Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Dimensionality Reduction via Hierarchical Analysis of Partially Ordered Structures: LAPOS

Elsayed Badr[1,2,3*], Luai Ibrahim Alharbi[4], Moaaz Badr[5], Ahmed Zaghlol[3], Ahmed Hagag[3]

[1] Department of Information Systems, College of Information Technology, Misr University for Science and Technology (MUST), P.O. BOX 77, Giza, Egypt
[2] Department of Scientific Computing, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt
[3] The Egyptian School of Data Science (ESDS), Benha, Egypt
[4] Department of Science and Artificial Intelligence, Faculty of Information Technology, Monash University, Clayton Victoria, 3800, Australia
[5] Department of Civil Engineering, Higher Institute of Engineering and Technology in Mansoura, Mansoura, Egypt
E-mail: alsayed.badr@must.edu.eg

Received: 26 September 2025; **Revised:** 10 November 2025; **Accepted:** 25 November 2025

Abstract: Domains such as artificial intelligence, statistical modeling, and signal interpretation. This method aims to retain the core characteristics of datasets while minimizing their dimensional scope, which contributes to improving the efficiency of algorithms and enhancing data understanding. The proposed framework relies on the principle of Level Analysis of Partially Ordered Sets (LAPOS). Instead of treating dimensions as independent entities or searching for linear projections, this method considers the ordinal relationships between different dimensions. A partial order structure is constructed that reflects the correlations and interactions between data attributes, allowing for a more accurate identification of the most influential and frequently occurring dimensions. LAPOS can reveal nonlinear relationships, better interpretability, and flexibility in dimension selection. Preliminary results have shown that this method surpasses traditional dimensionality reduction approaches in maintaining data integrity and minimizing information loss in subsequent tasks (such as classification and clustering). LAPOS, Principal Component Analysis (PCA), and Factor Analysis (FA) achieve 92.59, 86.1, and 81.02 accuracy respectively, when employing the Support Vector Machine (SVM) algorithm. This research opens new avenues for using partially ordered set theory to manage the complications arising from high-dimensional feature spaces.

***Keywords*:** dimensionality reduction, partially ordered sets, artificial intelligence models

MSC: 06A06, 68Q25, 68W40

1. Introduction

High-dimensional datasets, which are frequently encountered in practical and applied fields such as industrial processes and medical diagnostics, typically encompass a vast number of features, many of which are either redundant or irrelevant to the task under investigation. These redundant and irrelevant attributes add noise to the dataset, thereby complicating the learning process for machine learning algorithms and diminishing their predictive accuracy and

DOI: https://doi.org/10.37256/cm.7120268736

computational efficiency. Consequently, there arises a pressing need to determine and preserve only the most valuable and informative attributes, while excluding those that do not meaningfully support the classification or prediction objectives. Feature Selection (FS) emerges as a robust and effective strategy designed to address this dimensionality reduction challenge by isolating the specific subset of attributes possessing the highest predictive relevance, while eliminating redundant and irrelevant factors [1]. Lowering the dataset's dimensionality improves the efficiency of the learning process and decreases the likelihood of overfitting, resulting in better generalization performance for machine learning models. Once the essential features are identified and retained, various classification algorithms can be applied to perform the predictive tasks efficiently using the refined feature space. It is widely acknowledged in the literature that the problem of FS can be expressed as an optimization task, where the central objective is to enhance model performance, typically evaluated through accuracy, by selecting the optimal subset of attributes. This setup enables embedding feature selection into optimization schemes, with classification accuracy or any suitable performance metric—serving as the fitness function guiding the optimization process [2]. Feature Selection possesses demonstrated effectiveness across numerous practical domains. This approach finds applications across a variety of domains, including knowledge extraction from large datasets, identification and analysis of underlying data patterns, retrieval and management of relevant information, pharmaceutical research and development, optimization of production and scheduling processes, and the extension of operational lifetimes in wireless sensor networks [3–5].

The following paragraphs present previous studies of 16 methods [6] that can perform dimensionality reduction with varying efficiencies. Table 1 presents previous studies of 16 methods that can perform dimensionality reduction with varying efficiencies.

Table 1. Sixteen method for reducing the dimension [6]

Method	Input data	Method class	Nonlinear complexity	
Principal Component Analysis (PCA)	Continuous data	Unsupervised		$O(\max(n^2p, np^2))$
Correspondence Analysis (CA)	Categorical data	Unsupervised		$O(\max(n^2p, np^2))$
Multiple Correspondence Analysis (MCA)	Categorical data	Unsupervised		$O(\max(n^2p, np^2))$
Principal Coordinates Analysis (PCoA) (classical Multidimensional Scaling (cMDS))	Distance matrix	Unsupervised		$O(n^2p)$
Non-metric Multidimensional Scaling (NMDS)	Distance matrix	Unsupervised		$O(n^2h)$
Isomap	Continuous	Unsupervised	✓	$O(n^2(p+\log n))$
Diffusion map	Continuous	Unsupervised	✓	$O(n^2p)$
Kernel PCA	Continuous	Unsupervised	✓	$O(n^2p)$
t-distributed Stochastic Neighbor Embedding (t-SNE)	Continuous/distance	Unsupervised	✓	$O(n^2p+n^2h)$
Barnes-Hut t-SNE	Continuous/distance	Unsupervised	✓	$O(nh\log n)$
Linear Discriminant Analysis (LDA)	Continuous (X and Y)	Supervised		$O(np^2+p^3)$
Partial Least Squares (PLS) (Non-linear Iterative Partial Least Squares (NIPALS))	Continuous (X and Y)	Supervised		$O(npd)$
Neighborhood Component Analysis (NCA)	Distance matrix	Supervised	✓	$O(n^2h)$
Bottleneck Neural Network (NN)	Continuous/categorical	Supervised	✓	$O(nph)$
STATIS	Continuous	Multidomain		$O(n^2P, nP^2)$
DiSTATIS	Distance matrix	Multidomain		$O(n^2P, nP^2)$

