters for each approach

Method	Parameter
GR	Maximal depth = 10, normalize weight
FCBF	Threshold = 0.01, Classifier = Naive Bayes
FWS	maximal number of attributes = 50, Classifier = Naive Bayes
Optimize Selection (Evolutionary) / OSE	Min number of attributes = 1, Maximal number of generations = 50, population size = 5, normalize weight, Classifier = Naive Bayes
*Ensemble Feature Selection with FCBF (EFS-FCBF)	Ensemble size = 10, method = Top k, normalize weight, Classifier = Naive Bayes
* Ensemble feature selection with Optimize Weights Forward (EFS-OWF)	Ensemble size = 10, method = Top k, normalize weight Keep best =1, generations without improval =1, Classifier = Naive Bayes

The optimal value of k is identified through parameter optimization using RapidMiner, with the selected features accounting for approximately 40% to 60% of the total feature set. The optimal k values are 30 features for EFS-OWF and 25 features for EFS-FCBF.

Finally, the selected feature subsets from different methods are integrated using a voting mechanism to obtain a stable and robust final feature subset. This ensemble strategy effectively combines the strengths of filter and embedded methods, enhancing feature selection stability and reducing sensitivity to data variability. The parameter configurations for each approach are summarized in Table 1, while the overall conceptual framework of the proposed method is illustrated in Figure 1.

Classification With Machine Learning Algorithms

Following the construction of the classification model using the final selected feature subset described in step 3, a thorough evaluation process is conducted to assess and improve the model's classification accuracy while addressing variability and bias within the dataset. In this study, Decision Tree (DT), Random Forest (RF), and Stacking (SK) classifiers are utilized to evaluate the model's performance.

Step 5: Model Evaluation

Following model construction, a comprehensive evaluation is conducted to assess classification performance and ensure reliability. The entire dataset is evaluated using 10-fold cross-validation to obtain a robust and unbiased performance estimate. The dataset is randomly divided into ten equal subsets, where each subset is used once as a testing set while the remaining subsets are used for training. The final results are averaged across all folds to produce a reliable overall evaluation. Performance metrics, including accuracy, precision, recall, and F1-score, are used to provide a balanced assessment, particularly for imbalanced datasets.

Accuracy denotes the proportion of all instances that are correctly classified by the model and serves as a fundamental performance indicator in machine learning. It is derived from the counts of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), where TP and TN represent correct classifications of positive and negative instances, respectively, while FP and FN correspond to misclassified negative and positive instances.

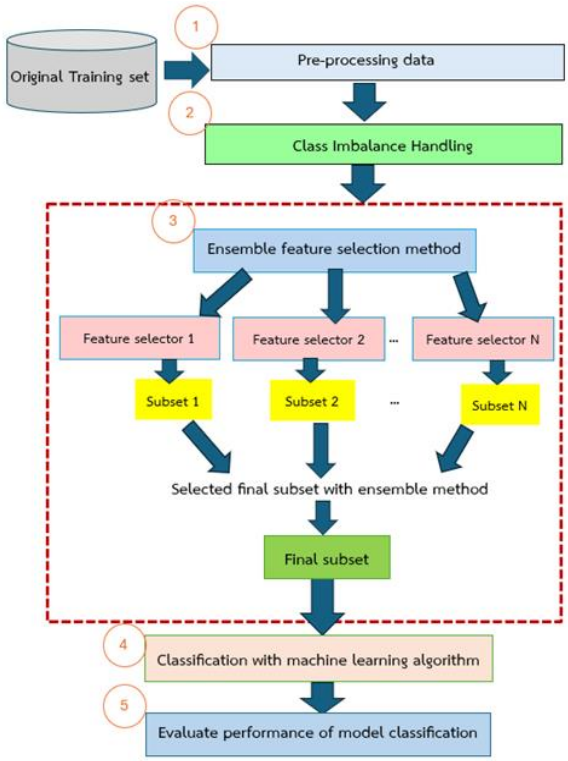

Fig. 1: Framework of ensemble feature selection for classification in assessing the health condition of the elderly

The accuracy metric is computed using these components to quantify the overall correctness of the classification model. (Sathe and Adamuthe, 2021):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Recall or sensitivity is the ratio of correctly classified positive instances and all instances designated to the positive class (Sathe and Adamuthe, 2021):

$$\text{Recall}_{\text{micro}} = \frac{TP_{c1}+TP_{c2}+\ldots+TP_{cn}}{TP_{c1}+FN_{c1}+TP_{c2}+FN_{c2}+\ldots+TP_{cn}+FN_{cn}}$$

$$\text{Recall}_{\text{macro}} = \frac{Recall_{c1+}Recall_{c2+\ldots}Recall_{cn}}{N}$$

Precision is the proportion of all true positives that are correctly labelled as the positive class (Sathe and Adamuthe, 2021):

$$\text{Precision}_{\text{micro}} = \frac{TP_{c1}+TP_{c2}+\ldots+TP_{cn}}{TP_{c1}+FP_{c1}+TP_{c2}+FP_{c2}+\ldots+TP_{cn}+FP_{cn}}$$

$$\text{Precision}_{\text{macro}} = \frac{Precision_{c1+}Precision_{c2+\ldots}Precision_{cn}}{N}$$

The F1-score is defined as the harmonic mean of precision and recall, providing a balanced evaluation of classification performance, particularly for imbalanced datasets (Hicks et al., 2022):

$$2 \cdot \frac{Precision \cdot Recall}{precision + Recall}$$

Results

In this section, the experimental results are presented and compared. The performance of the proposed method is evaluated using accuracy, recall, and precision metrics. The discussion is organized into several sub-sections.

