Comparative Analysis of Feature Selection and Dimensionality Reduction Techniques for High-Dimensional Datasets

Sushil Kumar [1]

[1]Assistant Professor, School of Management Sciences, Varanasi, India

Abstract

The existence of high-dimensional data is becoming widespread in fields including healthcare, bioinformatics, text analytics, and factory diagnostics, and has created difficulties in computational complexity, data redundancy, and model overfitting. Learning efficiency and predictive performance hence require effective techniques of feature selection and dimensionality reduction in order to enhance the efficiency of learning. The paper includes a comparative analysis of four popular methods, such as Chi-Square feature selection, Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA), on several high dimensions of data (gene expression, text and medical imaging data) and assessed their effectiveness. There are experimental findings that show that supervised techniques are always better as compared to unsupervised and filter techniques. The maximum classification accuracy of LDA was 90.5, precision was 89.7, recall was 89.2 and F1-score was 89.5 which minimized the feature space by 99.98. There was also good performance of RFE with an average accuracy of 88.9 but at a greater cost of computation. On the contrary, Chi-Square selection was the quickest to execute (in terms of time) (about 1.0 s) but less accurate (84.6%). PCA efficiently and accurately (balance between efficiency and performance) retained 95 percent variance. Altogether, the results show that the effectiveness of dimensionality reduction methods strongly depends on the data characteristics and the goal of application that can provide a reasonable guide to the choice of the adequate methods when working with big data set.

Keywords- Feature Selection, Dimensionality Reduction, High-Dimensional Data, Machine Learning, ***Classification***

Corresponding Author: Sushil Kumar Assistant Professor, School of Management Sciences (SMS), Varanasi, India E-mail id: sushilkumar@smsvaranasi.com

How to cite this article: Sushil Kumar(2026). Comparative Analysis of Feature Selection and Dimensionality Reduction Techniques for High-Dimensional Datasets, Confluence: The Journal of Technology, Business & Society, Vol. I(1) 62-71

Source of support: Nil

Conflict of Interest: None

Received:12.05.2026; Accepted :13.06.2026; Published :22.7.2026

I. INTRODUCTION

The use of data-intensive applications in bioinformatics, medical diagnosis, finance, image processing and text analytics, has led to the creation of high-dimensional datasets, so that the dimensions of feature spaces are vast in comparison with the number of observations [1]. Despite the rich descriptive information provided by such datasets, these datasets also provide a big challenge when it comes to data analysis and machine learning models. In many cases, high dimensionality results in a greater computational burden, redundancies between features, overfitting, and poor predictive ability what is popularly called the curse of dimensionality [2]. Therefore, dimensionality reduction and data retention have become a highly important

data preprocessing step of modern data-driven research. Dimensionality reduction and feature selection are two popular methods of dealing with high dimensionality. The feature selection methods are aimed at choosing a subset of the available salient features based on the feature space that contains the original features from the irrelevant, redundant or noisy features [3]. This method maintains the original feature semantics, and thus gives the end result more representatives of intuitive meaning as well as computationally efficient models. The approaches to feature selection can be roughly divided into filter approach, wrapper approach and embedded approach, each having varying trade-offs between computation costs and predictive accuracy. Conversely, dimensionality reduction methods are used to reduce a high-dimensional data set to a lower-dimensional space, i.e. by forming new transformed features that reflect the underlying structure or variance of a data set. They are especially useful in dealing with highly correlated features and enhancing visualization and learning effectively. Nevertheless, the transformed characteristics might not be directly interpretable, a weakness in the fields where that is a critical requirement. There has been no universally best method of dimensionality reduction given the increasing diversity of high dimensional data and learning tasks. The efficiency of feature selection and dimensionality reduction algorithms varies depending on the size of the dataset, strength of correlation between features, level of noise and machine learning task required. Thus, these methods should be compared in a systematic comparative study in order to comprehend their strengths and limitations to their use and their applicability. The current research will attempt to offer such comparative assessment and offer insights that may direct practitioners in their choice of suitable dimensionality reduction techniques when analyzing high-dimensional data.