The comparative characteristics of several dimensionality reduction methods are summarized below. Each method differs in the type of input data it accepts, its methodological class (unsupervised, supervised, or multidomain), its ability to capture nonlinear relationships, and its computational complexity. Principal Component Analysis (PCA) operates on continuous data within an unsupervised framework. It is a linear method with computational complexity of $O(\max(n^2p, np^2))$. Correspondingly, Correspondence Analysis (CA) and Multiple Correspondence Analysis (MCA) are also unsupervised linear techniques applied to categorical data, both sharing the same computational complexity as PCA, namely $O(\max(n^2p, np^2))$. Principal Coordinates Analysis (PCoA), also known as classical Multidimensional Scaling (cMDS), is an unsupervised method based on a distance matrix and exhibits a computational complexity of $O(n^2p)$. Non-metric Multidimensional Scaling (NMDS), which is also distance-based and unsupervised, generally has a complexity of $O(n^2h)$, where (h) denotes the number of iterations required for convergence. Among nonlinear methods, Isomap applies to continuous data in an unsupervised manner and has a complexity of $O(n^2(p+\log n))$. The Diffusion Map and Kernel PCA are likewise nonlinear, unsupervised approaches for continuous data, each with a complexity of $O(n^2p)$. The t-distributed Stochastic Neighbor Embedding (t-SNE) method handles both continuous and distance-based inputs in an unsupervised setting, is nonlinear, and requires $O(n^2p+n^2h)$ operations, while its optimized version, Barnes-Hut t-SNE, reduces this to $O(nh\log n)$. In the supervised category, Linear Discriminant Analysis (LDA) operates on continuous predictor and response variables (X and Y), with computational complexity $O(np^2+p^3)$. The Partial Least Squares method using the NIPALS algorithm (PLS-NIPALS) is also supervised, applied to continuous X and Y, and has complexity $O(nph)$, where (d) represents the number of latent components. Neighborhood Component Analysis (NCA), a nonlinear supervised approach using a distance matrix, typically requires $O(n^2h)$ operations. The Bottleneck Neural Network (Bottleneck NN) is another nonlinear supervised method applicable to both continuous and categorical data, with complexity $O(nph)$. Finally, in the multidomain category, STATIS and its distance-based variant DiSTATIS are both linear methods. STATIS operates on continuous data with complexity $O(n^2P, nP^2)$, while DiSTATIS, using distance matrices, has an identical complexity of $O(n^2P, nP^2)$.

Essential properties include the required input data, method classification, linearity or nonlinearity, and runtime complexity, which is expressed as a function of several parameters: Here, n denotes the overall sample size, p denotes the total number of attributes in the original dataset, k specifies the chosen neighborhood size, h refers to the number of iterations, and P represents the aggregate count of features obtained from n samples across various domains.

Isomap generally computes geodesic relationships among data instances through Euclidean metrics. In contrast, the diffusion-based embedding method and kernel-based principal component analysis apply normal distribution-based kernel functions in their computations that are primarily applicable to continuous-valued data. However, categorical variables can also be analyzed by applying alternative kernel functions or dissimilarity measures.

The list of abbreviations used in this study includes CA for Correspondence Analysis patterns, cMDS for classical Multidimensional Scaling, LDA for Linear Discriminant Analysis, MCA for Multivariate Correspondence Analysis, NCA for Neighborhood Component Analysis, NIPALS for Non-linear Iterative Partial Least Squares method, NMDS for Non-metric Multidimensional Scaling, NN stands for Neural Network, PCA denotes Principal Component Analysis, PCoA refers to Principal Coordinate Analysis, t-SNE represents t-distributed Stochastic Neighbor Embedding, and PLS signifies Partial Least Squares.

To formally represent relationships among features and to assist in the feature selection process, the mathematical concept of partially ordered sets (posets) proves highly useful. Collections with hierarchical orderings, denoted by $D = (F, \prec)$, consists of a non-empty set F coupled using a pairwise relational operator $\prec$ that satisfies transitivity and anti-symmetry across every constituent. Two elements b_i and b_j within the poset are considered equivalent in order if the condition $b_i \prec b_j$ holds; otherwise, they are said to be incomparable. A chain in a poset is defined asa non-empty subset $C = \{a_1, a_2, \ldots, a_k\} \subseteq F$ in which the elements can be organized sequentially where $\boldsymbol{D}$ if $a_1 \prec a_2 \prec \ldots \prec a_k$. On the other hand, an antichain is a subset of F where no two distinct elements are equivalent in order with each other. A cover relation within a poset refers to a direct connection between elements (m, n) such that m immediately succeeds n if no intermediate element exists between them o in F satisfying $o \in F$ such that $m \prec o \prec n$. The structure of a poset can be effectively visualized using Hasse diagrams, which are graphical representations where each vertex represents an element, and edges are drawn to indicate direct cover relations while omitting transitive edges for clarity. This ensures that if $m \prec o \prec n$, the

diagram includes edges for $m \prec n$ and $n \prec o$ without redundantly displaying the edge for $m \prec o$. A systematic method for encoding the relationships within a poset involves the use of the zeta (ζ) matrix, a square matrix where both rows and columns correspond to the elements of F. Each entry in the ζ matrix is marked as 1 if the corresponding pair of elements satisfies the relation $m \prec n$, and 0 otherwise. From the ζ matrix, one can derive the cover matrix by eliminating entries corresponding to transitive relations, thereby isolating only the immediate cover relations necessary for the construction of the Hasse diagram. An important property of posets is that the set of elements can always be partitioned into disjoint chains.

This process, commonly referred to as decomposition, involves dividing a partially ordered set into a collection of disjoint chains. The minimum decomposition corresponds to the smallest possible partition that can be formed using the least number of such chains. Based on Dilworth's Theorem, the number of chains in this minimal. The partitioning of a poset is equivalent to the cardinality of the largest antichain within the poset [7]. This theorem holds major importance in combinatorial mathematics and finds practical relevance in multiple optimization-related applications. Building upon this theoretical foundation, Badr et al. introduced an Integer Linear Programming (ILP) formulation capable of efficiently computing Dilworth's decomposition [8]. In addition, they developed an effective algorithm to determine the width of a poset based on the principles established by Dilworth's Theorem [9]. Further analytical studies on jump-critical ordered sets featuring specific jump numbers were also performed, providing deeper insight into the intrinsic properties of partially ordered structures [10].