From Table 2, the EFS-OWF method achieved the highest classification accuracy of 98.11% when applied with the SK learning algorithm, demonstrating its superior effectiveness. When combined with the RF classifier, EFS-OWF also outperformed other methods, achieving an accuracy of 92.24%. The EFS-FCBF method using the DT classifier showed competitive performance with an accuracy of 91.82%. Additionally, the single-feature selection method (FWS) delivered satisfactory results when used with the SK algorithm. These results highlight the robustness and effectiveness of the proposed feature selection methods across various classifiers and algorithms.

Figure 2 illustrates that the Ensemble Feature Selection (EFS) methods, including both EFS-FCBF and EFS-OWF, achieved higher average classification accuracy compared to single-feature selection methods. Among the single-feature approaches, the FCBF technique demonstrated a relatively good average accuracy, whereas the GR method exhibited lower accuracy than the other methods. These results clearly indicate that the ensemble feature selection methods provide superior data classification performance compared to single-feature selection techniques.

Figure 3 demonstrates that applying six feature selection techniques significantly improves data classification using the stacking algorithm.

Among them, the EFS-OWF method achieved the highest accuracy, underscoring its effectiveness in selecting key features to enhance model performance. This highlights the advantage of combining ensemble feature selection with stacking to achieve superior classification results.

Table 3 presents a comparison of classification performance in terms of recall using both macro and micro average values. According to the results, the EFS-OWF method achieved the highest recall performance when applied with the SK learning algorithm, yielding Recallu = 0.9811 and RecallM = 0.9836. This demonstrates the superior ability of EFS-OWF in identifying relevant instances compared to other feature selection methods. Furthermore, both EFS-FCBF and EFS-OWF showed consistently strong recall values when used with ensemble learning algorithms such as RF and SK.

In terms of precision, as shown in Table 4, the EFS-OWF method again delivered the highest performance with $Precision_{\mu}$ = 0.9811 and $Precision_{M}$ = 0.9776 when applied with the SK algorithm.

Table 2: presents the classification performance evaluated using the DT, RF, and SK learning algorithms

Method	DT	RF	SK
GR	85.17	86.20	87.06
FCBF	91.47	91.77	96.53
FWS	90.87	92.16	96.61
Optimize Selection (Evolutionary) / OSE	91.77	90.70	95.37
*Ensemble Feature Selection with FCBF (EFS-FCBF)	91.82	92.11	97.47
* Ensemble feature selection with Optimize Weights Forward (EFS-OWF)	91.60	92.24	98.11

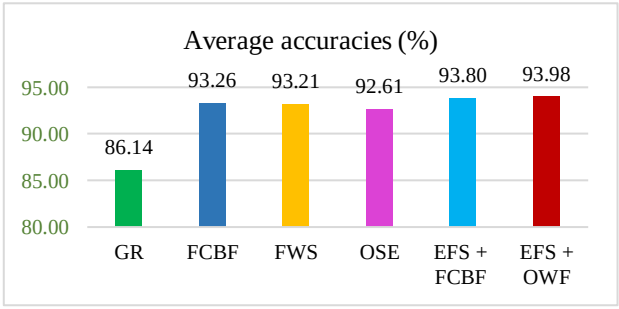

Fig. 2: Showing the comparison of average accuracies

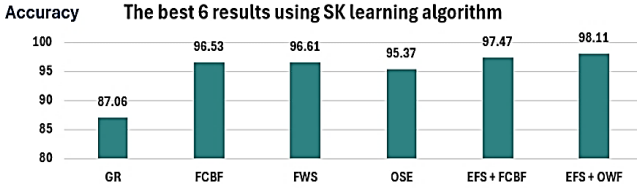

Fig. 3: The best six outputs of feature selection methods by SK classifier

Table 3: Recall comparison using micro-averaging and macro-averaging methods

Method	DT		RF		SK	
	$Recall_{\mu}$	$Recall_{M}$	$Recall_{\mu}$	$Recall_{M}$	$Recall_{\mu}$	$Recall_{M}$
GR	0.8517	0.8248	0.8620	0.8502	0.8706	0.8567
FCBF	0.9147	0.9072	0.9177	0.9098	0.9653	0.9667
FWS	0.9087	0.9015	0.9216	0.9137	0.9661	0.9667
OSE	0.9177	0.9111	0.8894	0.8729	0.9537	0.9543
EFS-FCBF	0.9181	0.9115	0.9211	0.9150	0.9747	0.9760
EFS-OWF	0.9160	0.9099	0.9224	0.9143	0.9811	0.9836

Table 4: Precision comparison using micro-averaging and macro-averaging methods

Method	DT		RF		SK	
	$Precision_{\mu}$	$Precision_{M}$	$Precision_{\mu}$	$Precision_{M}$	$Precision_{\mu}$	$Precision_{M}$
GR	0.8517	0.8323	0.8620	0.8436	0.8706	0.8510
FCBF	0.9147	0.9128	0.9177	0.9265	0.9653	0.9601
FWS	0.9087	0.9091	0.9216	0.9269	0.9661	0.9592
OSE	0.9177	0.9182	0.8894	0.8971	0.9537	0.9476
EFS-FCBF	0.9181	0.9181	0.9211	0.9279	0.9747	0.9704
EFS-OWF	0.9160	0.9137	0.9224	0.9321	0.9811	0.9776

Table 5: F1-score (%) comparison using micro-averaging and macro-averaging methods

