Acta Informatica Pragensia
2026, Volume 15, Issue 1, 157–172
https://doi.org/10.18267/j.aip.295

PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Article Open Access

Modular Local Classification via Cluster-Guided Feature Selection in Tabular Data

Leila Boussaad

Department of Management, Faculty of Economics, University of Batna 1, Batna, Algeria

Corresponding author: Leila Boussaad (leila.boussaad@univ-batna.dz)

Editorial Record

First submission received:
August 9, 2025

Revision received:
October 9, 2025

Accepted for publication:
October 24, 2025

Academic Editor:
Michal Munk
Constantine the Philosopher University in Nitra, Slovakia

This article was accepted for publication by the Academic Editor upon evaluation of the reviewers' comments.

How to cite this article:
Boussaad, L. (2026). Modular Local Classification via Cluster-Guided Feature Selection in Tabular Data. *Acta Informatica Pragensia*, 15(1), 157–172. https://doi.org/10.18267/j.aip.295

Abstract

Background: Many real-world tabular datasets are heterogeneous, with distinct regions of the feature space exhibiting different feature–label relationships. Conventional global classifiers often miss these local patterns, reducing both predictive accuracy and interpretability.

Objective: This study aims to design a modular classification framework that combines local specialization with global consistency to enhance predictive performance and interpretability in heterogeneous tabular data.

Methods: The author proposes Cluster-guided local feature selection with top-2 voting and fallback (CGLFS+), which integrates unsupervised clustering, cluster-specific feature selection and lightweight local models. Final predictions combine top-2 local decisions with a global fallback classifier for robustness. The framework was evaluated on five diverse benchmark datasets using repeated stratified cross-validation.

Results: CGLFS+ achieved consistent gains in accuracy and macro F1 over strong baselines, with statistically significant improvements and competitive inference times.

Conclusion: CGLFS+ successfully balances local adaptation and global consistency, providing a scalable and interpretable approach well suited to heterogeneous domains such as healthcare, chemistry and finance.

Index Terms

Local models; Feature selection; Clustering; Modular classification; Tabular data interpretable machine learning.

1 INTRODUCTION

In many real-world classification problems, relying on a single global model often fails to capture the inherent heterogeneity and complex structure of the data. Datasets derived from diverse sources, such as clinical diagnostics, consumer behaviour or industrial monitoring, frequently exhibit subpopulation-specific distributions, localized feature relevance and non-uniform decision boundaries. Despite this, most standard machine learning pipelines are designed under the global assumption: a single classifier trained over the entire input space, using a fixed set of features and shared decision rules.

Recent research into interpretable machine learning and modular architectures has begun to question this monolithic paradigm. Several studies have investigated local learning strategies such as mixture-of-experts (MoE) (Yuksel et al., 2012), hierarchical models (Ismail et al., 2022) or instance-based approaches (Kim et al., 2015).

While these models improve adaptability, they often suffer from rigid architecture constraints, limited interpretability or overlook the locality of feature relevance. In particular, feature selection – a critical step for both performance and transparency – is frequently applied globally, failing to account for regional differences in predictive structure (Hancer et al., 2020).

In this paper, we propose an enhanced modular classification framework named **cluster-guided local feature selection with top-2 voting and fallback (CGLFS+).** The core idea of CGLFS+ is to exploit structural heterogeneity in tabular data by combining three complementary mechanisms: localized feature relevance, hybrid inference and modular model specialization.

Firstly, the dataset is partitioned into coherent subregions via unsupervised clustering, without supervision labels. For each cluster, we perform local feature selection using mutual information to identify the most predictive features within that cluster. To improve robustness and coverage, we also apply global feature selection and inject the top-k globally informative features into every cluster-specific subset. This local-global feature fusion balances specificity and generalization, enhancing the expressive power of each submodel while preserving interpretability.

Secondly, within each cluster, the method automatically selects the most suitable classifier from a small pool (e.g., logistic regression, decision tree or SVM), using internal cross-validation. This ensures that model complexity is aligned with the local data distribution, avoiding overfitting or underfitting across diverse subpopulations.

Thirdly and critically, the inference process in CGLFS+ is not based on hard assignment. Instead, we introduce a **top-2 voting with fallback** strategy: at test time, each instance is evaluated by the two nearest cluster-specific models (based on distance to cluster centroids) and a consensus prediction is attempted. In the case of disagreement, a global fallback model is consulted to resolve the ambiguity. This layered inference mechanism increases robustness in noisy or borderline regions and yields more stable predictions, as confirmed by our empirical evaluation. This work introduces a modular classification framework, CGLFS+, with the following **key contributions**:

- We propose a hybrid classification strategy that combines unsupervised clustering, cluster-specific feature selection and adaptive model assignment to address data heterogeneity.
- We introduce a local-global feature fusion mechanism, allowing each local model to benefit from both globally and locally informative variables, thereby enhancing generalization and interpretability.
- We design a robust inference procedure based on top-2 cluster voting with fallback to a global model, improving predictive reliability in ambiguous regions.
- We conduct an extensive empirical evaluation on five diverse tabular datasets, demonstrating that CGLFS+ consistently outperforms standard global classifiers, with validated statistical significance and competitive inference time.

The rest of this paper is organized as follows. Section 2 provides an overview of existing research related to local models, feature selection and interpretable clustering techniques. Section 3 details the proposed methodology. Section 4 describes the experimental setup and presents the evaluation protocol. The findings and their implications, along with limitations, are discussed in Section 5. Finally, Section 6 summarizes the contributions and outlines potential directions for future research.

2 RELATED WORK

Many conventional supervised learning algorithms operate under the assumption of a uniform mapping from input features to target labels. However, real-world tabular datasets often violate this assumption due to underlying heterogeneity: different subpopulations may exhibit distinct patterns, different relevant variables and decision boundaries. This limitation has motivated a variety of approaches that introduce local adaptability into the modelling pipeline.

2.1 Local models and mixture-of-experts

The mixture-of-experts (MoE) framework (Yuksel et al., 2012) is a foundational technique for modelling heterogeneity by partitioning the input space and assigning a specialized predictor (expert) to each region, guided by a gating mechanism. While effective in complex settings, traditional MoEs often rely on non-transparent gating functions and act as black-box ensembles, limiting their interpretability. Recent advances such as interpretable MoEs

(Ismail et al., 2022) and relational MoEs (Oyamada and Nakadai, 2017) aim to improve transparency by enforcing sparse or structured gating, but typically do not support local feature selection.

Early efforts to integrate feature selection within local models include the work of Peralta and Soto (Peralta and Soto, 2014), who explored region-wise variable selection in MoE-style regression. However, their approach used a fixed model class across regions and did not address classification tasks or modular inference.

2.2 Feature selection and interpretability

Feature selection is a cornerstone of interpretable machine learning, with mutual information, chi-squared tests and wrapper-based methods widely used to improve generalization and reduce redundancy (Kheradpisheh et al., 2014). However, most of these techniques are applied globally, failing to account for the possibility that feature relevance may vary across the data space.