This research is chiefly focused on present a new dimensionality reduction framework founded on the concept of Hierarchical Analysis of Partially Ordered Sets. This proposed technique utilizes the hierarchical characteristics of posets to systematically guide the feature selection process, ensuring that only the most significant and informative attributes are retained for classification purposes. The experimental investigations and computational analysis confirm that the LAPOS approach surpasses many traditional and well-established feature selection methodologies across various datasets when integrated with advanced predictive modeling algorithms, such as ensemble decision tree model, categorical boosting model, Logistic-based models, and Margin-based classification models. LAPOS, PCA, and Factor Analysis (FA) achieve 92.59, 86.1, and 81.02 accuracy respectively, when employing the Support Vector Machine (SVM) algorithm. This research opens new avenues for using partially ordered set theory to tackle the issues associated with large-scale, high-dimensional datasets.

The rest of this manuscript is organized in the following manner. Chapter 2 explains the data and procedures utilized in the investigation. Chapter 3 offers an in-depth overview of the feature selection strategies. Chapter 4 describes the performance measures applied to evaluate the LAPOS approach. Chapter 5 illustrates the experimental design and interprets the outcomes derived from the study. Lastly, Chapter 6 summarizes the work, highlighting key insights and proposing avenues for subsequent research.

2. Data and methodological framework

In the following segment, the suggested technique includes its algorithm, an analysis of its time complexity, and a numerical demonstration.

2.1 *LAPOS*

This part explains the LAPOS dimensionality reduction method, which uses the structure of partially ordered sets. For each column, comparable pairs are identified using the $\leq$ relation. Intersections of these pairs form a relational (zeta) matrix and a cover matrix. A Hasse diagram is then created to represent the partial order. Based on its levels, k features are selected from top to bottom and left to right, while all others are excluded, as shown in Stepwise Process 1.

Pseudocode 1: LAPOS

Stage 1: Import the dataset $A[m][n]$.

Stage 2: Data cleaning and transformation

• Conduct data standardization (e.g., Min-Max, Z-score)

• Determine the transpose of the dataset.

Stage 3: Employ a technique for selecting relevant features

• Compute the pairwise comparisons for every feature by checking the relation (<=) between each pair of instances in that feature.

• Compute the intersection between the ordered pairs of the features to construct:

• The matrix representing feature relationships.

• The direct relation matrix for partial order representation.

• Construct the Hasse diagram from the cover matrix to graphically represent the partial order relationships among the features.

• Based on the hierarchical levels of the Hasse diagram, k relevant features are identified sequentially, progressing from the topmost to bottommost levels and in a horizontal sequence, while disregarding all other features.

Stage 4: Model Construction and Validation Using Selected Features

• The data collection is divided into development and testing subsets, employing exclusively the attributes identified by the suggested approach.

Stage 5: Develop a predictive classifier

• Train a classifier using the selected features on the training data.

Stage 6: Evaluate the Model

• Evaluate the classifier using metrics.

2.2 *Efficiency assessment of the LAPOS approach*

In this part, we examine the computational demands of the LAPOS approach across its three main operations outlined in Algorithm 1. Assume a dataset with size (R, F), where R represents the number of records (rows) and F specifies the total number of variables (features). The first procedure has a computational cost of $O(R \times F)$, and the second procedure exhibits a similar order of complexity. The final procedure incurs a complexity of $O(F^2)$, corresponding to the operations needed for pairwise feature comparisons and selection within the Hasse diagram framework, leading to an overall complexity of $O(R \times F^2)$.

Overall efficiency assessment of the proposed LAPOS algorithm takes into account all computational operations performed across its main procedures, including pairwise feature comparisons and selection within the Hasse diagram framework $O(RF^2)$.

2.3 *Illustrative example*

This part presents a practical example to illustrate how the LAPOS method, as outlined in Algorithm 1, operates in practice.

Stage 1: Load the initial dataset into the workspace, which contains three samples and ten attributes, as illustrated in Table 2.

Table 2. The original data set

	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$	$E7$	$E8$	$E9$	$E10$
$Q1$	11	20	14	13	17	13	14	12	14	9
$Q2$	0.2	0.4	0.3	0.3	0.3	0.4	0.3	0.6	0.4	0.4
$Q3$	31	55	41	63	45	51	41	41	45	29

Stage 2: Data preprocessing: We transposed the original dataset as shown in Table 3.

Table 3. The transposed original data set

	$Q1$	$Q2$	$Q3$
$E1$	11	0.2	31
$E2$	20	0.4	55
$E3$	14	0.3	41
$E4$	13	0.3	63
$E5$	17	0.3	45
$E6$	13	0.4	51
$E7$	14	0.3	41
$E8$	12	0.6	41
$E9$	14	0.4	45
$E10$	9	0.4	29

After that, we normalize the transposed original dataset as shown in Table 4.

Table 4. The normalized transposed original data set

	$Q1$	$Q2$	$Q3$
$E1$	0.350649	0	1
$E2$	0.358974	0	1
$E3$	0.336609	0	1
$E4$	0.202552	0	1
$E5$	0.373602	0	1
$E6$	0.249012	0	1
$E7$	0.336609	0	1
$E8$	0.282178	0	1
$E9$	0.304933	0	1
$E10$	0.300699	0	1

Stage 3: Implement the LAPOS method using the following procedure.