Method	DT		RF		SK	
	$F1_{\mu}$	$F1_{M}$	$F1_{\mu}$	$F1_{M}$	$F1_{\mu}$	$F1_{M}$
GR	85.17	82.85	86.20	84.68	87.06	85.38
FCBF	91.47	91.00	91.77	91.80	96.53	96.34
FWS	90.87	90.53	92.16	92.02	96.61	96.29
OSE	91.77	91.46	88.94	88.49	95.37	95.09
EFS-FCBF	91.81	91.48	92.11	92.14	97.47	97.32
EFS-OWF	91.60	91.18	92.24	92.31	98.11	98.06

These results highlight the effectiveness of the EFS-OWF method in achieving both high sensitivity and precision in classification tasks, reinforcing its value as a robust feature selection approach. In summary, the micro-averaging method assigns equal weight to each individual instance and reflects the overall average performance across all predictions. In multi-class classification tasks, micro-averaged precision and recall values are identical, making this method particularly suitable for datasets with balanced class distributions. In contrast, the macro-averaging method calculates the average performance across classes independently and is more appropriate for imbalanced datasets, as it treats all classes equally regardless of their size.

Table 5 presents the F1-score (%) using both micro-averaging and macro-averaging. The results show that the proposed EFS-OWF method achieves the highest performance, especially when employing the SK learning technique. The close values between $F1_{\mu}$ and $F1_{M}$ indicate that the model performs consistently well across both majority and minority classes, demonstrating its robustness in handling imbalanced data. Overall, the ensemble feature selection methods outperform single methods, confirming their effectiveness in improving classification performance and maintaining a good balance between precision and recall.

The results in Table 6 summarize the optimal feature subsets and their corresponding classification performance across different EFS methods and classifiers. The findings show that EFS methods effectively select compact and informative feature subsets while maintaining high accuracy. Notably, EFS-OWF combined with the Stacking (SK) classifier achieved the highest accuracy of 98.11% using 30 features, whereas EFS-FCBF with Decision Tree (DT) used fewer features (25) while maintaining competitive performance. These results demonstrate the effectiveness of the proposed methods in balancing dimensionality reduction and predictive performance.

The findings indicate that the EFS-OWF method yielded the highest classification accuracy when combined with ensemble learning algorithms, specifically Random Forest (RF) and the Stacking (SK) classifier. In contrast, the EFS-FCBF method demonstrated strong performance when used with the Decision Tree (DT) classifier.

Table 6: Details of the optimal feature subsets selected by the EFS methods for each classifier

Methods	Classifiers	Selected features	Accuracy
EFS-FCBF	DT	25	91.82
EFS-OWF	RF	30	92.24
EFS-OWF	SK	30	98.11

These results suggest that the EFS approach is particularly suitable for domains characterized by high-dimensional data. The integration of ensemble learning techniques facilitates the construction of feature subsets that enhance predictive performance. A detailed description of the selected features is provided in Table S1 in the supplementary material.

The optimal subset derived from the EFS-FCBF method comprises 25 features, categorized into three groups:

1) Living environment information: Gender, Living situation, Occupation, Education, Social role, Income
2) Health problem information: Place of treatment, Problem of eye, Problem of sleep, BP1, Urinary, Phydis1, Phydis5, Results_diabetes, Moderate physical activity, Assess_fall, Assess_remembers1, Cdis3, Cdis4, Cdis5
3) Behavioral information: Smoking, Alcohol consumption, ADL1, ADL2, ADL10

Similarly, the final subset obtained from the EFS-OWF method consists of 30 features, also grouped into three categories:

1) Living environment information: Gender, Living situation, Occupation, Education, Social role, Income
2) Health problem information: Place of treatment, Problem of eye, Problem of sleep, BP1

Body mass index, Urinary, Phydis1, Phydis2, Phydis4, Assess_fall, Phydis5, Results_diabetes, Results_hypertension, Teeth, Moderate physical activity, Assess_remembers1, Assess_remembers2, Cdis2, Cdis9, Cdis14, Cdis15.

Behavioral Information: Smoking, Alcohol consumption, ADL10

A comparative analysis of the selected features reveals that 20 features are commonly chosen by both EFS-FCBF and EFS-OWF methods, including: Gender, Living situation, Occupation, Education, Social role, Income, Place of treatment, Smoking, Alcohol consumption, Moderate physical activity, ADL10, Problem of eye, Problem of sleep, BP1, Urinary, Phydis1, Phydis5, Results_diabetes, Assess_fall, and Assess_remembers1. It is noteworthy that the EFS-FCBF method produced a smaller subset compared to EFS-OWF. This is attributed to the nature of the FCBF method, which prioritizes features with strong correlations to the class label, thereby eliminating redundant or less informative features. Although this results in fewer selected features, it may come at the cost of slightly lower accuracy, as observed when compared to EFS-OWF. Nevertheless, the reduced feature set enhances model interpretability and efficiency, which can be particularly advantageous in real-world applications.

The stability of the selected feature subsets was evaluated using the Jaccard Index, which measures the similarity between feature sets obtained from different runs. A higher Jaccard value indicates greater stability (Mera-Gaona et al., 2021):

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

To assess the stability of the selected feature subsets, the Jaccard Index was computed between feature sets obtained from EFS-FCBF and EFS-OWF. The result (J = 0.5714) indicates a moderate level of stability, suggesting that the ensemble approach provides a reasonable balance between consistency and diversity in feature selection.

Figure 4 illustrates the confusion matrix heatmap of classification performance across three classes, with most samples correctly classified and minimal misclassification observed. The dominant diagonal values indicate strong classification performance. Misclassifications are limited and mainly occur between Social Bound and Home Bound, suggesting a slight overlap between these classes. In contrast, Bed Bound is clearly distinguished, reflecting the model's ability to capture distinct patterns in severe health conditions.