Several authors have proposed instance- or class-specific feature selection frameworks (Ma and Lu, 2024; Shi et al., 2025), but these often require prior knowledge or lack modularity. Kheradpisheh et al. (2014) proposed expert-based models with different feature subsets, though their work lacked a clear integration of clustering or adaptive classifier selection.

2.3 Clustering-aware learning and interpretability

Several studies have explored combining unsupervised clustering with downstream supervised tasks. Law et al. (2004) introduced joint clustering and feature selection via Gaussian mixtures, though with limited classifier flexibility. Hu et al. (2024) provided a recent survey on interpretable clustering, highlighting the disconnect between unsupervised partitioning and predictive modelling.

Other approaches such as Alangari et al. (2023) and Yeganejou and Dick (2019) have proposed probabilistic or fuzzy models with enhanced transparency. However, these are often tied to specific architectures (e.g., GMMs, neural networks), reducing modularity and applicability to lightweight pipelines.

2.4 Positioning of our contribution

CGLFS+ integrates four key capabilities – clustering-based partitioning, local feature selection, adaptive model assignment and robust inference via top-2 voting with fallback – into a unified, end-to-end framework. While prior studies have proposed components of this approach, to our knowledge, CGLFS+ is the first to combine all four elements in a lightweight, interpretable and effective pipeline for tabular classification. Table 1 summarizes how CGLFS+ differentiates itself from existing approaches across core research dimensions.

Table 1. *Positioning of CGLFS+ relative to existing paradigms in heterogeneous tabular learning.*

Research axis	Existing approaches	Contributions of CGLFS+
Modelling granularity	– Global models: GBDT, TabNet, FT-transformer (Grinsztajn et al., 2022; McElfresh et al., 2023) – Local models: clusterwise regression, local linear forests (Kuang and Ooi, 2024; Freidberg et al., 2020)	– Prior local models fix the architecture across clusters – Global models lack flexibility in heterogeneous settings – CGLFS+: modular cluster-based architecture with per-cluster model selection
Clustering integration	– *k*-means, HDBSCAN (MacQueen, 1967; McInnes et al., 2017) – Used mainly for preprocessing or exploratory analysis	– Clustering used as active routing layer for prediction – CGLFS+: unsupervised clustering informs both model assignment and feature selection
Feature selection strategy	– Global MI-based ranking (Peng et al., 2005; Ross, 2014) – Class-/context-aware filtering (Ma an Lu, 2024; Aguilar-Ruiz, 2024)	– CGLFS+: cluster-specific MI filtering with local-global feature fusion (GFF) – Sparse yet informative representations tailored to each subregion

Research axis	Existing approaches	Contributions of CGLFS+
Model selection	– One global model selected via cross-validation (Kohavi, 1995)	– CGLFS+: lightweight adaptive selection per cluster among interpretable learners (LR, DT, SVM)
Inference routing strategy	– Learned gates in MoEs (Lepikhin et al., 2020; Fedus et al., 2022) – Sparse attention or top-*k* routing (Kai et al., 2025)	– CGLFS+: transparent routing via top-2 cluster proximity + global fallback – Enhances stability near cluster boundaries

Recent developments have further emphasized the benefits of combining clustering and modular classification for heterogeneous tabular data. For example, the cluster-based SMOTE boosting ensemble (CSBBoost) utilized cluster structures to improve robustness in imbalanced classification tasks (Salehi and Khedmati, 2024). Similarly, the cluster-based ensemble learning (CBEL) framework for congenital heart disease prediction demonstrated that pre-clustering can enhance both interpretability and adaptability (Kaur and Ahmad, 2024). More recently, the enhanced three-stage cluster-then-classify method (ETSCCM) integrated feature grouping and local modelling to improve predictive reliability in materials data (Yilmaz Eroglu and Guleryuz, 2025). These works share the same underlying philosophy as CGLFS+, highlighting the growing relevance of modular, interpretable and cluster-aware learning paradigms in modern tabular data analysis.

3 PROPOSED METHOD

We propose CGLFS+ (cluster-guided local feature selection with top-2 voting and fallback), a modular framework for classification in heterogeneous tabular data. Instead of relying on a single global model, CGLFS+ partitions the input space into interpretable subregions and trains specialized, lightweight classifiers tailored to each.

The method builds on the premise that relationships between features and target labels can vary across the data space. To exploit this structure, CGLFS+ performs:

- Unsupervised clustering to identify locally coherent groups in the data, with the number of clusters K selected automatically (e.g., via silhouette analysis), independently of the number of classes.
- Feature selection at both global and local levels to identify shared and region-specific predictive signals.
- Hybrid inference via top-2 voting among the nearest clusters, with a fallback to a global model when disagreement occurs.

Each cluster is assigned a compact classifier trained on a fused set of global and cluster-specific features, enhancing both interpretability and local adaptation. At inference, predictions from the two nearest cluster models are compared: if they agree, their output is returned; otherwise, a global fallback model resolves the tie, ensuring robustness near boundaries and in ambiguous regions.

3.1 Overview of CGLFS+ pipeline

Let $D = \{(\mathbf{x}_i, y_i)\}^N_{i=1}$ be the dataset, where $\mathbf{x}_i \in \mathbb{R}^d$ are feature vectors and $y_i \in \{1,\dots,C\}$ are class labels. The dataset is clustered into K partitions using an unsupervised method applied to the input matrix $X = [\mathbf{x}_1,\dots,\mathbf{x}_N]^T$.

CGLFS+ proceeds in four main steps:

- **Step 1: Clustering the data.** We first cluster the training data into K groups (e.g., via k-means). These clusters define regions in the input space where separate models will be trained. The number of clusters K is tuned using internal clustering metrics such as silhouette score.
- **Step 2: Selecting features locally and globally.** To ensure that models remain interpretable and efficient, we select features in two stages:
    - **Global feature selection:** A supervised criterion (e.g., mutual information) is used on the full dataset to select a fixed number of globally relevant features (F^{global}).

- **Local feature selection:** The same criterion is applied within each cluster to identify features that are locally informative (F_k^{local}).

The final feature set for cluster k is the union of the global and local selections: $F_k = F^{global} \cup F_k^{local}$.

- **Step 3: Training local models.** For each cluster, we train a lightweight classifier (e.g., logistic regression, decision tree or SVM) using only the features in F_k. Model selection is done via three-fold cross-validation within the cluster to choose the best-performing algorithm from a predefined pool. If a cluster contains samples from only one class, we store a constant predictor for that class instead of training a model.
- **Step 4: Making predictions (hybrid inference).** To predict the label of a new input $\boldsymbol{x}$:

(1) Compute distances from $\boldsymbol{x}$ to all cluster centroids and identify the two nearest clusters.

(2) Query the local classifiers from both clusters using their respective feature subsets.

(3) If both classifiers agree on the prediction, return that class.

(4) Otherwise, defer to the global fallback model trained on the full dataset.

This approach improves reliability near cluster boundaries and when local models disagree.

The complete procedure is summarized in Algorithm 1.

Algorithm 1. CGLFS+: Local-global feature fusion with top-2 voting and fallback.