Identify all comparable element pairs within each column based on the binary relation ($\leq$) used in this example. For instance, consider Column 1: ($E10$, $E1$), ($E10$, $E2$), ($E10$, $E3$), ($E10$, $E4$), ($E10$, $E5$), ($E10$, $E6$), ($E10$, $E7$), ($E10$, $E8$), ($E10$, $E9$), ($E1$, $E2$), ($E1$, $E3$), ($E1$, $E4$), ($E1$, $E5$), ($E1$, $E6$), ($E1$, $E7$), ($E1$, $E8$), ($E1$, $E9$), ($E1$, $E10$), ($E8$, $E2$), ($E8$, $E3$), ($E8$, $E4$), ($E8$, $E5$), ($E8$, $E6$), ($E8$, $E7$), ($E8$, $E9$), ($E4$, $E3$), ($E4$, $E5$), ($E4$, $E6$), ($E4$, $E7$), ($E4$, $E2$), ($E4$, $E9$), ($E6$, $E2$), ($E6$, $E3$), ($E6$, 5), ($E6$, $E7$), ($E6$, $E9$), ($E3$, $E7$), ($E3$, $E9$), ($E3$, $E5$), ($E3$, $E2$), ($E7$, $E9$), ($E7$, $E5$), ($E7$, $E2$), ($E9$, $E5$), ($E9$, $E2$), ($E5$, $E2$).

After determining the comparable pairs for all columns, compute their intersection and subsequently derive the zeta matrix, as illustrated in Table 5.

Based on the obtained zeta matrix, the cover matrix can be derived following the procedure outlined in the introduction section (Table 6).

Table 5. Zeta matrix

	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$	$E7$	$E8$	$E9$	$E10$
$E1$	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$E2$	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0
$E3$	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
$E4$	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
$E5$	1.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0
$E6$	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0
$E7$	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
$E8$	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0
$E9$	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	1.0
$E10$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0

Table 6. Zeta matrix

	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$	$E8$	$E9$	$E10$
$E1$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$E2$	0.0	0.0	0.0	0.0	1.0	1.0	0.0	1.0	0.0
$E3$	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$E4$	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$E5$	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
$E6$	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
$E8$	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
$E9$	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0
$E10$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

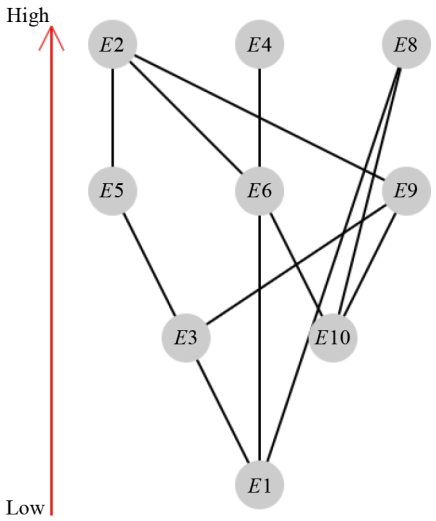

Figure 1. Final results of Algorithm 1 using a Hasse diagram for example 1

Ultimately, a Hasse diagram depicting the partially ordered set was created, as shown in Figure 1. Using this diagram, k features are chosen by moving sequentially from the top to the bottom levels and from left to right, following their relative importance.

3. Feature selection methods

Techniques for relevant Features Selecting (FS) are commonly divided into three primary categories: independent feature assessment, predictive model evaluation, and embedded selection frameworks [11]. Each category employs a distinct strategy to recognize and keep the most informative features while eliminating irrelevant or redundant ones, thus effectively reducing the dataset's dimensionality and refining the effectiveness and generalization ability of machine learning systems.

Filter methods are characterized by their independence from any specific machine learning algorithm during the feature selection process. These approaches assess the significance of individual attributes by employing statistical indicators or dependency measures that capture the relationship between predictors and the response variable. By assessing the strength of association between each feature and the output variable, filter methods can systematically eliminate irrelevant or redundant features prior to the model development stage. An example of a widely used filter method is the Pearson Correlation Coefficient, which quantifies the linear association between two variables and assists in detecting features exhibiting a strong linear connection with the target attribute. The primary advantage of filter methods lies in their computational performance and adaptability, rendering them appropriate for high-dimensional datasets. However, these techniques overlook the interdependencies among features, which may limit their effectiveness in capturing complex feature dependencies.

Wrapper methods, in contrast to filter methods, incorporate the feature selection process within the model development stage by employing a predictive model to investigate the effectiveness of different feature subsets. These methods employ an iterative process to search for the chosen attributes that optimizes the performance of the model based on a chosen evaluation metric, such as *accuracy* or *F1-score*. Wrapper methods frame feature selection as a search-based optimization problem, often leveraging strategies like progressive inclusion, stepwise exclusion, and iterative feature pruning to determine the subset of features that provides the highest informative value. An example of a wrapper method is SelectKBest isolates the k most influential features according to the weighting scheme derived from the applied evaluation measure. While wrapper methods can provide a more accurate and tailored feature selection process by considering feature interactions, they are computationally intensive, particularly when applied to large-scale datasets with numerous features.

Embedded methods, also known as model-driven feature selection techniques, merge the feature evaluation stage as an inherent part of the model training routine. These techniques intrinsically conduct feature identification within the framework of model formation, thus concurrently refining the model parameters while pinpointing the key features that improve the model's forecasting performance. Embedded methods benefit from being less computationally demanding than wrapper methods while still considering the interactions between features during the selection process. Illustrations of integrated approaches comprise the inherent attribute weighting mechanisms found in ensemble-based classifiers such as Random Forests, which prioritize variables based on their contribution to model accuracy, as well as L1-penalized linear models like Lasso, where less significant predictors are constrained toward null values, effectively excluding them from the analysis. By embedding the selection mechanism within the training framework, these techniques achieve an optimal compromise between computational economy and selection precision.

In summary, the selection between preprocessing-based, model-evaluation, and integrated feature selection techniques for feature selection is determined by the characteristics and structure of the data collection, the computational resources available, and the complexity of feature interactions within the data. Filter methods are preferred when computational efficiency is prioritized, wrapper methods are suitable when a high level of accuracy and feature interaction analysis is desired, and embedded methods provide an effective middle ground by integrating feature selection within the model training process to enhance predictive performance while managing computational complexity.