Figure 5 illustrates the confusion matrix heatmap for three classes, showing that most samples are correctly classified, as indicated by the dominant diagonal values. Minor misclassifications are observed, primarily between Social Bound and Home Bound and between Home Bound and Bed Bound. No misclassification occurs between Social Bound and Bed Bound. These results demonstrate strong overall performance, with clear class separation, particularly for the Bed Bound class.

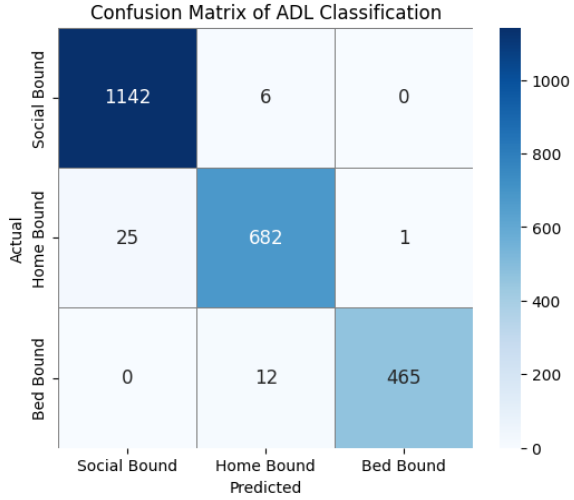

Fig. 4: Confusion matrix heatmap of the EFS-OWF with the SK model

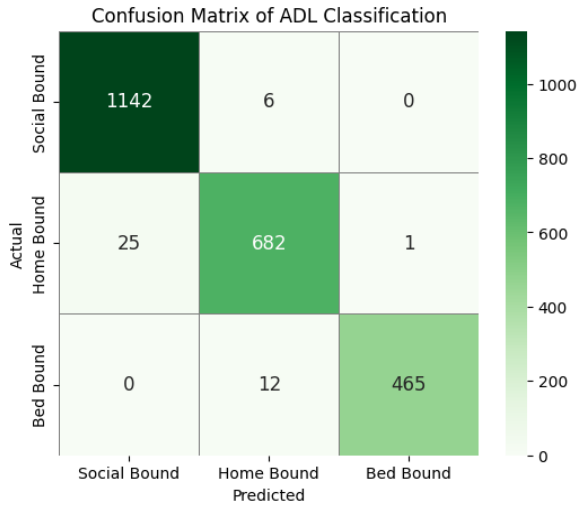

Fig. 5: Confusion matrix heatmap of the EFS-FCBF with the SK model

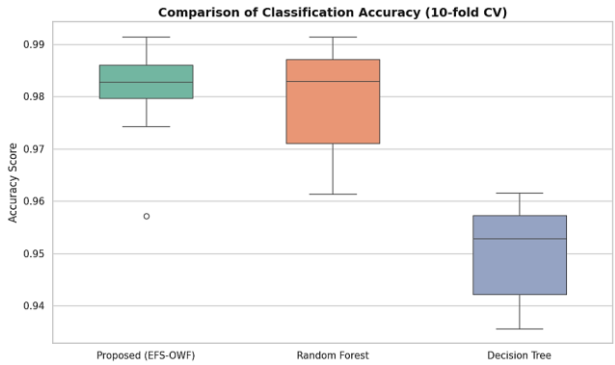

Fig. 6: Comparison of classification accuracy between the proposed EFS-OWF framework and various feature selection methods using 10-fold cross-validation

The proposed EFS-OWF framework demonstrates superior performance and stability compared to GR, FCBF, FWS, and OSE methods. By integrating dual-layered feature selection, the model achieves higher accuracy and lower variance across all folds. This empirical evidence, as illustrated in Figure 6, is further supported by statistical analysis where p-values less than 0.05 confirm that the performance improvements are statistically significant. Such results validate the framework's effectiveness in optimizing complex geriatric datasets for more reliable clinical classification.

Discussion

The experimental results demonstrate that the Ensemble Feature Selection with Optimized Weights Forward (EFS-OWF) method achieved the highest classification accuracy of 98.11% when combined with the Stacking (SK) classifier, while EFS-FCBF also yielded strong performance, reaching 97.47% with the Random Forest (RF) classifier. These results validate the effectiveness of the proposed ensemble feature selection framework in identifying highly relevant and diverse features for predicting elderly health conditions and classifying Activities of Daily Living (ADL) levels.

These findings are consistent with previous studies. For example, Kolukisa and Bakir-Gungor (2023) demonstrated the applicability of ensemble feature selection in medical diagnosis tasks, emphasizing its scalability and performance across clinical datasets. Similarly, Claude et al. (2024) and Natarajan et al. (2025) highlighted the stability and accuracy advantages of combining multiple feature selection techniques in complex biomedical data scenarios. In addition, the findings are consistent with Le et al. (2024) and highlight the effectiveness of integrating hybrid optimization techniques with ensemble frameworks to address the challenges of high-dimensional data complexity. The proposed EFS-OWF method effectively balances feature relevance and redundancy, indicating that optimized weighting strategies can significantly improve predictive robustness and outperform traditional feature selection methods. Accordingly, this study supports the idea that combining filter and wrapper approaches particularly FCBF and OWF can produce more stable, generalizable, and interpretable feature subsets, especially when integrated with advanced ensemble learning models. The discussion section has been revised by providing a more in-depth analysis of the results and strengthening the linkage with relevant studies. In alignment with these studies, the present work confirms that integrating filter and wrapper approaches specifically FCBF and OWF leads to more stable, generalizable, and interpretable feature subsets, particularly when used with ensemble learning algorithms such as RF and SK.

