InSight Bulletin: A Multidisciplinary Interlink International Research Journal
Peer Reviewed International, Open Access Journal.
ISSN : 3065-7857 / Website: https://ibrj.us / Volume-3, Issue-1 / January - 2026

Original Article

A Comparative Analysis of Embedded and Wrapper Feature Selection Methods for High-Dimensional Data

Pooja Ravi Kamble[1], Kiran Hrushikesh Chavan [2], Ratnamala Ravindra Borole[3]
[1,2,3] Computer Science, Assistant Professor, ATSS College of Business Studies and Computer Application, Chinchwad, Pune

Manuscript ID:
IBMIIRJ -2026-030132

Submitted: 09 Dec. 2025

Revised: 13 Dec. 2025

Accepted: 08 Jan. 2026

Published: 31 Jan. 2026

ISSN: 3065-7857

Volume-3

Issue-1

Pp. 170-172

January 2026

Correspondence Address:
Pooja Ravi Kamble, Computer Science, Assistant Professor, ATSS College of Business Studies and Computer Application, Chinchwad, Pune
Email: poojakamble@atsscollege.edu.in

OPEN ACCESS

Quick Response Code:

Web. https://ibrj.us

DOI: 10.5281/zenodo.18954429

DOI Link:
https://doi.org/10.5281/zenodo.18954429

Abstract

The rapid increase in data volume across modern applications has resulted in datasets with very high dimensionality, creating challenges such as increased computational cost, overfitting, and limited interpretability. Feature selection addresses these issues by identifying a subset of informative attributes while preserving the original meaning of the data [1], [10]. Among existing techniques, wrapper and embedded methods are widely used due to their strong integration with predictive models [2], [7]. This paper presents a comparative analysis of wrapper and embedded feature selection approaches for high-dimensional datasets, focusing on their principles, computational complexity, scalability, and applicability across domains. The discussion draws on established theoretical foundations and empirical studies to provide guidance for selecting suitable feature selection strategies in complex learning environments.

Keywords: *Feature Selection, Embedded Methods, Wrapper Methods, High-Dimensional Data, Dimensionality Reduction, Machine Learning, Model Performance, Overfitting, Regularization, LASSO, Random Forest, Recursive Feature Elimination (RFE), Cross-Validation, Computational Complexity, Predictive Modeling, Data Mining.*

Introduction

Recent advances in data acquisition and storage technologies have led to the availability of high-dimensional datasets in domains such as bioinformatics, healthcare, text mining, image processing, and IoT systems [3], [5]. Although a large number of features can capture complex patterns, many features are often irrelevant or redundant, which negatively impacts learning performance [4], [6]. High dimensionality increases computational complexity and raises the risk of overfitting, particularly when the number of samples is limited [3]. Feature selection has emerged as an effective preprocessing step to mitigate these challenges by reducing dimensionality without transforming the original feature space [1], [10]. By selecting only relevant features, models become more efficient, interpretable, and robust. Feature selection techniques are commonly categorized into filter, wrapper, and embedded methods [11]. While filter methods are computationally efficient, wrapper and embedded methods are often preferred in supervised learning due to their higher predictive accuracy [7], [10]. This paper focuses on a comparative study of wrapper and embedded feature selection techniques, highlighting their strengths, limitations, and suitability for high-dimensional datasets.

Background

High-dimensional datasets are characterized by a feature space that is significantly larger than the number of observations. This situation leads to the well-known curse of dimensionality, which adversely affects learning algorithms [3]. Increased dimensionality results in higher computational costs and memory requirements, making many algorithms impractical for large-scale applications [2]. Another major concern is overfitting, where models fit noise rather than meaningful patterns [6]. This issue reduces generalization performance and reliability, especially in critical applications such as medical diagnosis [5]. Additionally, models trained on large feature sets often lack interpretability, limiting their adoption in decision-sensitive environments.

How to cite this article:
Kamble, P. R., Chavan, K. H., & Borole, R. R. (2026). A Comparative Analysis of Embedded and Wrapper Feature Selection Methods for High-Dimensional Data. InSight Bulletin: A Multidisciplinary Interlink International Research Journal, 3(1), 170–172. https://doi.org/10.5281/zenodo.18954429

Feature selection addresses these issues by eliminating irrelevant and redundant features while maintaining interpretability [1]. Wrapper and embedded methods are particularly effective because they evaluate feature relevance in conjunction with model performance [7], [8]. However, these methods differ significantly in terms of scalability and computational requirements, necessitating a systematic comparison.

Methodology

Feature selection techniques are broadly classified into three categories [10], [11]:

A. Filter Methods

Filter methods rank features using statistical measures such as correlation, mutual information, or variance, independent of any learning algorithm [4], [13].

B. Wrapper Methods

Wrapper methods evaluate subsets of features by training and testing a predictive model, selecting the subset that yields the best performance [7].

C. Embedded Methods

Embedded methods integrate feature selection into the model training process itself, typically using regularization or intrinsic feature importance mechanisms [8], [9].

This study concentrates on wrapper and embedded methods due to their effectiveness in supervised learning tasks involving high-dimensional data.

Wrapper Feature Selection Methods

Wrapper methods treat feature selection as a search problem, where different feature subsets are evaluated using a learning algorithm [7]. Common search strategies include forward selection, backward elimination, recursive feature elimination, and metaheuristic approaches such as genetic algorithms [12], [15].

Wrapper methods are known for their ability to capture complex feature interactions and often achieve high predictive accuracy [7]. However, they are computationally expensive and do not scale well to extremely high-dimensional datasets, increasing the risk of overfitting [2].