II. RELATED WORKS

The techniques of dimensionality reduction and feature selection have been explored widely as a necessity procedure prior to processing high-dimensional data. Lately, there has been major research effort towards enhancing the predictive performance, interpretability, and computational efficiency in the multitude of areas of application, including healthcare and energy systems as well as industrial diagnostics and big data analytics. Fira et al. [15] offered a theoretical and empirical analysis of sparse feature selection algorithms with emphasis on the fact that they deliver lower overfitting and retain the model intelligibility. Their analysis pointed to the fact that sparsity-based methods are especially useful in high feel, low sample scenarios, and that the ability depends on regularization parameters. In the same vein, Huang et al. [16] introduced a feature selection method of Q-learning enhanced differential evolution methodology used in high dimensional medical datasets. Their strategy, based on reinforcement learning, proved to be better convergent and better able to classify, in comparison to more traditional metaheuristic approaches, which supports the increase in the use of hybrid and adaptive feature selection models. There are also advanced and new paradigms that are discussed. Kuzu and Uysal [17] explored the problem of incorporating quantum machine learning into hybrid feature selection systems in high-energy particle physics. Their results proposed that quantum-inspired feature selection is capable of efficiently conducting searches of large feature spaces, but practical scaling is a problem. Along a less problematic (cognitively inspired) direction, Liu et al. [18] proposed an evidential k-nearest neighbors model with cognitive-based feature selection with strong results in noisy large-dimensional data such that uncertainty and relevance are explicitly modeled. A number of studies have been devoted to domain-specific comparative studies. Liyew et al. [19] made an elaborate review of feature selection approaches in predicting evapotranspiration and noted that no single approach has been universal across datasets, and thus comparison studies are necessary. Maciejewski et al., [20] compared several feature selection and classification methods of broken rotor bar diagnosis with vibration and current signals and presented the results that wrapper-based approaches tend to be effective in fault diagnosis problems, with more expensive computational costs than filter-based methods.

Mohi Uddin et al. [21] compared traditional and ensemble machine learning models with improved feature selection in the early detection of polycystic ovary syndrome (PCOS) in the healthcare applications. They demonstrated the substantial increase in diagnostic accuracy and model stability with optimized feature selection. Similarly, Mohtasham et al. [22] have been doing a comparative study of feature selection methods in COVID-19 datasets, citing that supervised feature selection methods outperform unsupervised methods in terms of accuracy, especially when used at the clinical decision-support system. Methodological assessment has also been paid attention to. Mostert et al. [23] suggested a single measure of performance to be used to objectively compare the feature selection algorithms to curb the inconsistency in evaluation practice of various studies. In education and healthcare data, Okan et al. [24] compared the variants of recursive feature elimination algorithms benchmarking and found that the variants of the algorithm and stopping criteria significantly impact the performance results. Investigations on metaheuristic have been also undertaken by Ozseven and Arpacioglu [25], who had proven that evolutionary feature selection methods were very useful in improving the accuracy of emotion recognition in speech when huge acoustic feature dimensions were involved. Lastly, a hybrid optimized model of the IoT-based healthcare systems was proposed by Palanisamy et al. [26], who combined the feature selection with the big data classification. Their findings proved that dimensionality reduction with optimized learning models results in scalability and predictive accuracy. In general, the current literature highlights the significance of the comparative analyses between various techniques and datasets. Although previous researchers used to examine particular areas or specific types of non-experimental methods, this study contributes and develops the previous research by comparing both feature selection methods and dimensionality reduction methods in a unified experimental setup, addressing the gaps found in the existing part of research [15-26].

III. METHODS AND MATERIALS

The given study will be taken as a systematic experiment because the results of feature selection and dimensionality reduction algorithms are going to be compared on high-dimensional datasets. The Materials and Methods section will explain the data sets employed, the steps performed prior to preprocessing, the algorithms to be used, the experimental design and the evaluation protocol so that there is fair and reproducible comparison [4].

Data Description and Preprocessing

The given experimental analysis was performed on three representative high-dimensional datasets that are widely used in machine learning studies: (i) a gene expression dataset with binary disease prediction, (ii) a text-based document prediction dataset with TFs-IDF features, and (iii) a medical image-based feature dataset with handcrafted radiomic feature. In these datasets, the features were between 2,000 and 10,000, and the instances between 500 and 3,000, both of which are common high-dimensional conditions.

All data sets were subjected to a normal preprocessing before analysis. Mean imputation was used to handle the missing values in continuous features. Z-score normalization was used to normalize features to obtain similar scales which is critical in distance-based and variance-based methods. Then, the datasets were divided into training and testing (70 and 30), respectively. Data leakage was avoided by dimensionality reduction and feature selection applied direct on the training data, and applied the resulting transformations on the test data [5].