In particular, the superior performance in terms of accuracy, recall, precision, and F1-score can be attributed to several key factors. First, the ensemble selection framework effectively removes redundant and irrelevant features, thereby simplifying the model and mitigating overfitting. Second, the implementation of hybrid resampling (SMOTE and undersampling) during the data preparation stage helps balance class distributions, enhancing the model's ability to learn from underrepresented classes. Third, by selecting features with strong correlations to the target class, the framework fosters more efficient learning and greater predictive reliability.

Importantly, the final selected feature subsets obtained from both EFS-OWF and EFS-FCBF encompass a broad spectrum of variables, including living environment, health conditions, behavioral factors, and psychological status. This aligns with the multidimensional nature of geriatric assessment, which is essential for personalized

health planning and monitoring, especially for patients with varying levels of ADL dependency. These diverse and informative subsets not only enhance predictive accuracy but also provide valuable insights for healthcare providers to develop targeted interventions. When compared with single-method approaches (e.g., Gain Ratio, FWS, and OSE), the proposed EFS-based methods consistently demonstrate superior performance across all evaluation metrics and classifiers. Although standalone methods such as FWS and FCBF yield competitive results, they are surpassed by the ensemble approaches in both macro and micro-averaging for recall and precision. This further supports the conclusion that ensemble feature selection offers not only enhanced performance but also greater stability and interpretability in real-world healthcare applications.

The superior performance of the proposed ensemble method can be attributed to its ability to integrate multiple feature selection strategies, effectively balancing feature relevance and redundancy. In contrast, individual methods tend to rely on a single criterion, which often limits their performance when handling complex datasets. The ensemble feature selection approach reduces the risk of selecting unstable feature subsets by aggregating feature importance from diverse techniques. Instead of relying on a single method, the proposed framework produces more consistent and robust feature selection results, as evidenced by stable performance across multiple evaluation metrics, including accuracy and F1-score.

Furthermore, the stability of the selected feature subsets is supported by a Jaccard Index of 0.57, indicating a moderate level of consistency across different runs. This suggests that while the ensemble approach maintains robustness, some variability in feature selection remains, which is expected in ensemble-based methods where diversity can enhance generalization performance. Regarding practical applications, the proposed method shows strong potential to assist healthcare professionals in assessing elderly health conditions and planning personalized care. However, several limitations must be noted. The dataset was collected from a single institution, which may limit the generalizability of the findings. In addition, class imbalance and the limited dataset size may still affect model performance. Further validation using larger, more diverse datasets, as well as real-world deployment, is required to ensure long-term reliability and scalability.

In conclusion, the integration of EFS-OWF and EFS-FCBF with ensemble classifiers such as SK and RF resulted in superior classification models. These models effectively reduced dimensionality, maintained classification performance across all folds, and provided clinically relevant insights. The study thus confirms the value of ensemble feature selection in complex and high-dimensional healthcare datasets and suggests its broader applicability in similar domains.

Conclusion

This study proposes two ensemble feature selection methods, EFS-OWF and EFS-FCBF, by integrating multiple techniques, including FCBF, OWF, and FWS. Classification and evaluation were conducted using Decision Tree (DT), Random Forest (RF), and Stacking (SK) classifiers to identify the optimal feature subset for elderly health condition assessment. The results show that both EFS-OWF and EFS-FCBF outperform single feature selection methods, with EFS-OWF combined with the SK learning approach achieving the highest accuracy of 98.11% (Recallu = 0.9811, $Recall_M$ = 0.9836, Precisionμ = 0.9811, and $Precision_M$ = 0.9776). The high F1-score further confirms the balanced performance of the model across all classes. The confusion matrix analysis reveals that most instances are correctly classified, as indicated by dominant diagonal values, with only minor misclassification observed between Social Bound and Home Bound, suggesting slight overlap between these classes. In contrast, Bed Bound is clearly distinguished, reflecting the model's ability to capture distinct patterns in severe health conditions. The selected feature subsets include diverse variables critical for classifying elderly health and Activities of Daily Living (ADL). These subsets reflect the multidimensional aspects of geriatric health, encompassing health conditions, living environment, behaviors, and psychological factors like depression. Such comprehensive data supports personalized care planning and monitoring, enhancing health outcomes for elderly individuals. Overall, the proposed framework provides a robust and reliable solution for feature selection and classification in elderly health assessment. Future work will focus on incorporating advanced stability evaluation techniques and extending the proposed approach to other real-world datasets.

Acknowledgment

This research was supported by the Thailand Science Research and Innovation Fund and the University of Phayao. The authors would like to express their sincere gratitude to the experts from Mae Ka Health Promotion Hospital, Phayao, Thailand, for their valuable guidance in interpreting the health data.

Funding Information

This research received generous support from the Thailand Science Research and Innovation Fund and the University of Phayao.

Ethics

This manuscript is original and has not been published elsewhere. The corresponding author affirms that all authors have reviewed and approved the final version, and that no ethical concerns are associated with this work.

Supplementary Material

The supplementary material is available online. Table S1 provides a detailed breakdown of each attribute within the health condition dataset.