4. Evaluation metrics

The results derived from the proposed methodology, together with a detailed discussion, are presented in this section. To facilitate the development, execution, and visualization of the machine learning workflows, the contemporary Interactive Development Environment (IDE) Jupyter Lab was employed owing to its flexibility and seamless integration with a wide range of data science libraries.

During the experimental phase of this study, Various classification algorithms—including Logistic-based Model (LM), Support Vector Machine Classifier (SVMC), k-Nearest Neighbor Method (kNNM), Random Forest Model (RFM), Linear Discriminant Classifier (LDC), and Decision Tree Model (DTM)—were employed to evaluate the effectiveness and accuracy of the suggested feature selection methodology. These algorithms were selected due to their proven robustness, interpretability, and extensive use in classification problems across various data domains.

To comprehensively assess model performance, standard evaluation metrics were adopted, the evaluation framework employed performance indicators such as accuracy, precision, recall, and the F1-score. Accuracy measures the proportion of correctly classified samples relative to the total dataset, serving as a general gauge of model effectiveness. Precision, which expresses the proportion of correctly identified positive cases among all predicted positives, indicates the model's ability to reduce false alarms. Recall, defined as the proportion of true positives detected out of all actual positives, evaluates the model's capability to capture relevant instances. The F1-score, computed as the harmonic mean of precision and recall, offers a unified metric that balances the effects of both false positives and false negatives in model evaluation.

Collectively, these performance indicators provide a comprehensive evaluation of the predictive strength and operational efficiency of the machine learning models developed using the features determined by the proposed method. The outcomes obtained from this assessment emphasize the effectiveness of the feature selection framework in enhancing classification accuracy while maintaining computational efficiency and interpretability throughout the learning procedure.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{1}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positives}} \tag{2}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{3}$$

$$\text{F1 Score} = \frac{2(\text{True Positive})}{2(\text{True Positive}) + \text{False Positive} + \text{False Negative}} \tag{4}$$

The Python (3.10), 64-bit Windows 8.1, Core (TM) i5 CPU M 460 @2.53 GHz, and 4.00 GB of memory are the components of the software environment.

5. Findings and analysis

In this section, the proposed LAPOS framework was tested on medical datasets due to their availability and reliability for dimensionality reduction evaluation. However, the methodology itself is general and can be applied to other data domains, such as industrial or environmental data. We have intentionally kept the paper focused on medical datasets to maintain clarity and coherence. Here, the processes of feature selection and dimensionality reduction utilized several recognized methodologies aimed at improving the efficiency and accuracy of the classification models. Techniques applied in this context included Principal Component Analysis (PCA), Independent Component Analysis (ICA), and

Factor Analysis (FA)—each functioning as a dimensionality reduction strategy aimed at transforming data with numerous attributes into a lower-dimensional form while preserving the most meaningful and distinguishing structures contained in the dataset.

Independent Component Analysis (ICA) was specifically applied to isolate mutually independent components, proving valuable in disentangling composite signals and revealing latent elements underlying the observed data patterns. Through the extraction of non-Gaussian and statistically independent attributes, ICA strengthens the distinguishing ability of classification algorithms, allowing them to better model intricate and hidden dependencies within the dataset.

Furthermore, Factor Analysis (FA) was applied to identify latent variables that account for the correlations among observed features. Through dimensionality reduction based on these latent constructs, FA streamlines the feature domain by reducing redundancy and preserving the fundamental interconnections among variables, thus enhancing the clarity of data interpretation and boosting computational effectiveness. To capture the complex nonlinear relationships within high-dimensional datasets, t-distributed Stochastic Neighbor Embedding (t-SNE) was employed as a nonlinear dimensionality reduction approach. This method projects the data into a reduced-dimensional space while maintaining the proximity of neighboring data points, effectively preserving local structural relationships, which facilitates the visualization of clusters and enhances the separability of different classes. Discriminant Component Analysis (DCA) was incorporated to maximize class separability by projecting the data onto directions that highlight differences between distinct classes while reducing variability within each class, thus enhancing the overall effectiveness of the subsequent classification models. For feature selection, information gain and the Chi-square statistical test were employed to evaluate and prioritize the most relevant and informative attributes within the dataset. These statistical measures identify attributes that play the most significant role in enhancing the predictive performance of the models, and were also combined with PCA (Chi-square + PCA) to obtain a reduced yet highly informative feature set. To mitigate issues arising from class imbalance in the datasets, the preprocessing workflow incorporated the Synthetic Minority Oversampling Technique (SMOTE) and Adaptive Synthetic Sampling (ADASYN) were employed. SMOTE tackles class imbalance by producing artificial samples for the minority category through interpolation between existing minority observations, resulting in a more balanced dataset while avoiding direct data replication.

Meanwhile, ADASYN extends this concept by dynamically producing artificial samples, with greater emphasis on minority instances that are more difficult to classify determined by their proximity to majority class samples. This adaptive approach enables the model to better capture complex decision boundaries and enhances the classifier's ability to generalize, particularly in scenarios with severe class imbalance.

The incorporation of these resampling techniques ensures that the training process remains unbiased toward the majority class, leading to more reliable and equitable model performance across all classes, thereby balancing the class distribution without merely replicating existing samples. ADASYN, an extension of SMOTE, further enhances this procedure by concentrating on producing artificial samples for minority instances that are more challenging to classify, thereby strengthening the model's capacity to effectively differentiate between categories in imbalanced datasets. Following the feature selection and data balancing steps, the prepared datasets were utilized to train and evaluate two widely adopted machine learning algorithms. These classifiers were selected for their proven effectiveness in handling classification tasks across a range of domains and their ability to provide interpretable results. The integration of these dimensionality reduction and data balancing techniques with the classification algorithms ensures that the models are capable of learning from a compact, balanced, and informative set of features, consequently enhancing the overall predictive accuracy and robustness of the proposed methodology.

Table 7, Figures 2 and 3 illustrate that the suggested LAPOS approach surpasses both PCA and FA when employing the SVM algorithm. The proposed LAPOS approach surpasses ICA solely when employing the Linear Regression (LR) algorithm. On the other hand, LAPOS overcomes PCA and ICA when employing the K-Nearest Neighbor (KNN) algorithm. Finally, LAPOS overcomes FA when employing the Decision Tree (DT) algorithm.