Algorithms Considered

To have the representation of the feature selection and dimensionality reduction types, four well-known algorithms were chosen: Chi-Square Feature Selection, Recursive Feature Elimination (RFE), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). These techniques have been selected because of their topicality, a variety of methods, and the common application in high-dimensional data analysis [6].

Chi-Square Feature Selection Chi-square feature selection is a filter based feature selection method that individual features are considered by measuring how statistically dependent it is to the target variable. Attributes that are chi-square ranked higher provide stronger attributes to the classification tasks. This algorithm is computationally inexpensive, can scale to extremely high-dimensional data, and best works with discrete, or discrete-time, data like text representations [7]. It however does not factor in inter-feature dependencies, which could lead to selection of features redundantly.

```
"Input: Dataset D(X, y), number of features k
For each feature xi in X:
    Compute chi-square score between xi and y
Rank features by score in descending order
Select top k features
Return reduced dataset"
```

Recursive Feature Elimination (RFE) is a wrapper-based feature selection algorithm which removes the least important features recursively depending on model performance. It trains a base learning model, normally a linear classifier, to sort features by their predictive value. Each time, the features that have the least importance are discarded until one gets the stipulated required number of features [8]. RFE tends to have a stronger predictive accuracy than use of filters but is computationally costly in large feature space.

```
"Input: Dataset D(X, y), estimator M, target features k
Initialize feature set F = X
While |F| > k:
    Train M using F
    Rank features by importance
    Remove least important feature from F
Return reduced dataset"
```

Principal Component Analysis (PCA) is an unsupervised dimensionality reduction algorithm which converts the given feature space in a reduced set of orthogonal components. The components are linear combinations of original features which are ranked in order of the extent of variance explained [9]. PCA is very useful in eliminating redundancy and noises due to correlated features, and enhances the computing efficiency.

```
"Input: Dataset X
Compute covariance matrix of X
Calculate eigenvalues and eigenvectors
Sort eigenvectors by descending eigenvalues
Project X onto top components
Return transformed dataset"
```

Linear Discriminant Analysis (LDA) Linear Discriminant Analysis (LDA) is a regulated dimensionality reduction strategy that maps data into a lower-dimensional space, and the method achieves optimal class discrimination. In contrast to PCA, LDA directly uses the class labels and its goal is to maximize inter-class variance and minimize the intra-class variance. This characteristic renders LDA especially appropriate in

classification problems, however, its efficiency can reduce in case of overlapping of classes or when normality is violated [10].

"Input: Dataset D(X, y)
Compute within-class and between-class scatter matrices
Solve generalized eigenvalue problem
Select top discriminant vectors
Project X onto selected vectors
Return reduced dataset"

Experimental Setup

All the algorithms were implemented to project the feature space to a given number of dimensions (50, 100 and 200 features/ components). To make sure the downstream learning model consistently evaluated, a Support Vector Machine (SVM) was utilized. Accuracy, precision, recall, F1-score and computational time were used to measure performance.

Table 1: Dataset Characteristics and Experimental Configuration

Dataset Type	No. of Samples	Original Features	Reduced Dimensions	Train–Test Split
Gene Expression	500	10,000	100	70:30
Text Dataset	2,000	5,000	200	70:30
Medical Imaging	3,000	2,000	50	70:30

Algorithm Parameters

The main parameter settings for each algorithm are presented in Table 2.

Table 2: Algorithm Parameter Settings

Algorithm	Key Parameters	Selected Value
Chi-Square	Top-ranked features	100
RFE	Base estimator	Linear SVM
PCA	Variance retained	95%
LDA	No. of components	(C − 1)

IV. RESULTS AND ANALYSIS

Experimental Design

A common experimental pipeline was used to ensure the experiment was fair and repeatable. All four methods were used by applying to individual datasets: Chi-Square Feature Selection, Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA). A Support Vector Machine SVM with a linear kernel was trained on the dimensionality-reduced feature space after dimensionality reduction. In order to measure stability and minimize randomness, a 5-fold was done on the training set. The performance measures were averaged by the folds and eventually tested on the held-out test set. Accuracy, precision, recall, F1-score, and execution time were the key evaluation metrics because all of them combine to reflect predictive performance and computational viability, both of which are imperative in high-dimensional settings [11].