References

Akhy, S. A., Mia, Md. B., Mustafa, S., Chakraborti, N. R., Krishnachalitha, K. C., & Rabbany, G. (2024). A Comprehensive Study on Ensemble Feature Selection Techniques for Classification. *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*, 1319–1324. https://doi.org/10.23919/indiacom61295.2024.10498364

Bolón-Canedo, V., & Alonso-Betanzos, A. (2019). Ensembles for feature selection: A review and future trends. *Information Fusion*, *52*, 1–12. https://doi.org/10.1016/j.inffus.2018.11.008

Brahim, A. B., & Limam, M. (2018). Ensemble feature selection for high dimensional data: a new method and a comparative study. *Advances in Data Analysis and Classification*, *12*(4), 937–952. https://doi.org/10.1007/s11634-017-0285-y

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077

Chen, R.-C., Dewi, C., Huang, S.-W., & Caraka, R. E. (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, *7*(1), 52. https://doi.org/10.1186/s40537-020-00327-4

Claude, E., Leclercq, M., Thébault, P., Droit, A., & Uricaru, R. (2024). Optimizing hybrid ensemble feature selection strategies for transcriptomic biomarker discovery in complex diseases. *NAR Genomics and Bioinformatics*, *6*(3), lqae079. https://doi.org/10.1093/nargab/lqae079

Collin, C., Wade, D. T., Davies, S., & Horne, V. (1988). The Barthel ADL Index: A Reliability Study. *International Disability Studies*, *10*(2), 61–63. https://doi.org/10.3109/09638288809164103

El-Mageed, A. A. A., Abohany, A. A., & Elashry, A. (2023). Effective Feature Selection Strategy for Supervised Classification based on an Improved Binary Aquila Optimization Algorithm. *Computers & Industrial Engineering*, *181*, 109300. https://doi.org/10.1016/j.cie.2023.109300

El-Mageed, A. A. A., Abohany, A. A., & Hosny, K. M. (2025). Enhanced Binary Kepler Optimization Algorithm for effective feature selection of supervised learning classification. *Journal of Big Data*, *12*(1), 93. https://doi.org/10.1186/s40537-025-01125-6

Gaye, A., Diongue, A. K., Sylla, S. N., Diarra, M., Diallo, A., Talla, C., & Loucoubar, C. (2024). Supervised Classification of High-Dimensional Correlated Data: Application to Genomic Data. *Journal of Classification*, *41*(1), 158–169. https://doi.org/10.1007/s00357-024-09463-5

Gong, H., Li, Y., Zhang, J., Zhang, B., & Wang, X. (2024). A new filter feature selection algorithm for classification task by ensembling pearson correlation coefficient and mutual information. *Engineering Applications of Artificial Intelligence*, *131*, 107865. https://doi.org/10.1016/j.engappai.2024.107865

Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, *12*(1), 5979. https://doi.org/10.1038/s41598-022-09954-8

Hoque, N., Singh, M., & Bhattacharyya, D. K. (2018). EFS-MI: an ensemble feature selection method for classification. *Complex and Intelligent Systems*, *4*(2), 105–118. https://doi.org/10.1007/s40747-017-0060-x

Hu, J., Pan, K., Song, Y., Wei, G., & Shen, C. (2023). An improved feature selection method for classification on incomplete data: Non-negative latent factor-incorporated duplicate MIC. *Expert Systems with Applications*, *212*, 118654. https://doi.org/10.1016/j.eswa.2022.118654

Huda, S., Yearwood, J., Jelinek, H. F., Hassan, M. M., Fortino, G., & Buckland, M. (2016). A Hybrid Feature Selection With Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis. *IEEE Access*, *4*, 9145–9154. https://doi.org/10.1109/access.2016.2647238

Joodaki, M., Dowlatshahi, M. B., & Joodaki, N. Z. (2021). An ensemble feature selection algorithm based on PageRank centrality and fuzzy logic. *Knowledge-Based Systems*, *233*, 107538. . https://doi.org/10.1016/j.knosys.2021.107538

Kar, B., & Sarkar, B. K. (2022). A Hybrid Feature Reduction Approach for Medical Decision Support System. *Mathematical Problems in Engineering*, *2022*, 1–20. https://doi.org/10.1155/2022/3984082

Kolukisa, B., & Bakir-Gungor, B. (2023). Ensemble feature selection and classification methods for machine learning-based coronary artery disease diagnosis. *Computer Standards & Interfaces*, *84*, 103706. https://doi.org/10.1016/j.csi.2022.103706

Kuzudisli, C., Bakir-Gungor, B., Bulut, N., Qaqish, B., & Yousef, M. (2023). Review of feature selection approaches based on grouping of features. *PeerJ*, *11*, e15666. https://doi.org/10.7717/peerj.15666

Le, P., Gong, X., Ung, L., Yang, H., Keenan, B. P., Zhang, L., & He, T. (2024). A robust ensemble feature selection approach to prioritize genes associated with survival outcome in high-dimensional gene expression data. *Frontiers in Systems Biology*, *4*, 5595. https://doi.org/10.3389/fsysb.2024.1355595

Lubis, A. I., & Chandra, R. (2023). Forward Selection Attribute Reduction Technique for Optimizing Naïve Bayes Performance in Sperm Fertility Prediction. *Sinkron*, *8*(1), 275–285. https://doi.org/10.33395/sinkron.v8i1.11967

Maseno, E. M., & Wang, Z. (2024). Hybrid wrapper feature selection method based on genetic algorithm and extreme learning machine for intrusion detection. *Journal of Big Data*, *11*(1), 24. https://doi.org/10.1186/s40537-024-00887-9

Mera-Gaona, M., López, D. M., Vargas-Canas, R., & Neumann, U. (2021). Framework for the Ensemble of Feature Selection Methods. *Applied Sciences*, *11*(17), 8122. https://doi.org/10.3390/app11178122

Mohtasham, F., Pourhoseingholi, M., Nazari, S. S. H., Kavousi, K., & Zali, M. R. (2024). Comparative analysis of feature selection techniques for COVID-19 dataset. *Scientific Reports*, *14*(1), 18627. https://doi.org/10.1038/s41598-024-69209-6