Table 7. Comparison results of backbone improvements

	SVM			Logistic regression		
	PCA [12]	FA [12]	LAPOS	PCA [13]	ICA [13]	LAPOS
Acc	93.04	89.57	**95.61**	**97.6**	75.8	95.61
Prec	93.03	89.64	**97.5**	96.9	75.6	**97.5**
Rec	**93.11**	89.47	90.69	**98.5**	89.7	90.69
F1	93.04	89.53	**93.97**	**96.4**	61.5	93.97
Time (Sc)	-	-	1.953125	-	-	0.1875
	LDA			KNN		
	PCA [14]	ICA [14]	LAPOS	FA [15]	T-SNE [15]	LAPOS
Acc	81.45	84.22	**93.85**	66.7	73.3	**94.73684**
Prec	-	-	**92.69**	66.7	73.8	**90.96573**
Rec	79.85	82.97	**100**	66.7	73.3	**96.05263**
F1	-	-	**86.39**	66.7	73.5	**86.3905**
Time (Sc)	-	-	0.453125	-	-	0.21875
	Rnadom forest			Decision tree		
	Without [16]	SMOTE + PCA [16]	LAPOS	FA [15]	T-SNE [15]	LAPOS
Acc	96.05	**98.32**	96.49123	80.00	**93.3**	91.22807
Prec	96.00	98.00	**100**	80.8	94.4	**100**
Rec	95.00	98.00	**100**	80.00	93.3	**100**
F1	96.00	98.00	**100**	79.4	93.3	**100**
Time (Sc)	-	-	0.46875	-	-	0.234375

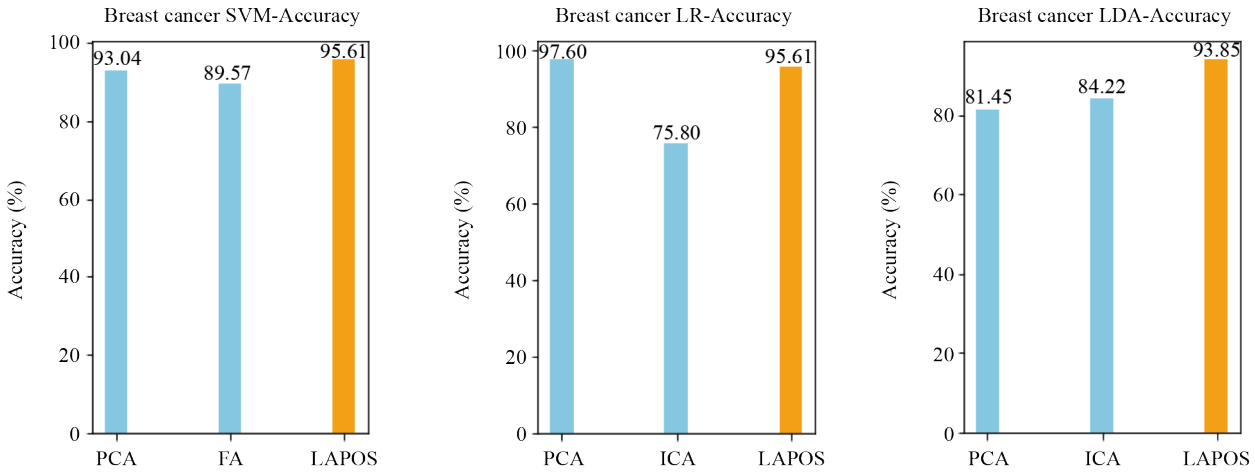

Figure 2. Accuracies for 3 models (SVM, LR, and LDA) to evaluate LAPOS with other techniques (PCA and ICA) for breast cancer dataset

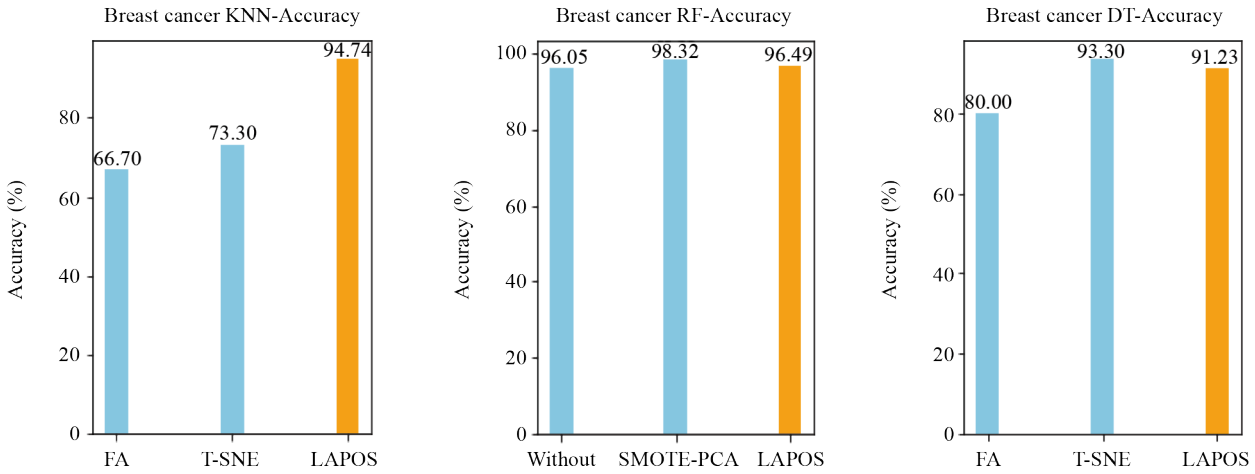

Figure 3. Accuracies for 3 models (KNN, RF, and DT) to evaluate LAPOS with other techniques (FA and T-SNE) for breast cancer dataset

Table 8. Run 6 models on heart disease dataset to evaluate LAPOS with other techniques