Require: Dataset $D = \{(\mathbf{x}_i, y_i)\}$, number of clusters K, feature selector FS, #features m_g (global), m_ℓ (local), candidate models H
Ensure: Local models $\{h_k\}$, feature sets $\{F_k\}$, fallback model h_{global}
Apply k-means clustering to obtain assignments $\{c_i\}$ for K clusters
Select top-m_g global features F^{global} using FS on entire dataset
for each cluster $k = 1,\dots,K$ **do**
    Extract local dataset $D_k = \{(\mathbf{x}_i, y_i) \mid c_i = k\}$
    if D_k contains only one class **then**
        Set h_k as a constant classifier for that class
        Set $F_k \leftarrow$ all features
    else
        Select top-m_ℓ local features F_k^{local} using FS on D_k
        Set $F_k \leftarrow F^{global} \cup F_k^{local}$
        Train all models in H on $D_k[F_k]$ using 3-fold CV
        Assign $h_k \leftarrow$ best-performing model
    end if
end for
Train a global model h_{global} on full dataset with all features
function Predict($\mathbf{x}$)
    Find top-2 nearest clusters $\{k_1, k_2\}$ by centroid distance
    Predict $y_1 = h_{k1}(\mathbf{x}[F_{k1}])$, $y_2 = h_{k2}(\mathbf{x}[F_{k2}])$
    if $y_1 = y_2$ **then return** y_1
    elsereturn $h_{global}(\mathbf{x})$
    end if
end function

4 EXPERIMENT METHODOLOGY DESCRIPTION

4.1 Datasets

To evaluate the performance of our CGLFS+ framework, we experimented with five well-known tabular datasets from a variety of domains, including biology, healthcare, chemistry, computer vision and mechanical systems. These datasets were selected for their diversity in terms of size, number of features, number of classes and class distributions. A summary of their key properties is provided in Table 2.

- **Iris[1] (Dua and Graff, 2019)**: This classic dataset contains 150 samples of iris flowers, each described by four measurements of sepal and petal size. The goal is to classify the flowers into one of three species: *Iris setosa*, *I. versicolor* or *I. virginica*.
- **Breast cancer Wisconsin (diagnostic)[2] (Dua and Graff, 2019)**: This medical dataset includes 569 samples from breast tumour biopsies. Each sample is characterized by 30 numerical features extracted from cell images and labelled as benign or malignant.
- **Red wine quality[3] (Cortez et al., 2009)**: Contains 1599 red wine samples from Portugal, with 11 physicochemical features per sample. The target variable is a quality score from 3 to 8, making this a multi-class and highly imbalanced problem, as most wines are rated 5 or 6.
- **Digits (Scikit-learn)[4] (Scikit-learn developpers, 2025)**: A smaller, easier-to-use version of MNIST with 1797 images of handwritten digits. Each image is 8×8 pixels and represented as a vector of 64 features. The task is to classify digits from 0 to 9.
- **Vehicle silhouettes[5] (Dua and Graff, 2019)**: This dataset includes 846 examples of vehicle shapes described by 18 geometric features. The goal is to identify the type of vehicle: Opel, Saab, Bus or Van.

Table 2. *Summary of datasets used in the evaluation.*

Dataset	Samples	Features	Classes	Class balance
Iris	150	4	3	Balanced
Breast cancer (WDBC)	569	30	2	Slightly imbalanced (62% benign)
Red wine quality	1599	11	6*	Highly imbalanced (mostly 5–6)
Digits	1797	64	10	Balanced
Vehicle silhouettes	846	18	4	Moderately imbalanced

*Note: Labels range from 3 to 8, giving six distinct classes.

4.2 Evaluation metrics

To evaluate both the predictive performance and practical efficiency of our framework, we report three key metrics:

- **Accuracy:** The proportion of correctly classified examples in the test set. While widely used, it can be misleading when classes are imbalanced.
- **Macro F1-score:** The average F1-score computed independently for each class, then averaged across all classes. This means that all classes are treated equally, regardless of their size, which makes it well-suited for imbalanced multi-class datasets. The F1-score itself balances precision and recall.
- **Inference time:** The average time (in seconds) that the model takes to predict the full test set, averaged over several runs. It reflects how efficient and scalable the method is in practice.

4.3 Baselines

To assess the performance of our approach, we compare it against a set of standard global classifiers, all trained on the full feature space without clustering or local feature selection. These baselines represent common choices in tabular classification tasks:

- **Logistic regression (LR):** A simple and interpretable linear model with L2 regularization. The regularization strength is selected via five-fold cross-validation using the liblinear solver.

[1] See, https://archive.ics.uci.edu/ml/datasets/iris

[2] See, https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

[3] See, https://archive.ics.uci.edu/ml/datasets/Wine+Quality

[4] See, https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_digits.html

[5] See, https://archive.ics.uci.edu/ml/datasets/Statlog+(Vehicle+Silhouettes).

- **Support vector machine (SVM):** A non-linear classifier using an RBF kernel. The parameters γ and C are tuned through nested five-fold cross-validation on the training set.
- ***K*-nearest neighbours (KNN):** A distance-based method with k = 3, uniform weighting and Euclidean distance.
- **Decision tree (DT):** A rule-based classifier using Gini impurity, with a maximum depth of 10 unless specified otherwise.
- **Random forest (RF):** An ensemble of 100 decision trees trained on bootstrapped samples with feature subsampling (max features set to sqrt). Default Scikit-learn settings are used unless tuning is applied.
- **Gradient boosting (GB):** A boosted tree model with 100 estimators and a learning rate of 0.1, implemented via the gradient boosting classifier. Early stopping is performed using a 5% validation split.
- **Naive Bayes (NB):** A generative classifier assuming Gaussian distributions for each feature.

All models are evaluated on identical training splits, with hyperparameters optimized via five-fold cross-validation to ensure fairness.

4.4 Proposed method configuration (CGLFS+)

This section details the configuration choices underlying the implementation of CGLFS+ (cluster-guided local feature selection with top-2 voting and fallback):

- **Clustering:** The training set is partitioned into K clusters using k-means. The number of clusters is determined automatically based on the silhouette score, selecting the value of K that maximizes intra-cluster cohesion and inter-cluster separation. This allows the model to adapt to hidden substructures in the data, independently of the number of classes. The only exception is the digits dataset, where K is fixed at 10 to match the number of digit classes and avoid unstable clustering.
- **Local feature selection:** Within each cluster, features are ranked by their mutual information (MI) with class labels. Rather than fixing the number of features, we retain those whose MI exceeds the cluster-specific average, allowing the feature set to adapt to the local structure (Peng et al., 2005).
- **Global feature fusion:** Globally informative features (those with MI above the global average) are added to the local subset of each cluster. This balances local specialization with shared context and avoids arbitrary feature limits. The selection is based on information-theoretic criteria such as max-relevance (Peng et al., 2005).
- **Candidate models:** For each cluster, we perform model selection among three lightweight classifiers:
    - **Logistic regression (LR):** Linear, with L2 regularization and the liblinear solver.
    - **Decision tree (DT):** A shallow tree with a maximum depth of 3, offering concise decision rules.
    - **SVM (RBF kernel):** Non-linear, suitable for clusters with complex boundaries.
- **Model selection:** Classifiers are evaluated using three-fold stratified cross-validation within each cluster. The model with the highest average accuracy is retrained on the full cluster data and retained.
- **Prediction:** At test time, each instance is assigned to its two nearest clusters (via Euclidean distance). If both local models agree, the prediction is returned. Otherwise, a fallback prediction is provided by a global model.
- **Fallback model:** We use a global random forest with 100 estimators trained on the full dataset and all features to resolve prediction conflicts. The fallback mechanism is activated only for a small fraction of samples located near ambiguous cluster boundaries, where the top-2 local models disagree. This selective use prevents overreliance on the global model while ensuring consistent decisions across heterogeneous subspaces. Random forests were chosen because they provide robust generalization and enable feature importance inspection, offering a practical balance between stability and interpretability. As a result, the fallback component reinforces prediction reliability without undermining the local transparency of the CGLFS+ architecture.