Natarajan, K., Baskaran, D., & Kamalanathan, S. (2025). An adaptive ensemble feature selection technique for model-agnostic diabetes prediction. *Scientific Reports*, *15*(1), 6907. https://doi.org/10.1038/s41598-025-91282-8

Patrizio, E., Calvani, R., Marzetti, E., & Cesari, M. (2021). Physical Functional Assessment in Older Adults. *The Journal of Frailty & Aging*, *10*(2), 141–149. https://doi.org/10.14283/jfa.2020.61

Prasetiyo, B., Alamsyah, Muslim, M. A., & Baroroh, N. (2021). Evaluation of feature selection using information gain and gain ratio on bank marketing classification using naïve bayes. *Journal of Physics: Conference Series*, *1918*(4), 042153. https://doi.org/10.1088/1742-6596/1918/4/042153

Pudjihartono, N., Fadason, T., Kempa-Liehr, A. W., & O'Sullivan, J. M. (2022). A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Frontiers in Bioinformatics*, *2*, 27312. https://doi.org/10.3389/fbinf.2022.927312

Qaraad, M., Amjad, S., Manhrawy, I. I. M., Fathi, H., Hassan, B. A., & Kafrawy, P. E. (2021). A Hybrid Feature Selection Optimization Model for High Dimension Data Classification. *IEEE Access*, *9*, 42884–42895. https://doi.org/10.1109/access.2021.3065341

Saeys, Y., Abeel, T., & Peer, Y. V. de. (2008). Robust Feature Selection Using Ensemble Feature Selection Techniques. *Machine Learning and Knowledge Discovery in Databases, ECML/PKDD 2008*, 313–325. https://doi.org/10.1007/978-3-540-87481-2_21

Sathe, T. M., & Amol, C. A. (2021). Comparative Study of Supervised Algorithms for Prediction of Students' Performance. *International Journal of Modern Education and Computer Science*, *13*(1), 1–21. https://doi.org/10.5815/ijmecs.2021.01.01

Senliol, B., Gulgezen, G., Yu, L., & Cataltepe, Z. (2008). Fast Correlation Based Filter (FCBF) with a different search strategy. *2008 23rd International Symposium on Computer and Information Sciences*, 1–6. https://doi.org/10.1109/iscis.2008.4717949

Spooner, A., Mohammadi, G., Sachdev, P. S., Brodaty, H., & Sowmya, A. (2023). Ensemble feature selection with data-driven thresholding for Alzheimer's disease biomarker discovery. *BMC Bioinformatics*, *24*(1), 9. https://doi.org/10.1186/s12859-022-05132-9

Theng, D., & Bhoyar, K. K. (2024). Feature selection techniques for machine learning: a survey of more than two decades of research. *Knowledge and Information Systems*, *66*(3), 1575–1637. https://doi.org/10.1007/s10115-023-02010-5

Tsymbal, A., Pechenizkiy, M., & Cunningham, P. (2005). Diversity in search strategies for ensemble feature selection. *Information Fusion*, *6*(1), 83–98. https://doi.org/10.1016/j.inffus.2004.04.003

Tukhtaev, A., Turimov, D., Kim, J., & Kim, W. (2024). Feature Selection and Machine Learning Approaches for Detecting Sarcopenia Through Predictive Modeling. *Mathematics*, *13*(1), 98. https://doi.org/10.3390/math13010098

Uçar, M. K. (2020). Classification Performance-Based Feature Selection Algorithm for Machine Learning: P-Score. *IRBM*, *41*(4), 229–239. https://doi.org/10.1016/j.irbm.2020.01.006

Venkatesh, B., & Anuradha, J. (2019). A Review of Feature Selection and Its Methods. *Cybernetics and Information Technologies*, *19*(1), 3–26. https://doi.org/10.2478/cait-2019-0001

Wang, J., Xu, J., Zhao, C., Peng, Y., & Wang, H. (2019). An ensemble feature selection method for high-dimensional data based on sort aggregation. *Systems Science & Control Engineering*, *7*(2), 32–39. https://doi.org/10.1080/21642583.2019.1620658

Wang, L., Jiang, S., & Jiang, S. (2021). A feature selection method via analysis of relevance, redundancy, and interaction. *Expert Systems with Applications*, *183*, 115365. https://doi.org/10.1016/j.eswa.2021.115365

Yeo, D., Faleiro, R., & Lincoln, N. (1995). Barthel ADL Index: a comparison of administration methods. *Clinical Rehabilitation*, *9*(1), 34–39. https://doi.org/10.1177/026921559500900105

Yousef, M. M. (2021). A Hybrid Feature Selection Method for Improve the Accuracy of Medical Classification Process. *International Journal of Innovative Technology and Exploring Engineering*, *11*(1), 50–55.