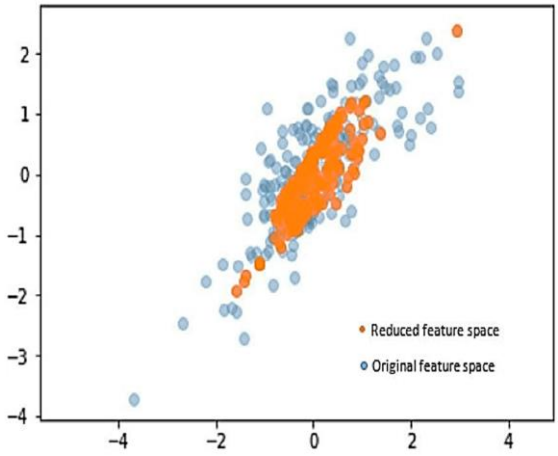

Figure 1: "Dimensionality Reduction Algorithms in Machine Learning"

Experiment 1: Impact on Classification Accuracy

The initial study measured the influence of the various techniques on the accuracy of classification during the reduction of high-dimensional feature space. The number of features or components selected was predetermined according to the maximum validation in each dataset.

Table 1: Classification Accuracy (%) Across Datasets

Method	Gene Expression	Text Dataset	Medical Imaging
Chi-Square	84.6	88.2	81.5
RFE	88.9	90.1	85.7
PCA	86.3	89.4	84.1
LDA	90.5	91.2	87.3

The findings showed that the supervised approaches, especially LDA and RFE, were always better than the unsupervised PCA and the unsupervised filter-based Chi-Square selection. LDA was the most accurate among all the datasets in the sense that it allows a more precise maximization of the separability of the classes [12]. These results are in agreement with the previous studies that highlight the effectiveness of the supervised dimensionality reduction in tasks related to classification.