All inputs are standardized (zero mean, unit variance). Evaluation is performed using five-fold cross-validation, repeated over five random splits. Inference time is reported as the average prediction time per test instance, excluding training.

Choice of candidate models. We select three candidate classifiers for local model selection: logistic regression (LR), decision tree (DT) and support vector machine with RBF kernel (SVM-RBF). This choice is motivated by the following factors:

- **Model diversity:** The three models offer complementary inductive biases: linear (LR), rule-based (DT) and non-linear kernel-based (SVM), which allows adaptation to various local patterns (Domingos, 2012).
- **Interpretability:** LR and DT are inherently transparent and well-suited for locally interpretable modelling, unlike ensemble or deep models (Molnar, 2020; Rudin, 2019).
- **Efficiency:** All selected models are lightweight and fast to train, which is important when fitting multiple models independently in each cluster (Pedregosa et al., 2011).
- **Robustness to small data:** Simpler models generalize better in low-data regimes, which often occur in small clusters (Hastie et al., 2009).
- **Proven effectiveness:** LR, DT and SVM have shown competitive performance on tabular datasets in empirical studies, particularly when interpretability is a priority (Grinsztajn et al., 2022; Fernández-Delgado et al., 2014).

We exclude the following model families for practical reasons:

- **Ensemble methods (e.g., random forest, GBDT):** High accuracy but low interpretability and higher inference cost (Lundberg et al., 2020).
- ***K*-nearest neighbours (KNN):** Sensitive to feature scaling and inefficient at inference time (Peterson, 2009).
- **Neural networks:** Require large amounts of data and offer limited interpretability in small tabular clusters (Shwartz-Ziv and Armon, 2022).

This configuration offers a balanced trade-off between model variety, transparency, efficiency and generalization capability. In summary, this model set offers a good balance of diversity, interpretability, efficiency and generalization, making it well-suited for adaptive classification in locally clustered tabular data.

5 RESULTS AND DISCUSSION

We evaluate the performance of the proposed method CGLFS+ on five diverse tabular datasets: iris, breast cancer, digits, red wine quality and vehicle (see subsection 4.1). All experiments are conducted using five-fold stratified cross-validation to ensure balanced and reliable results. For each dataset and model, we report the mean and standard deviation of the following metrics: accuracy, macro-averaged F1-score and inference time (in seconds).

Detailed results are presented in Tables 3 to 7. For clarity, we report the following variants of our method:

- **CGLFS:** base model without global feature fusion or top-2 voting + fallback;
- **CGLFS+GFF:** CGLFS extended with global feature fusion only;
- **CGLFS+**: full model with both enhancements.

Table 3. *Performance on iris dataset.*

Model	Accuracy (%)	F1-score (%)	Time (s)
Logistic regression	95.32 ± 1.52	95.32 ± 1.53	0.0138
SVM (RBF)	95.35 ± 1.52	95.37 ± 1.53	0.0094
KNN (k=3)	94.00 ± 2.46	94.01 ± 2.44	0.0117
Random forest	94.67 ± 2.67	94.64 ± 2.68	0.2157
Gradient boosting	95.33 ± 2.40	95.31 ± 2.41	0.3612
Decision tree	95.36 ± 2.41	95.31 ± 2.43	0.0031
Naive Bayes	94.67 ± 1.00	94.65 ± 1.01	0.0031
CGLFS (ours)	96.14 ± 1.15	95.91 ± 1.21	0.2412
CGLFS+GFF (ours)	96.91 ± 1.75	96.01 ± 1.13	0.2612
CGLFS+ (ours)	97.03 ± 1.13	96.82 ± 1.16	0.3127

Table 4. *Performance on breast cancer dataset.*

Model	Accuracy (%)	F1-score (%)	Time (s)
Logistic regression	97.37 ± 0.66	97.14 ± 0.83	0.0138
SVM (RBF)	97.54 ± 0.95	97.36 ± 1.09	0.0466
KNN (k=3)	96.83 ± 0.54	96.55 ± 0.70	0.0197
Random forest	95.61 ± 0.23	95.29 ± 0.35	0.3132
Gradient boosting	95.26 ± 1.26	94.86 ± 1.49	0.6269
Decision tree	91.04 ± 1.79	90.28 ± 2.17	0.0122
Naive Bayes	92.97 ± 0.99	92.44 ± 1.08	0.0040
CGLFS (ours)	98.07 ± 0.47	97.67 ± 0.21	0.5380
CGLFS+GFF (ours)	98.15 ± 0.25	97.86 ± 0.75	0.5712
CGLFS+ (ours)	98.71 ± 0.48	97.91 ± 1.21	0.6022

Table 5. *Performance on digits dataset.*

Model	Accuracy (%)	F1-score (%)	Time (s)
Logistic regression	97.11 ± 0.38	97.10 ± 0.37	0.1971
SVM (RBF)	98.39 ± 0.60	98.39 ± 0.60	0.9353
KNN (k=3)	97.50 ± 0.68	97.49 ± 0.69	0.0587
Random forest	97.61 ± 0.38	97.60 ± 0.38	0.5751
Gradient boosting	96.38 ± 0.50	96.39 ± 0.49	12.9223
Decision tree	85.64 ± 1.49	85.64 ± 1.51	0.0583
Naive Bayes	78.30 ± 1.80	78.02 ± 1.99	0.0161
CGLFS (ours)	98.16 ± 0.67	98.12 ± 0.58	5.1730
CGLFS+GFF (ours)	98.17 ± 0.53	98.12 ± 0.61	5.8512
CGLFS+ (ours)	98.51 ± 1.07	98.41 ± 1.19	6.0005

Table 6. *Performance on wine quality dataset.*

Model	Accuracy (%)	F1-score (%)	Time (s)
Logistic regression	59.91 ± 1.37	28.43 ± 1.88	0.2680
SVM (RBF)	62.54 ± 2.21	28.93 ± 1.54	1.1983
KNN (k=3)	57.66 ± 2.63	31.08 ± 3.76	0.0587
Random forest	69.92 ± 1.97	36.62 ± 3.88	0.5926
Gradient boosting	65.48 ± 2.23	35.10 ± 3.12	3.2421
Decision tree	61.85 ± 0.38	34.99 ± 1.93	0.0202
Naive Bayes	54.91 ± 1.95	31.86 ± 3.32	0.0061
CGLFS (ours)	71.03 ± 1.67	40.12 ± 1.98	1.8710
CGLFS+GFF (ours)	71.47 ± 1.26	40.91 ± 1.78	1.9512
CGLFS+ (ours)	72.14 ± 1.15	40.92 ± 1.21	2.1023