Yu, L., & Liu, H. (2003). Feature selection for high-dimensional data: A fast correlation-based filter solution. *Springer Nature Link*, 21–24. https://doi.org/10.1007/978-3-540-75175-5_30

Table S1: Health condition survey data and ability assessment in doing their daily activities of the elderly

Feature	Descriptions	Feature	Descriptions
1. Gender	1 = Male 2 = Female	2. Living Situation	1 = Alone 2 = Spouse taking Care 3 = Children Taking Care 4 = Brother/Sister Taking Care 5 = Cousin Taking Care 6 = No one
3. Education	1 = No education 2 = Primary 3 = Secondary 4 = High school 5 = Vocational Education Level 6 = Higher Vocational Level ./ Diploma 7 = bachelor's Degree 8 = Graduate Level	4. Occupation	1 = farmer 2 = Labor 3 = Entrepreneur 4 = Animal Keeper 5 = pensioner 6 = others
5. Social Role	1 = Local Leader 2 = Village Volunteer 3 = Village Safety Carer 4 = Housewife 5 = Temple Helper 6 = Village Committee 7 = Temple Committee 8 = School Committee 9 = Village Fund Committee 10 = Others Please identify 11 = Not have	6. Place of treatment	1 = Phayao Hospital 2 = Phayao University Hospital 3 = Maeka Health Promotion Hospital 4 = Private hospital 5 = Private Clinic 6 = Buy medicine and take it yourself 7 = Others (Please identify)
7. Income per Month	1 = Not more than 1,500 baht 2 = Between 1,501 – 3,000 baht 3 = Between 3,001 – 4,500 baht	8. MPA (Movement activities at moderate level)	1 = Always at least 3 times a week 2 = Sometimes but not regular, less than 3 times a week 3 = Never because I can't do it
9. Smoking	4= More than 4,500 baht 1= Never 2= Ever but not smoke now 3=Still smoke	10. Alcohol	1 = Never drink 2 = Ever but not drink now 3 = Still drink
11. Phydis1 (Physical and movement disability)	1= Have 0=Not have	12. Phydis2 (Eyesight disability)	1= Have 0=Not have
13. Phydis3 (Ear disability)	1= Have 0=Not have	14. Phydis3 (Learning and intellect disability)	1= Have 0=Not have
15. Phydis5 (Mental disability/ behavior)	1= Have 0=Not have	16. Phydis6	1= Have 0=Not have
17. BP1 (Blood pressure check, first time)	1=normal (not more than 120/80), 2=starting high (120-139/80-89), 3=high (140-159/9099), 4=very high (>160/100) 0= no information	18. BMI (body mass index (nutritional status)	1=obese range 2=healthy weight range 3=overweight range 4=obese range 5=very obese range 0=Unknown
19. ADL1 (Feeding)	0= unable,	20. ADL2 (Grooming)	0= needs help with personal care

	1= needs help cutting, spreading butter 2=independent (food provided within reach)		1= independent face/hair/teeth/shaving (implements provided)
21. ADL3 (Transfer)	0= unable – no sitting balance 1= major help (one or two people, physical), can sit 2= minor help (verbal or physical) 3 = independent	22. ADL4 (Toilet use)	0 = dependent 1= needs some help, but can do something alone 2 = independent (on and off, dressing, wiping)
23. ADL5 (Mobility)	0 = immobile 1=wheelchair independent, including corners 3= walks with help of one person (verbal or physical) 4= independent	24. ADL6 (Dressing)	2 = dependent 3 = needs help, but can do about half unaided 2 = independent (including buttons, zips, laces)
25. ADL7 (Stairs)	4 = unable 5 = needs help (verbal, physical, carrying aid) 2 = independent up and down	26. ADL8 (Bathing)	0 = dependent, 1 = independent (or in shower)
27. ADL9 (Bowels)	6 = Incontinent (or needs to be given enema) 1 = occasional accident (once/week) 2 = continent	28. ADL10 (Bladder)	0 = incontinent, or catheterized and unable to manage 1 = occasional accident (max. once per 24 hours) 2 = continent (for over 7 days)
29. diabetes (Yearly diabetes checkup result)	1=Normal 2= Abnormal 3= Never check up	30. blood pressure (Yearly high blood pressure result)	1 = Normal 2= Abnormal
31. Teeth (You have at least twenty teeth now)	1= Have 2= Not have	32. Assess_fall (The assessment of falling condition)	1= less than 30 seconds 2=more than 30 seconds 3= unable to walk
33. Sleep problem diagnosis	1=Insomnia 2= Oversleeping 3=Sleepwalking 4=Snoring 5=Not have	34. Eye problem (Eyesight testing result)	1=Short-sighted 2= Long-sighted 3=Cataract 4=Glaucoma 5=Pterygium 6= Pinguecula 7= Not have
35. Feel_depressed (Feel sad, frustrated and disappointed)	1= Have 0=Not have	36. Feel_bored (Feel bored not enjoyable)	1= Have 0=Not have
37. Assess_remem 1 (The assessment of the dementia condition by remembering and telling all words of things correctly.	1= true 2= false	38. Assess_remem 2 (The assessment of the dementia condition by telling one's own name and age correctly)	1=true 2= false
39. Urinary (The screening of urinary incontinence)	1=yes, 0= No	40. Cdis1 (Diabetes Mellitus)	1= Yes 0=No
41. Cdis2 (hypertension)	1= Yes 0=No	42. Cdis3 (Dyslipidemia)	1= Yes 0=No
43. Cdis4 (Cardiovascular disease)	1= Yes 0=No	44. Cdis5 (Stroke)	1= Yes 0=No
45. Cdis6 (Kidney failure)	1= Yes 0=No	46. Cdis7 (Asthma)	1= Yes 0=No
47. Cdis8 (Emphysema)	1= Yes 0=No	48. Cdis9 (Cancer)	1= Yes 0=No
49. Cdis10 (Anemia)	1= Yes 0=No	50. Cdis11 (Psychiatric disease)	1= Yes 0=No
51. Cdis12 (Epilepsy)	1= Yes 0=No	52. Cdis13 (Gout)	1= Yes 0=No
53. Cdis14 (Knee Osteoarthritis)	1= Yes 0=No	54. Cdis15 (Others)	1= Yes 0=No

55. Elderly group (Group of Elderly which is divided by their abilities in doing their daily lives)	Class0= Social Bound Class1= Home Bound Class2= Bed Bound