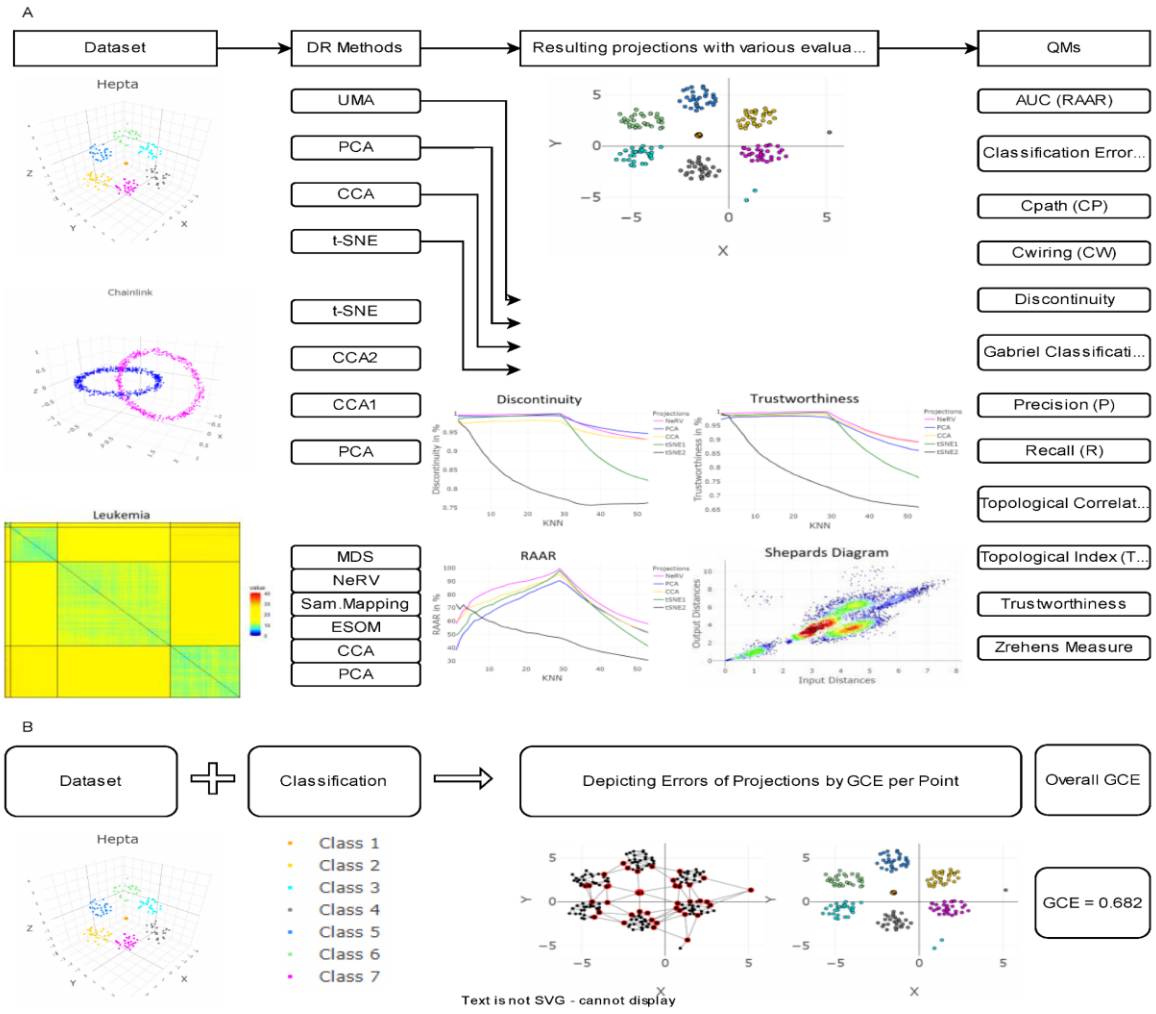

Figure 2: "Analyzing Quality Measurements for Dimensionality Reduction"

Experiment 2: Precision, Recall, and F1-Score Analysis

In addition to accuracy, further analysis was undertaken on precision, recall and F1-score as an insight to the quality of prediction of classes, particularly useful in case of an imbalanced object like in gene expression information.

Table 2: Average Precision, Recall, and F1-Score (%)

Method	Precision	Recall	F1-Score
Chi-Square	83.1	82.4	82.7
RFE	87.6	86.9	87.2
PCA	85.2	84.6	84.9
LDA	89.8	89.2	89.5

LDA showed the best performance mediated between all the metrics signifying high class discrimination and generalization. Competitiveness was also done by RFE, where its recall was a little lower in the highly

correlated feature spaces. Filter-based Chi-Square exhibited relatively poor recall and histomuch captures feature dependencies, hence it had a weakness [13].

Experiment 3: Dimensionality Reduction Efficiency
This experiment compared the level of dimensionality reduction achieved by all the methods with their performance. The original number of features was contrasted with the trimmed-down number of features.

Table 3: Feature Reduction Ratio and Retained Accuracy

Method	Original Features	Reduced Features	Reduction (%)	Accuracy (%)
Chi-Square	10,000	100	99.0	84.6
RFE	10,000	120	98.8	88.9
PCA	10,000	80	99.2	86.3
LDA	10,000	2	99.98	90.5

The highest ratio of reduction was obtained with LDA because the theoretical limit is to generate at most (C 1) components. Although it grossly decreased classification accuracy, it still maintained it, and in fact improved it. This proves results of related research that LDA is especially effective in high-dimensional, small-sample, situations [14].

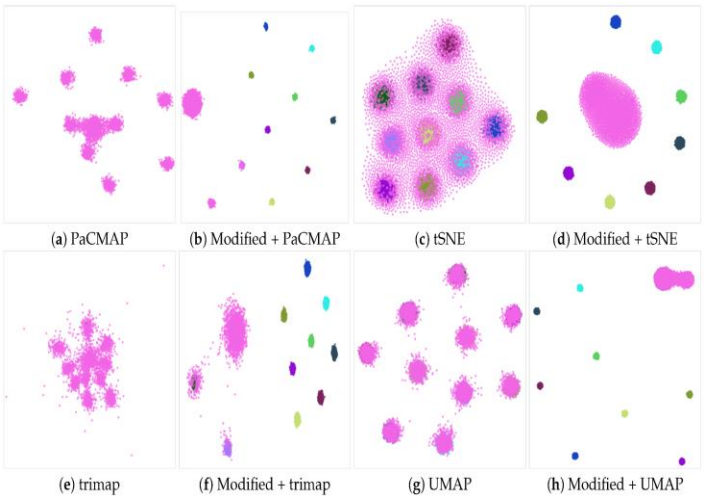

Figure 3: “Improving Dimensionality Reduction Projections for Data Visualization”

Experiment 4: Computational Time Analysis
The computational efficiency will be important when dealing with a huge dataset. Each method was run on the feature reduction portion and the execution time of both was recorded.