Table 7. *Performance on vehicle dataset.*

Model	Accuracy (%)	F1-score (%)	Time (s)
Logistic regression	78.73 ± 1.68	78.56 ± 1.78	0.1159
SVM (RBF)	75.42 ± 2.29	74.82 ± 2.41	0.2725
KNN (k=3)	71.40 ± 1.31	71.10 ± 1.47	0.0322
Random forest	73.17 ± 2.38	72.67 ± 2.69	0.3984
Gradient boosting	76.60 ± 1.89	76.52 ± 1.70	1.7268

Model	Accuracy (%)	F1-score (%)	Time (s)
Decision tree	69.51 ± 2.40	69.77 ± 2.31	0.0125
Naive Bayes	45.86 ± 3.32	42.76 ± 3.01	0.0076
CGLFS (ours)	78.91 ± 2.13	78.83 ± 2.98	0.7730
CGLFS+GFF (ours)	80.37 ± 2.13	79.12 ± 2.98	0.8743
CGLFS+ (ours)	82.14 ± 2.17	81.91 ± 2.31	0.9112

5.1 Overall performance

CGLFS+ consistently achieves the best or near-best performance across all five benchmark datasets. These results confirm the effectiveness of combining local modelling, adaptive feature selection and robust inference.

- **Iris:** CGLFS+ reaches **97.03%** accuracy and **96.82%** macro F1-score, outperforming strong baselines such as SVM (95.35%) and gradient boosting (95.33%) as shown in Table 3.
- **Breast cancer:** On this binary classification task with moderate class imbalance and nonlinear relationships, CGLFS+ achieves **98.71%** accuracy and **97.91%** F1-score (Table 4), reflecting its ability to adapt to clustered feature structures.
- **Digits:** Despite overall strong performance from most models, CGLFS+ delivers the top results with **98.51%** accuracy and **98.41%** F1-score (Table 5).
- **Red wine quality:** This multi-class, imbalanced dataset poses a greater challenge. CGLFS+ achieves **72.14%** accuracy and **40.92%** F1-score, significantly outperforming the next-best model (random forest at 69.92%) (Table 6).
- **Vehicle:** For this balanced, four-class dataset with subtle class differences, CGLFS+ sets a new performance benchmark with **82.14%** accuracy and **81.91%** F1-score (Table 7).

5.2 Ablation insights

To better understand the contribution of each component, we compare three variants:

- **CGLFS:** Base model without global feature fusion or top-2 voting + fallback;
- **CGLFS+GFF:** Adds global feature fusion, enriching the feature space of each cluster with dataset-wide signals;
- **CGLFS+:** Full model combining both GFF and top-2 voting with fallback.

Global feature fusion provides consistent improvements, while top-2 voting with fallback helps resolve ambiguous predictions near cluster boundaries; particularly evident in the wine and vehicle datasets.

5.3 Computational efficiency

Despite being modular, CGLFS+ remains practical in terms of inference time. For instance, on the digits dataset, it completes predictions in **6.00 s**, which is faster than gradient boosting (**12.92 s**) and comparable to other traditional models. This shows that the framework achieves a favourable balance between accuracy and scalability.

While the modular design introduces a modest increase in training complexity compared to simple classifiers, this cost is offset by enhanced predictive stability and adaptability in heterogeneous datasets. Moreover, because each local model operates independently within its assigned cluster, the framework naturally supports parallel training, resulting in near-linear scalability with respect to data size. These characteristics make the trade-off between computational costs and predictive benefits favourable in most realistic tabular settings.

5.4 Visual summary

To complement the quantitative tables, Figures 1 to 3 provide a visual overview of average performance across all the datasets.