Embedded Feature Selection Methods

Embedded methods perform feature selection as part of the learning process itself [8]. Regularization-based techniques such as Lasso and Elastic Net enforce sparsity by shrinking less important feature coefficients toward zero [8]. Tree-based models, including decision trees and random forests, provide feature importance measures during training [9]. Compared to wrapper methods, embedded approaches are computationally more efficient and scalable, making them suitable for large feature spaces [10]. Nevertheless, their performance may be influenced by model-specific assumptions [17].

Comparative Analysis

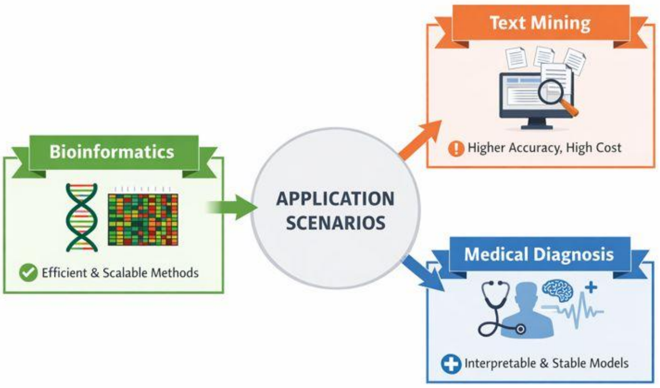
Wrapper methods generally achieve high accuracy when strong feature interactions exist but incur significant computational costs due to repeated model training [7]. Embedded methods reduce this overhead by integrating feature selection into training, resulting in better scalability and generalization for high-dimensional datasets [8], [9].

Performance in High-Dimensional Datasets

Empirical studies in bioinformatics and text classification demonstrate that wrapper methods perform well for moderately sized feature sets, whereas embedded methods are more practical for datasets with thousands of features [5], [14]. Hybrid approaches combining embedded and wrapper techniques have been proposed to balance accuracy and efficiency [10], [18].

Application Scenarios

Wrapper methods are suitable when accuracy is the primary concern and computational resources are sufficient. Embedded methods are preferred for large-scale applications requiring efficiency and interpretability.



Discussion

The choice between wrapper and embedded feature selection methods depends on dataset size, feature dimensionality, computational resources, and interpretability requirements [10], [11]. No single method is universally optimal, highlighting the need for context-aware selection strategies.

Conclusion

This paper presented a comparative analysis of wrapper and embedded feature selection methods for high-dimensional datasets. While wrapper methods offer strong predictive performance, their scalability limitations restrict their applicability to very large feature spaces. Embedded methods provide a balanced trade-off between efficiency and robustness, making them more suitable for large-scale applications. The insights presented in this study can assist researchers and practitioners in selecting appropriate feature selection techniques for complex data analysis tasks.

Acknowledgement

I would like to express my sincere gratitude to my supervisor for their invaluable guidance, insightful feedback, and continuous encouragement throughout the development of this study on embedded and wrapper feature selection methods for high-dimensional data. Their expertise and constructive suggestions significantly improved the quality and clarity of this work. I am also deeply thankful to the faculty members and colleagues who provided thoughtful discussions, technical support, and academic resources that contributed to the successful completion of this research. Special appreciation goes to my peers for their collaboration, motivation, and shared learning experiences.

Financial support and sponsorship

Nil.

Conflicts of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. Guyon and A. Elisseeff, "An introduction to variable and feature selection," Journal of Machine Learning Research, vol. 3, pp. 1157–1182, 2003.
2. H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 4, pp. 491–502, 2005.
3. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed., Springer, 2009. M. Dash and H. Liu, "Feature selection for classification," Intelligent Data Analysis, vol. 1, no. 1–4, pp. 131–156, 1997. Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," Bioinformatics, vol. 23, no. 19, pp. 2507–2517, 2007.
4. L. Yu and H. Liu, "Efficient feature selection via analysis of relevance and redundancy," Journal of Machine Learning Research, vol. 5, pp. 1205–1224, 2004.
5. R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artificial Intelligence, vol. 97, no. 1–2, pp. 273–324, 1997. J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models," Journal of Statistical Software, vol. 33, no. 1, 2010.
6. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
7. Chandrashekar and F. Sahin, "A survey on feature selection methods," Computers & Electrical Engineering, vol. 40, no. 1, pp. 16–28, 2014.
8. J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," CRC Press, 2014.
9. P. Pudil et al., "Floating search methods in feature selection," Pattern Recognition Letters, vol. 15, no. 11, pp. 1119–1125, 1994.
   M. A. Hall, Correlation-Based Feature Selection for Machine Learning, PhD Thesis, Univ. of Waikato, 1999. G. Forman, "An extensive empirical study of feature selection metrics for text classification," Journal of Machine Learning Research, vol. 3, pp. 1289–1305, 2003.
10. S. Maldonado and R. Weber, "A wrapper method for feature selection using SVM," Information Sciences, vol. 179, no. 13, pp. 2208–2217, 2009.
11. K. Kira and L. Rendell, "A practical approach to feature selection," Proc. ICML, 1992.
12. J. Ang et al., "Supervised, unsupervised, and semi-supervised feature selection," Pattern Recognition, vol. 47, no. 12, pp. 4074–4095, 2014.
13. Z. Zhao et al., "On similarity preserving feature selection," IEEE Transactions on Knowledge and Data Engineering, vol. 25, no. 3, pp. 619–632, 2013.