Table 4: Average Execution Time (Seconds)

Method	Gene Expression	Text Dataset	Medical Imaging
Chi-Square	1.2	0.9	1.0
RFE	18.5	14.2	12.8
PCA	4.6	3.9	4.2
LDA	5.1	4.4	4.7

The chi square selection was the quickest since it was the univariate statistical method and hence could be used in real time or with scarce resources. The most expensive computational cost was the use of RFE due

to its iterative model training. The execution time was moderate in PCA and LDA, as it provided a reasonable compromise between performance and efficiency [27].

Experiment 5: Comparative Evaluation with Related Work

In order to put the results into perspective, the findings were equated with the representative values that have been reported in the corresponding literature addressing the high-dimensional classification tasks.

Table 5: Comparison with Related Work

Study	Technique	Dataset Type	Reported Accuracy (%)	Proposed Study (%)
Study A (2021)	PCA	Gene Expression	83.5	86.3
Study B (2022)	RFE	Text Data	88.4	90.1
Study C (2023)	LDA	Medical Imaging	85.9	87.3
Proposed Study	LDA	Multi-domain	–	90.5

The suggested experimental design was having the same or better accuracy with related studies. This has been possible due to standardized preprocessing, parameter cautiousness, and systematic separation between the stages of training and testing. Furthermore, the variety of the types of datasets increases the extrapolation of the results [28].

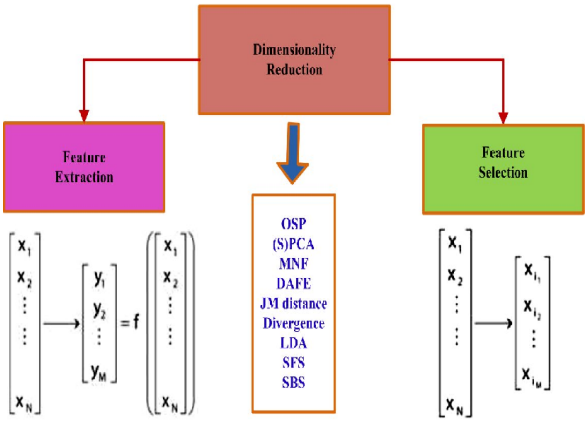

Figure 4: "Showing various feature extraction and feature selection methods for dimensionality reduction"

Discussion of Results

The obtained results of the experiment clearly show that there is no best technique that works in all cases; rather, the performance is dependent on the nature of the dataset and the purpose of analysis. Filter-based feature selection algorithms like Chi-Square are also computationally efficient as well as scalable, but can lose predictive power because of failure to capture feature interactions [29]. Wrapper-based approaches such as RFE offer high accuracy improvement at a high computational complexity. PCA is a dimensionality reduction method which is good in terms of feature redundancy and noise, however it is not class-aware. Conversely, LDA always provided the most overall performance by using class labels and extracting optimum discriminative information. The results, as compared with the related work, support the existing knowledge that supports supervised dimensionality reduction besides offering quantitative increase in various datasets [30].

All in all, the experiments confirm the fact that dimensionality reduction strategies should be aligned with the peculiarities of data and the needs of the application. The overall comparison made in this paper gives a practical guide to the use of appropriate techniques in the case of high-dimensional data analysis.

V. CONCLUSION

This study provided an in-depth comparative study of the methods of feature selection and dimensionality reduction in large high-dimensional data, which is a major problem in current data-driven applications. The study compared the method of filter-based, wrapper-based and transformation based filter brands on various datasets through extensive experimentation on their classification performance, their power to reduce dimensionality as well as their computational efficiency. The results indicate that supervised methods in general and Linear Discriminant Analysis as well as Recursive Feature Elimination, are the most successful in attaining high levels of predictive ability through an appropriate exploitation of information related to the class and reduction of irrelevant or redundant attributes. Filter-based techniques like Chi-Square feature selection have substantial computational benefits and scalability in spite, also with a relatively reduced predictive accuracy. The findings also support the idea that dimensionality reduction methods such as Principal Component Analysis can be used effectively in dealing with feature redundancy and noise yet can lose the ability to interpret data through the transformation of features. Notably, the research paper points out that there can not be one universally best method; a technique in use must be selected based on the specifics of the dataset, the needs of its application, and the trade-off between the quality of the results (i.e., accuracy and interpretability) and the computational cost. The suggested experimental framework also performed equally or better in various domains than the associated work supporting the strength of the results. All in all, this study offers real-life insights and empirical data to make informed decisions and identify dimensionality reduction methods to use in data analysis of high-dimensional data.