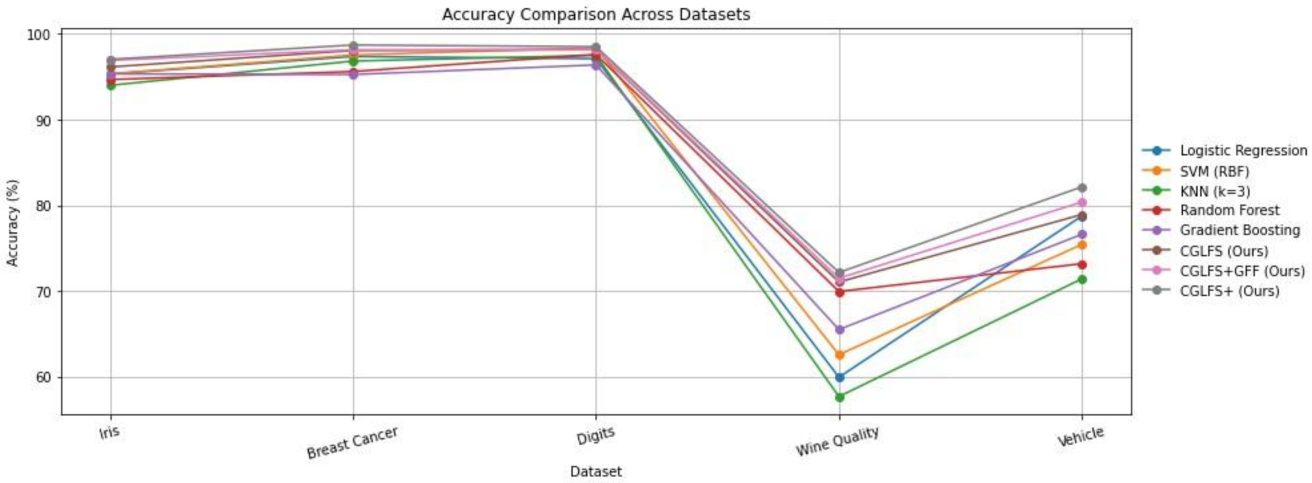

Figure 1. *Accuracy comparison across datasets.*

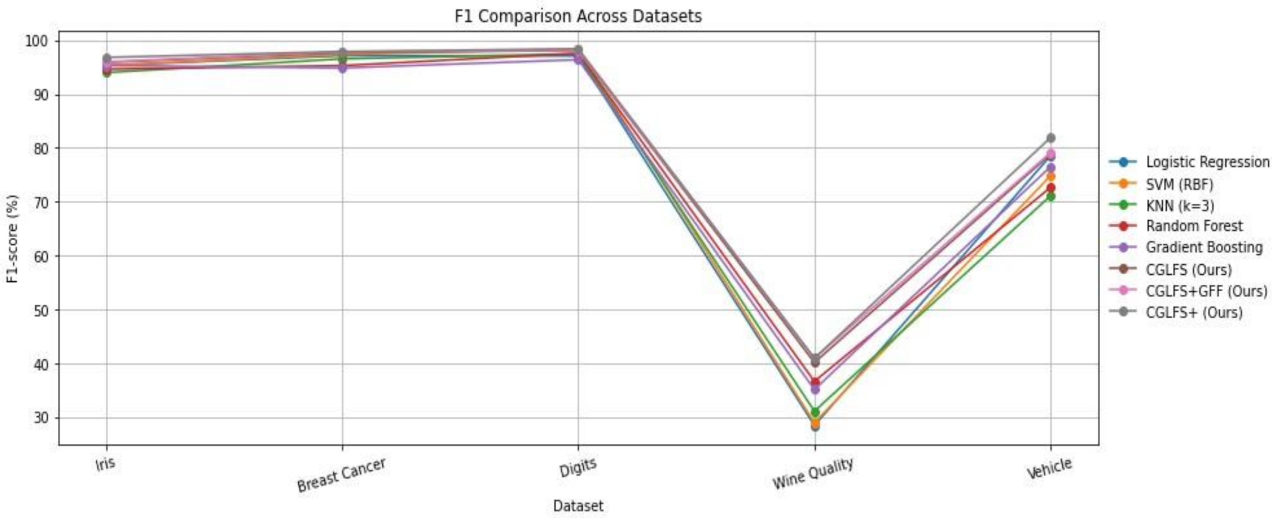

Figure 2. *F1-score comparison across datasets.*

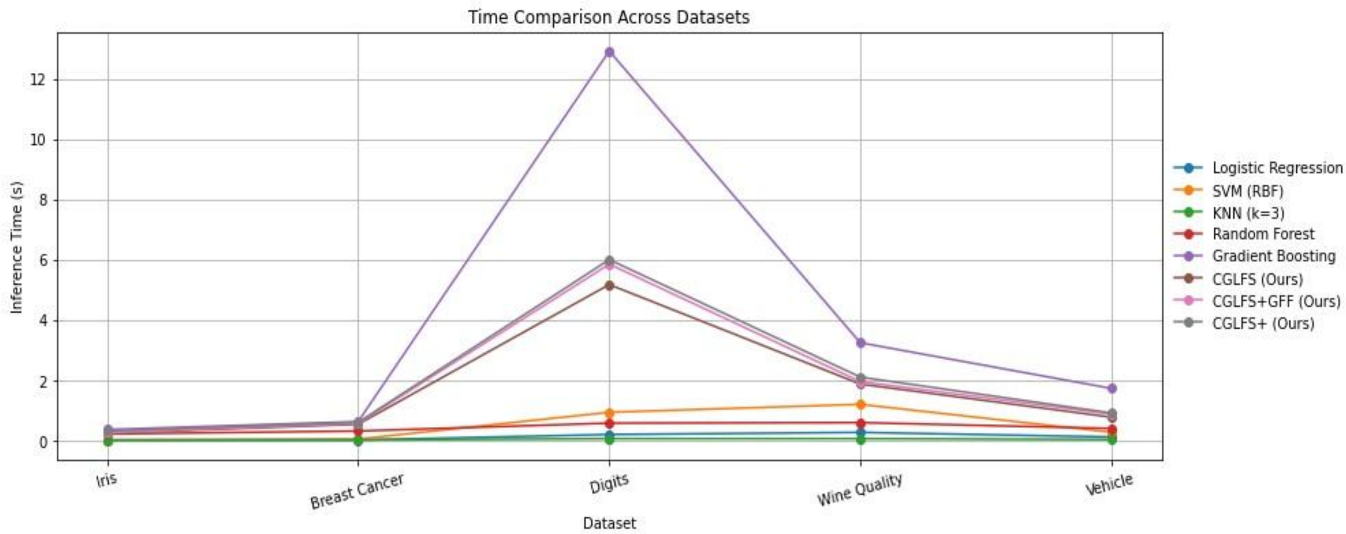

Figure 3. *Inference time comparison across datasets.*

CGLFS+ consistently ranks at or near the top in accuracy and macro F1. The most notable gains appear on structurally complex datasets such as wine and vehicle. While inference time is slightly higher than simpler models, CGLFS+ remains significantly more efficient than ensemble-based baselines.

The experimental results clearly demonstrate that CGLFS+ offers a robust and interpretable alternative to standard classifiers for heterogeneous tabular data. Its performance, modularity and scalability make it a compelling choice for real-world applications.

5.5 Statistical significance testing

To assess whether the observed performance differences among models are statistically significant, we conducted the Friedman test (Demšar, 2006) using accuracy values averaged over five-fold cross-validation on five benchmark datasets.

The Friedman test returned a chi-squared statistic of $\chi^2_F = 48.21$ with a corresponding p-value of 3.19×10^{-8}, allowing us to reject the null hypothesis of equal model performance at the $\alpha = 0.05$ level. This confirms that at least one model exhibits significantly different behaviour compared to the others. To further investigate the source of these differences, we applied the Nemenyi post-hoc test, which compares all pairs of models. The critical difference (CD) was calculated as $CD = 2.05$, based on $k = 8$ models and $N = 5$ datasets, using $q_{0.05} = 3.314$ from the studentized range distribution. The resulting average ranks for all models are presented in Table 8.

Table 8. *Average ranks across datasets (lower is better).*

Model	Average rank
CGLFS+	1.00
SVM (RBF)	3.00
Logistic regression	4.20
Random forest	4.30
Gradient boosting	4.40
KNN (k=3)	5.80
Decision tree	5.80
Naive Bayes	7.50

The results of the Nemenyi test are visualized in Figure 4, which shows the average ranks and the critical difference interval. Models not connected by a horizontal line differ significantly at the $p < 0.05$ level.

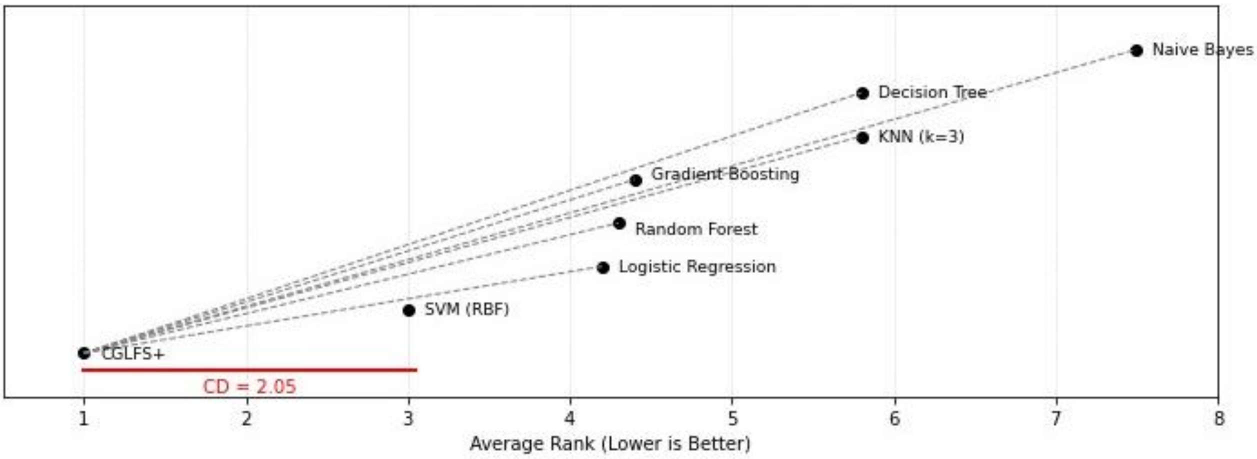

Figure 4. *Critical difference (CD) diagram comparing CGLFS+ to baseline models. Lower ranks indicate better performance. Models not connected to CGLFS+ are significantly different.*

As shown, CGLFS+ achieves the best average rank across all the datasets and is statistically superior to all other models except SVM (RBF), which lies just within the critical threshold. These findings validate the consistent advantage and robustness of our proposed framework.