	SVM			Logistic regression		
	PCA [17]	FA [17]	LAPOS	PCA [13]	ICA [13]	LAPOS
Acc	86.1	81.02	**92.59**	84.1	80.5	**92.59**
Prec	81.00	75.00	**100**	83.7	86.5	**100**
Rec	**90.4**	88.89	80.95	**87.8**	54.2	80.95
F1	86.00	81.00	**89.47**	80.6	87.7	**89.47**
Time (Sc)	-	-	0.218	-	-	0.23
	LDA			KNN		
	Without [18]	PCA [18]	LAPOS	PCA [19]	SMOTE [19]	LAPOS
Acc	92.2	**96.5**	88.88	**94.26**	80.04	74.07
Prec	92.00	**96.4**	82.9	**91.68**	71.02	83.00
Rec	91.8	**96.3**	85.1	**98.30**	98.25	82.17
F1	91.9	**96.3**	80.8	**94.88**	82.45	83.83
Time (Sc)	-	-	0.375	-	-	0.234375
	Rnadom forest			Decision tree		
	SMOTE [19]	SMOTE-ENN [19]	LAPOS	Info-gain [20]	Chi-Square + PCA [20]	LAPOS
Acc	89.64	**92.16**	87.03	**89.00**	87.00	79.62
Prec	88.52	90.57	**100**	87.00	81.00	**100**
Rec	89.94	97.39	**100**	94.00	99.00	**100**
F1	89.23	93.85	**100**	90.00	89.00	**100**
Time (Sc)	-	-	0.375	-	-	0.203125

Table 8 and 9 illustrates that the suggested LAPOS approach surpasses both PCA and FA when employing the SVM algorithm. The proposed LAPOS approach surpasses PCA and ICA when employing the LR algorithm. On the other hand, Table 8 illustrates that LAPOS overcomes PCA and DCA when employing the KNN algorithm. Finally, Table 10 illustrates that LAPOS overcomes PCA and ICA when employing the LR algorithm.

Table 9. Run 6 models to evaluate LAPOS with other techniques for IONOSPHERE dataset

	SVM			Logistic regression			KNN		
	SMOTE [21]	ADASYN [21]	LAPOS	PCA [13]	ICA [13]	LAPOS	PCA [22]	DCA [22]	LAPOS
Acc	90.88	**91.69**	80.95	**88.9**	87.8	83.09	87.89	88.57	**88.73239**
Prec	88.87	93.1	**95.00**	89.2	87.8	**95.00**	-	-	**95.28796**
Rec	**85.2**	82.88	73.07	76.4	**76.9**	73.07	-	-	**91**
F1	86.85	**87.62**	82.60	**91.8**	91.00	73.91	-	-	**100**
Time (Sc)	-	-	0.257	-	-	0.25	-	-	0.25

Table 10. Run 6 models to evaluate LAPOS with other techniques for SONAR dataset

	SVM			Logistic regression		
	PCA [23]	ICA [23]	LAPOS	PCA [13]	ICA [13]	LAPOS
Acc	**82.7**	71.7	80.95	75.1	65.6	**78.57**
Prec	-	-	**95.00**	77.5	67.7	**90.47**
Rec	-	-	**73.07**	72.00	62.3	**73.07**
F1	-	-	**82.60**	76.3	67.4	**80.85**
Time (Sc)	-	-	0.234	-	-	0.28125

6. Conclusive insights and research outlook

This research introduces a novel dimensionality reduction method referred to as Level Analysis of Partially Ordered Sets (LAPOS), which is conceptually based on the intrinsic structural characteristics derived from the principles of partial order theory. The LAPOS framework exploits the hierarchical dependencies among features, offering a structured and interpretable mechanism for compressing data dimensions while maintaining the highly representative and discriminative attributes necessary for reliable classification outcomes. To examine the efficiency and reliability of the developed LAPOS technique, extensive computational experiments were carried out using two well-established predictive supervised learning techniques, namely—Support Vector Machine (SVM) and Linear Regression (LR)—both of which are widely esteemed for their robust predictive capability in handling high-dimensional classification tasks. The results obtained from these experiments demonstrated that LAPOS exhibits superior performance in handling medical datasets when compared to conventional dimensionality reduction methods. The technique effectively enhanced the predictive capabilities of the classifiers, confirming its potential in real-world data analysis applications, particularly in the medical and healthcare domains where high-dimensional data is prevalent. Looking forward, the future direction of subsequent analysis will be devoted to extending the application of LAPOS across a broader spectrum of datasets and machine learning classifiers to comprehensively assess its versatility and robustness. Further, an in-depth exploration of the theoretical underpinnings of LAPOS will be undertaken to refine and optimize its performance, ensuring its adaptability to various data distributions and complexities. In addition to enhancing the methodological framework of LAPOS, the future work will prioritize

the integration of LAPOS with advanced artificial intelligence models, including ensemble learning techniques and deep learning architectures, to harness the combined strengths of interpretable feature selection and powerful predictive modeling. This integration aims to further improve classification accuracy while maintaining runtime efficiency and model comprehensibility, thereby expanding the applicability of LAPOS when dealing with extensive and complex datasets analysis scenarios. Through these future endeavors, LAPOS is expected to evolve into a comprehensive and efficient dimensionality reduction tool, contributing significantly to the advancement of machine learning pipelines in both research and practical applications [24–26].

Acknowledgement

The author would like to thank Deanship of Scientific Research at University of Science and Technology (MUST) for supporting this work. We also thank the referees for suggestions to improve the presentation of this paper.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Neggaz N, Houssein EH, Hussain K. An efficient Henry gas solubility optimization for feature selection. *Expert Systems with Applications*. 2020; 152: 113364. Available from: https://doi.org/10.1016/j.eswa.2020.113364.

[2] Yi JH, Deb S, Dong J, Alavi AH, Wang GG. An improved NSGA-III algorithm with adaptive mutation operator for big data optimization problems. *Future Generation Computer Systems*. 2018; 88: 571-585. Available from: https://doi.org/10.1016/j.future.2018.06.008.