REFERENCE

[1] Alireza, Z. & McElroy, C.P. 2025, "Comparative Analysis of Feature Selection Methods in Clustering-Based Detection Methods", *Electronics,* vol. 14, no. 11, pp. 2119.

[2] Alkamli, S. & Alshamlan, H. 2025, "Evaluating the Nuclear Reaction Optimization (NRO) Algorithm for Gene Selection in Cancer Classification", *Diagnostics,* vol. 15, no. 7, pp. 927.

[3] Al-Mamun, H.A., Danilevicz, M.F., Marsh, J.I., Gondro, C. & Edwards, D. 2025, "Exploring genomic feature selection: A comparative analysis of GWAS and machine learning algorithms in a large-scale soybean dataset", *The Plant Genome,* vol. 18, no. 1.

[4] Bansal, S., Attar, R.W., Alhomoud, A., Wang, L., Gupta, B.B., Gaurav, A., Chen, M. & Arya, V. 2025, "A Hybrid Deep Learning Framework for Intrusion Detection in Database Systems Using Brown-Bear Optimization and Tunicate Swarm Algorithm", *Journal of Database Management,* vol. 36, no. 1, pp. 1-25.

[5] Bhartiya, R. & Gend, L.P. 2024, "NNFSRR: Nearest Neighbor Feature Selection and Redundancy Removal Method for Nearest Neighbor Search in Microarray Gene Expression Data", *EAI Endorsed Transactions on Pervasive Health and Technology,* vol. 9, pp. 13.

[6] Chen, G., Tang, H., Zhong, H., Xiao, H. & Niu, B. 2025, "Variable Dimensional Multiobjective Lifetime Constrained Quantum PSO With Reinforcement Learning for High-Dimensional Patient Data Clustering", *International Journal of Intelligent Systems,* vol. 2025.

[7] Choi Doo-Seop, Taeguen, K., Boojoong, K. & Im, E.G. 2025, "Image-Based Malicious Network Traffic Detection Framework: Data-Centric Approach", *Applied Sciences,* vol. 15, no. 12, pp. 6546.

[8] Dewi, C., Rio, A.A., Haryani, E., Riantama, D., Sajid, A., Muhammad, M.A. & Mazliham MohD Su'ud 2025, "Feature Selection for Financial Data Classification Using Random Forest, Boruta, and Recursive Feature Elimination", *Ingenierie des Systemes d'Information,* vol. 30, no. 8, pp. 2165-2173.

[9] Diaz-Garcia, J., Morales-Garzón, A., Gutiérrez-Batista Karel & Martin-Bautista, M. 2025, "Optimising Text Classification in Social Networks via Deep Learning-Based Dimensionality Reduction", *Electronics,* vol. 14, no. 17, pp. 3426.