5.6 Qualitative evaluation of interpretability

To complement the quantitative evaluation, a qualitative inspection of the locally selected features was performed for representative datasets.

Table 9. *Representative top 5 locally selected features across clusters.*

Dataset	Cluster ID	Top 5 selected features
Breast cancer	C1	texture_mean, symmetry_mean, compactness_mean, smoothness_mean, fractal dimension_mean
Breast cancer	C2	texture_mean, symmetry_mean, fractal dimension_mean, radius_mean, compactness_mean
Red wine quality	C1	alcohol, sulphates, fixed_acidity, chlorides, pH
Red wine quality	C2	density, residual sugarvolatile_acidity, sulphates, pH

As shown in Table 9, the breast cancer dataset reveals consistent emphasis on texture_mean, symmetry_mean and compactness_mean, which are well-established diagnostic indicators in recent explainable AI models for breast cancer prediction (Kalangi et al., 2025). For the red wine quality dataset, clusters C1 and C2 respectively emphasize alcohol, volatile acidity, density and pH, aligning with recent findings that these physicochemical properties are major determinants of perceived wine quality (Luque Sendra et al., 2023). These results confirm that the local feature selection mechanism in CGLFS+ identifies domain-relevant and interpretable features, reinforcing both its explainability and practical validity.

5.7 Discussion and limitations

The performance of CGLFS+ across all the datasets demonstrates the benefits of combining local modelling with structured decision fusion. Unlike traditional global classifiers, CGLFS+ captures fine-grained patterns by tailoring feature selection and decision boundaries to each cluster. This localized specialization enables the model to adapt more effectively to heterogeneity in tabular data, contributing to its consistently top-ranked performance.

Moreover, the design of CGLFS+ offers interpretability by exposing the structure of cluster-specific models. This can be useful in domains where understanding local decision logic is essential. However, our current evaluation remains focused on accuracy and efficiency, without formal analysis of interpretability or feature consistency across clusters.

Although the datasets employed in this study are of moderate size, they span diverse domains, including biology, healthcare, chemistry, computer vision and mechanical systems, providing a broad basis for evaluating heterogeneous tabular structures. This diversity ensures fair comparison with existing benchmarks while maintaining controlled experimental conditions. Importantly, the modular architecture of CGLFS+ allows efficient parallelization and linear scalability with respect to data size, making the framework readily applicable to larger real-world datasets.

In line with recent advances in cluster-based ensemble learning (Salehi and Khedmati, 2024; Kaur and Ahmad, 2024; Yilmaz Eroglu and Guleryuz, 2025), CGLFS+ adopts a modular approach that emphasizes localized feature selection and decision fusion. This design is consistent with trends observed in the literature from the period 2024–2025, further supporting the relevance of the framework for heterogeneous tabular domains.

That said, there are a few important limitations to consider:

- **Dependency on clustering:** The approach assumes that meaningful substructures exist in the data. If the clustering fails to reflect actual task relevant divisions, the overall model may degrade.

- **Alternative clustering strategies:** While CGLFS+ currently employs K-means for partitioning, alternative clustering algorithms could potentially enhance robustness when data exhibit complex or non-linear substructures. Density-based methods such as ***HDBSCAN*** can adaptively detect irregularly shaped clusters and handle noise, while ***spectral clustering*** and ***Gaussian mixture models (GMMs)*** may better capture overlapping or manifold-structured regions. However, these approaches often involve higher computational costs and additional hyperparameters, which could reduce the simplicity and reproducibility of the present framework. Exploring these alternatives and their impact on local model stability and interpretability represents an interesting direction for future research.
- **Added complexity:** In simple or linearly separable datasets, the overhead introduced by clustering and local modelling may not be justified when compared to efficient global models. Nevertheless, these additional computational costs are most warranted in domains where feature–label relationships vary across subpopulations, such as healthcare diagnostics, chemistry-based toxicity prediction or credit scoring, where local specialization provides tangible benefits in both interpretability and predictive reliability. In contrast, for simple and homogeneous datasets, global classifiers may remain a more computationally efficient alternative.

6 CONCLUSION AND FUTURE WORK

In this work, we proposed CGLFS+, a modular classification framework that integrates cluster-guided local modelling with global decision fusion. The method balances interpretability, flexibility and predictive power, and achieved strong performance across five diverse tabular datasets. It significantly outperformed widely used classifiers, as confirmed by statistical testing.

By separating the learning process into interpretable local units and enhancing decision consistency through global fusion and fallback voting, CGLFS+ introduces a promising direction for modelling structured data. For future work, we plan to explore the following directions:

- **Adaptive clustering:** Investigate automatic methods for choosing the number and granularity of clusters based on data complexity.
- **Interpretability validation:** Design human-centred experiments to evaluate how local models and selected features are perceived and understood by domain experts.
- **Scalability and deployment:** Extend the framework to handle large-scale datasets and real-time prediction scenarios, particularly in applications such as finance or medical decision support.

Overall, CGLFS+ offers a practical and interpretable alternative to complex black-box classifiers, especially for structured data scenarios where both accuracy and understanding matter.

ADDITIONAL INFORMATION AND DECLARATIONS

Conflict of Interests: The author declares no conflict of interest.

Author Contributions: The author confirms being the sole contributor of this work.

Statement on the Use of Artificial Intelligence Tools: The author states that no generative artificial intelligence tools were used to create the textual or visual content of this article. Artificial intelligence-based language assistance was used only for limited linguistic revisions, such as grammar and spelling corrections of text written by the author. All concepts, methodologies, experiments, analyses, and interpretations presented in this article were developed solely by the author.

Data Availability: The data that support the findings of this study are available from the corresponding author.

REFERENCES

Aguilar-Ruiz, J. S. (2024). Class-specific feature selection for classification explainability. arXiv:2411.01204. https://doi.org/10.48550/arXiv.2411.01204

Alangari, N., Menai, M. E. B., Mathkour, H., & Almosallam, I. (2023). Intrinsically interpretable gaussian mixture model. *Information*, 14(3), 164. https://doi.org/10.3390/info14030164

Cai, W., Jiang, J., Wang, F., Tang, J., Kim, S., & Huang, J. (2025). A survey on mixture of experts in large language models. *IEEE Transactions on Knowledge and Data Engineering*, 37(7), 3896–3915. https://doi.org/10.1109/TKDE.2025.3554028

Cortez, P., Cerdeira, A., Almeida, F., Matos, T., & Reis, J. (2009). Modeling wine preferences by data mining from physicochemical properties. *Decision support systems*, 47(4), 547–553. https://doi.org/10.1016/j.dss.2009.05.016

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of machine learning research*, 7, 1–30.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87. https://doi.org/10.1145/2347736.2347755

Dua, D. & Graff, C. (2019). UCI machine learning repository. https://archive.ics.uci.edu/ml/index.php

Fedus, W., Zoph, B., & Shazeer, N. (2022). Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity. *Journal of Machine Learning Research*, 23, 1–39.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. *Journal of machine learning research*, 15(1), 3133–3181.