[3] Sayed SAF, Nabil E, Badr A. A binary clonal flower pollination algorithm for feature selection. *Pattern Recognition Letters*. 2016; 77: 21-27. Available from: https://doi.org/10.1016/j.patrec.2016.03.014.

[4] Martin-Bautista MJ, Vila MA. A survey of genetic feature selection in mining issues. In: *Proceedings of the 1999 Congress on Evolutionary Computation (CEC99)*. Washington, DC, USA: IEEE; 1999. p.1314-1321.

[5] Piramuthu S. Evaluating feature selection methods for learning in data mining applications. *European Journal of Operational Research*. 2004; 156(2): 483-494. Available from: https://doi.org/10.1016/S0377-2217(02)00911-6.

[6] Nguyen LH, Holmes S. Ten quick tips for effective dimensionality reduction. *PLoS Computational Biology*. 2019; 15(6): e1006907. Available from: https://doi.org/10.1371/journal.pcbi.1006907.

[7] Dilworth RP. A decomposition theorem for partially ordered sets. *Annals of Mathematics*. 1950; 51(1): 161-166. Available from: https://doi.org/10.2307/1969503.

[8] Badr E, Selim IM, Mostafa H, Attiya H. An integer linear programming model for partially ordered sets. *Journal of Mathematics*. 2022; 2022: 7660174. Available from: https://doi.org/10.1155/2022/7660174.

[9] Badr E, EL-Hakeem M, El-Sharawy EE, Ahmed TE. An efficient algorithm for decomposition of partially ordered sets. *Journal of Mathematics*. 2023; 2023: 9920700. Available from: https://doi.org/10.1155/2023/9920700.

[10] Badr EM, Moussa MI. On jump-critical ordered sets with jump number four. *Journal of Advanced Applied and Computational Mathematics*. 2014; 1: 8-13.

[11] Chen CW, Tsai YH, Chang FR, Lin WC. Ensemble feature selection in medical datasets: combining filter, wrapper, and embedded feature selection results. *Expert Systems*. 2020; 37(5): e12553. Available from: https://doi.org/10.1111/exsy.12553.

[12] Khan AA, Bakr MA. Enhancing breast cancer diagnosis with integrated dimensionality reduction and machine learning techniques. *Journal of Computational and Biomedical Informatics*. 2024; 7(2): 1-17.

[13] Musa AB. A comparison of ℓ1-regularization, PCA, KPCA and ICA for dimensionality reduction in logistic regression. *International Journal of Machine Learning and Cybernetics*. 2014; 5: 861-873. Available from: https://doi.org/10.1007/s13042-013-0171-7.

[14] Costa DD, Campos LF, Barros AK. Classification of breast tissue in mammograms using efficient coding. *Biomedical Engineering Online*. 2011; 10: 55. Available from: https://doi.org/10.1186/1475-925X-10-55.

[15] Al-Qazzaz NK, Mohammed IK, Al-Qazzaz HK, Ali SHBM, Ahmad SA. Comparison of the effectiveness of various classifiers for breast cancer detection using data mining methods. *Applied Sciences*. 2023; 13(21): 12012. Available from: https://doi.org/10.3390/app132112012.

[16] Oktafiani R, Sela EI. Breast cancer classification with principal component analysis and SMOTE using random forest method and support vector machine. *International Journal of Computer Applications*. 2024; 186: 16.

[17] Shilaskar S, Ghatol A. Dimensionality reduction techniques for improved diagnosis of heart disease. *International Journal of Computer Applications*. 2013; 61(5): 1-8.

[18] Ali S, Hassan M, Kim JY, Farid MI, Sanaullah M, Mufti H. FF-PCA-LDA: intelligent feature fusion based PCA-LDA classification system for plant leaf diseases. *Applied Sciences*. 2022; 12(7): 3514. Available from: https://doi.org/10.3390/app12073514.

[19] Wei X, Shi B. Reducing bias in coronary heart disease prediction using SMOTE-ENN and PCA. *PLoS One*. 2025; 20(8): e0327569. Available from: https://doi.org/10.1371/journal.pone.0327569.

[20] Alwakid G, Ul Haq F, Tariq N, Humayun M, Shaheen M, Alsadun M. Optimized machine learning framework for cardiovascular disease diagnosis: a novel ethical perspective. *BMC Cardiovascular Disorders*. 2025; 25(1): 123. Available from: https://doi.org/10.1186/s12872-025-04550-w.

[21] Huang PJ. *Classification of imbalanced data using synthetic over-sampling techniques*. Master Thesis. Los Angeles: University of California; 2015.

[22] Chen L, Wei P, Long Z, Yu Y. Adaptive metric learning for dimensionality reduction. *Journal of Computer and Communications*. 2022; 10(12): 95-112. Available from: https://doi.org/10.4236/jcc.2022.1012008.

[23] El Ferchichi S, Zidi S, Maouche S, Laabidi K, Ksouri M. Feature clustering based MIM for a new feature extraction method. *International Journal of Computers, Communications and Control*. 2013; 8(5): 699-707.

[24] Ahmad PN, Ullah I, Aboelenein NM, Singh KS, Jiang W, Al-Khasawneh MA, et al. Deep neural network-based feature encoding for automated health monitoring using large ai models in online communication systems. *ACM Transactions on Internet of Things*. 2025; 7(1): 1-20. Available from: https://doi.org/10.1145/3744754.

[25] Zang X, Zheng Z, Zheng H, Liu X, Khan MK, Jiang W. HyperEye: a lightweight features fusion model for unknown encrypted malware traffic detection. *IEEE Transactions on Consumer Electronics*. 2025; 71(2): 5079-5089. Available from: https://doi.org/10.1109/TCE.2025.3558353.

[26] Yang B, Zhou J, Zhang S, Xing Y, Jiang W, Xu L. A lightweight knowledge distillation and feature compression model for user CTR prediction in edge computing scenarios. *IEEE Internet of Things Journal*. 2025; 12(3): 2295-2308. Available from: https://doi.org/10.1109/JIOT.2024.3446640.