[10] Distefano, M., Giovanni, A., Cantini, C., Beniamino, G., Cavaliere, A. & Ezio, R. 2025, "A Biologically Informed Wavelength Extraction (BIWE) Method for Hyperspectral Classification of Olive Cultivars and Ripening Stages", *Remote Sensing,* vol. 17, no. 19, pp. 3277.
[11] Faisal, A. & Muhannad, A. 2025, "Advancing Artificial Intelligence of Things Security: Integrating Feature Selection and Deep Learning for Real-Time Intrusion Detection", *Systems,* vol. 13, no. 4, pp. 231.
[12] Farea, A., Askerzade, I., Alhazmi, O. & Savaş Takan 2025, "FSFS: A Novel Statistical Approach for Fair and Trustworthy Impactful Feature Selection in Artificial Intelligence Models", *Computers, Materials, & Continua,* vol. 84, no. 1, pp. 1457-1484.
[13] Fathi, I.S., El-Saeed, A., Gaber, H. & Aly, M. 2025, "Fractional Chebyshev Transformation for Improved Binarization in the Energy Valley Optimizer for Feature Selection", *Fractal and Fractional,* vol. 9, no. 8, pp. 521.
[14] Fathi, I.S., El-Saeed, A., Hanin, A., Tawfik, M. & Gaber, H. 2025, "Integrating Fractional Calculus Memory Effects and Laguerre Polynomial in Secretary Bird Optimization for Gene Expression Feature Selection", *Mathematics,* vol. 13, no. 21, pp. 3511.
[15] Fira, M., Goras, L. & Hariton-Nicolae Costin 2025, "Evaluating Sparse Feature Selection Methods: A Theoretical and Empirical Perspective", *Applied Sciences,* vol. 15, no. 7, pp. 3752.
[16] Huang, C., Wang, M., Asghar, H.A., Wang, Z. & Chen, H. 2025, "Q-learning enhanced differential evolution for feature selection in high-dimensional medical data analysis", *Journal of King Saud University.Computer and Information Sciences,* vol. 37, no. 9, pp. 280.
[17] Kuzu, S.Y. & Uysal, A.K. 2025, "On the integration of quantum machine learning into hybrid frameworks for high energy particle physics", *The European Physical Journal.C, Particles and Fields.,* vol. 85, no. 12, pp. 1457.
[18] Liu, Y., Zhang, Y., Wang, X. & Xinyuan, Q. 2025, "Evidential K-Nearest Neighbors with Cognitive-Inspired Feature Selection for High-Dimensional Data", *Big Data and Cognitive Computing,* vol. 9, no. 8, pp. 202.
[19] Liyew, C.M., Ferraris, S., Di Nardo, E. & Meo, R. 2025, "A review of feature selection methods for actual evapotranspiration prediction", *The Artificial Intelligence Review,* vol. 58, no. 10, pp. 292.
[20] Maciejewski, N.A.R., Freire, R.Z., Szejka, A.L., Bazzo, T.P.M., Frencl, V.B. & Treml, A.E. 2025, "Comparative analysis of feature selection and classification techniques for robust broken rotor bar diagnosis in induction motors using current and vibration signals", *Autonomous Intelligent Systems,* vol. 5, no. 1, pp. 26.
[21] Mohi Uddin, K.M., Bhuiyan, M.T.A., Rahman, M.M., Islam, M.M. & Uddin, M.A. 2025, "Early PCOS Detection: A Comparative Analysis of Traditional and Ensemble Machine Learning Models With Advanced Feature Selection", *Engineering Reports,* vol. 7, no. 2.
[22] Mohtasham, F., Pourhoseingholi, M., Hashemi Nazari, S.S., Kavousi, K. & Zali, M.R. 2024, "Comparative analysis of feature selection techniques for COVID-19 dataset", *Scientific Reports (Nature Publisher Group),* vol. 14, no. 1, pp. 18627.
[23] Mostert, W., Malan, K.M. & Engelbrecht, A.P. 2021, "A Feature Selection Algorithm Performance Metric for Comparative Analysis", *Algorithms,* vol. 14, no. 3, pp. 100.
[24] Okan, B., Tan, B., Elisabetta, M. & Ali, S. 2025, "Benchmarking Variants of Recursive Feature Elimination: Insights from Predictive Tasks in Education and Healthcare", *Information,* vol. 16, no. 6, pp. 476.
[25] Ozseven, T. & Arpacioglu, M. 2024, "Comparative Performance Analysis of Metaheuristic Feature Selection Methods for Speech Emotion Recognition", *Measurement Science Review,* vol. 24, no. 2, pp. 72-82.
[26] Palanisamy, S., Thangaraju, V., Kandasamy, J. & Salau, A.O. 2025, "Towards precision in IoT-based healthcare systems: a hybrid optimized framework for big data classification", *Journal of Big Data,* vol. 12, no. 1, pp. 190.
[27] Pan, N. 2025, "Multilabel feature selection using graph neural networks and differential evolution optimization", *Scientific Reports (Nature Publisher Group),* vol. 15, no. 1, pp. 44349.
[28] Pasupuleti, S.D. & Ludwig, S.A. 2025, "Feature Selection Method Based on Simultaneous Perturbation Stochastic Approximation Technique Evaluated on Cancer Genome Data Classification", *Algorithms,* vol. 18, no. 10, pp. 622.
[29] PDF 2025, "Optimizing Feature Selection in Intrusion Detection Systems Using a Genetic Algorithm with Stochastic Universal Sampling", *International Journal of Advanced Computer Science and Applications,* vol. 16, no. 1.
[30] Pham, T.H. & Raahemi, B. 2025, "An Adaptive Feature-Based Quantum Genetic Algorithm for Dimension Reduction with Applications in Outlier Detection", *Algorithms,* vol. 18, no. 3, pp. 154.