Friedberg, R., Tibshirani, J., Athey, S., & Wager, S. (2020). Local linear forests. *Journal of Computational and Graphical Statistics*, 30(2), 503–517. https://doi.org/10.1080/10618600.2020.1831930

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data?. In *36th Conference on Neural Information Processing Systems*, (pp. 507–520). NIPS.

Hancer, E., Xue, B., & Zhang, M. (2020). A survey on feature selection approaches for clustering. *Artificial intelligence review*, 53(6), 4519–4545. https://doi.org/10.1007/s10462-019-09800-w

Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The Element of Statistical Learning: Data Mining, Inference, and Prediction*. Springer. https://doi.org/10.1007/978-0-387-84858-7

Hu, L., Jiang, M., Dong, J., Liu, X., & He, Z. (2024). Interpretable clustering: A survey. arXiv:2409.00743. https://doi.org/10.48550/arXiv.2409.00743

Ismail, A. A., Arik, S. Ö., Yoon, J., Taly, A., Feizi, S., & Pfister, T. (2022). Interpretable mixture of experts. arXiv:2206.02107. https://doi.org/10.48550/arXiv.2206.02107

Kalangi, P. K., Rachuri, G., Saleem, D., Chandana, P., Goud, B. P., & Kumar, S. V. (2025). A Hybrid Approach to Accurate Breast Cancer Prediction Integrating: Explainable AI and Machine Learning. In *2025 5th International Conference on Intelligent Technologies (CONIT)* (pp. 1–7). IEEE. https://doi.org/10.1109/CONIT65521.2025.11167733

Kaur, I., & Ahmad, T. (2024). A cluster-based ensemble approach for congenital heart disease prediction. *Computer Methods and Programs in Biomedicine*, 243, 107922. https://doi.org/10.1016/j.cmpb.2023.107922

Kheradpisheh, S. R., Sharifizadeh, F., Nowzari-Dalini, A., Ganjtabesh, M., & Ebrahimpour, R. (2014). Mixture of feature specified experts. *Information Fusion*, 20, 242–251. https://doi.org/10.1016/j.inffus.2014.02.006

Kim, B., Shah, J. A., & Doshi-Velez, F. (2015). Mind the gap: A generative approach to interpretable feature selection and extraction. In *28th Conference on Neural Information Processing Systems,* (pp. 1–9). NIPS.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (II),* (pp. 1137–1145). IJCAI.

Kuang, Y. C., & Ooi, M. (2024). Performance Characterization of Clusterwise Linear Regression Algorithms. *Wiley Interdisciplinary Reviews: Computational Statistics*, 16(5), e70004. https://doi.org/10.1002/wics.70004

Law, M. H., Figueiredo, M. A., & Jain, A. K. (2004). Simultaneous feature selection and clustering using mixture models. *IEEE transactions on pattern analysis and machine intelligence*, 26(9), 1154–1166. https://doi.org/10.1109/TPAMI.2004.71

Lepikhin, D., Lee, H., Xu, Y., Chen, D., Firat, O., Huang, Y., ... & Chen, Z. (2020). Gshard: Scaling giant models with conditional computation and automatic sharding. arXiv:2006.16668. https://doi.org/10.48550/arXiv.2006.16668

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence*, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Luque, A., Mazzoleni, M., Zamora-Polo, F., Ferramosca, A., Lama, J. R., & Previdi, F. (2023). Determining the importance of physicochemical properties in the perceived quality of wines. *IEEE Access*, 11, 115430–115449. https://doi.org/10.1109/access.2023.3325676

Ma, X. A., & Lu, K. (2024). Class-specific feature selection using neighborhood mutual information with relevance-redundancy weight. *Knowledge-Based Systems*, 300, 112212. https://doi.org/10.1016/j.knosys.2024.112212

MacQueen, J. (1965). Some methods for classification and analysis of multivariate observations [C]. In *Proc. of Berkeley Symposium on Mathematical Statistics & Probability,* (pp. 281–297). University of California Press.

McElfresh, D., Khandagale, S., Valverde, J., Prasad C, V., Ramakrishnan, G., Goldblum, M., & White, C. (2023). When do neural nets outperform boosted trees on tabular data?. In *Proceedings of the 37th International Conference on Neural Information Processing System,* (pp. 76336–76369). NIPS.

McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11), 205. https://doi.org/10.21105/joss.00205

Molnar, C. (2020). Interpretable machine learning. https://christophm.github.io/interpretable-ml-book/

Oyamada, M., & Nakadai, S. (2017). Relational mixture of experts: Explainable demographics prediction with behavioral data. In *2017 IEEE International Conference on Data Mining (ICDM)* (pp. 357–366). IEEE. https://doi.org/10.1109/ICDM.2017.45

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12, 2825–2830.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8), 1226–1238. https://doi.org/10.1109/TPAMI.2005.159

Peralta, B., & Soto, A. (2014). Embedded local feature selection within mixture of experts. *Information Sciences*, 269, 176–187. https://doi.org/10.1016/j.ins.2014.01.008

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2), 1883. https://doi.org/10.4249/scholarpedia.1883

Ross, B. C. (2014). Mutual information between discrete and continuous data sets. *PloS one*, 9(2), e87357. https://doi.org/10.1371/journal.pone.0087357

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Salehi, A. R., & Khedmati, M. (2024). A cluster-based SMOTE both-sampling (CSBBoost) ensemble algorithm for classifying imbalanced data. *Scientific Reports*, 14(1), 5152. https://doi.org/10.1038/s41598-024-55598-1

Scikit-learn developers. (2025). Digits dataset — scikit-learn documentation. https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_digits.html

Shi, Y., Zeng, H., Gong, X., Cai, L., Xiang, W., Lin, Q., Zheng, H., & Zhu, J. (2025). Consensus Guided Multi-View Unsupervised Feature Selection with Hybrid Regularization. *Applied Sciences*, 15(12), 6884. https://doi.org/10.3390/app15126884

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81, 84–90. https://doi.org/10.1016/j.inffus.2021.11.011

Yeganejou, M., & Dick, S. (2019). Improved deep fuzzy clustering for accurate and interpretable classifiers. In *2019 IEEE international conference on fuzzy systems (FUZZ-IEEE)* (pp. 1–7). IEEE. https://doi.org/10.1109/FUZZ-IEEE.2019.8858809

Yilmaz Eroglu, D., & Guleryuz, E. (2025). Enhanced Three-Stage Cluster-Then-Classify Method (ETSCCM). *Metals*, 15(3), 318. https://doi.org/10.3390/met15030318

Yuksel, S. E., Wilson, J. N., & Gader, P. D. (2012). Twenty years of mixture of experts. *IEEE transactions on neural networks and learning systems*, 23(8), 1177–1193. https://doi.org/10.1109/TNNLS.2012.2200299

Acta Informatica Pragensia is published by the Prague University of Economics and Business, Czech Republic | eISSN: 1805-4951