STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, February 2026, pp 1000–1011.
Published online in International Academic Press (www.IAPress.org)

Comparison of Filter Techniques for Feature Selection in High-dimensional Data

Safaa Bouamira*, Hasna Chamlal, Tayeb Ouaderahman

Computer Science and Systems Laboratory (LIS), Department of Mathematics and Computer Science, Faculty of Sciences Ain Chock, University Hassan II of Casablanca, Morocco

Abstract Feature selection constitutes a fundamental challenge within machine learning, which has garnered heightened attention owing to the proliferation of high-dimensional datasets. Filtering-based feature selection methods hold crucial importance as they can be seamlessly integrated with any machine learning model and significantly accelerate the runtime of such algorithms. This study investigates the performance of eight distinct filter methods, examining their efficacy across nine high-dimensional datasets, the classification accuracy was assessed through the employment of support vector machines and k-nearest neighbor classifiers, and the Wilcoxon test statistic was applied to confirm the observed results regarding classification accuracy.

Keywords Feature Selection, Filters, High Dimensional Datasets

AMS 2010 subject classifications 62-07, 97K80, 68T20, 65C60

DOI: 10.19139/soic-2310-5070-2548

1. Introduction

Feature selection has emerged as a crucial component of data analysis and machine learning, particularly for applications in diverse fields such as healthcare and credit scoring [1]. In high-dimensional data, feature selection is essential to identify and remove irrelevant and redundant features by selecting a suitable subset of relevant features, this helps mitigate the risk of overfitting and addresses the challenges posed by the curse of dimensionality.

Over the past several decades, numerous feature selection techniques have been introduced, these methods can be broadly categorized into three distinct groups: filter-based feature selection methods, which initially conduct a filtering process on the feature variables and subsequently train the classifier using a reduced subset of features; wrapper feature selection approaches, which require a predefined classifier to identify a feature subset that is most conducive to the learning performance of this classifier by utilizing the classification performance as an evaluation metric for the significance of the features; and embedded feature selection techniques, which intertwine the feature selection process with the classifier training process, whereby feature selection and classification learning are concurrently executed within the same optimization framework.

Filters are generally based on criteria that can be used to measure the feature relevance [2] [3], [4] or redundancy [5], and the wrapper models investigated encompass Bat-inspired algorithms [6], Moth flame algorithms [7]. Predictive methods incorporating embedded feature selection techniques include Lasso regression [8] and Elastic net [9]. Moreover, various hybrid feature selection approaches that combine wrappers and filters have been proposed, such as a specific pre-ordonnances-based memetic algorithm [10], a maximal cliques-based hybrid method with interaction screening [11], and a graph partitioning-based hybrid feature selection method [12].

*Correspondence to: Safaa Bouamira (Email: safaa.bouamira-etu@etu.univh2c.ma)

ISSN 2310-5070 (online) ISSN 2311-004X (print)

This study examines the performance of eight filter-based feature selection approaches on high-dimensional classification datasets, the evaluated filter methods, including univariate and multivariate techniques, represent prominent general strategies for filter-based feature selection, the key aspects of this research are as follows:

- A comparative analysis of four univariate and four multivariate filter methods was conducted.
- The performance of the filter methods was evaluated using leave-one-out cross-validation accuracy with SVM and KNN classifiers.
- The number of selected features ranged from 2 to half the number of samples for each dataset, with increments of 2.
- The Wilcoxon signed-rank test was performed to confirm statistically significant differences in classification accuracy between the filter methods.
- The time complexity of each filter was evaluated.

The present study is structured as follows: Section 2 provides an overview of various filtering approaches, and Section 3 elaborates on the experimental setup used to compare and evaluate the performance of these filtering methods. Finally, Section 4 summarizes the findings and presents the concluding remarks of this investigation.

2. Filter methods

We describe two types of filter methods: univariate filters, which do not take into account interactions among features, and multivariate filters, which do account for feature interactions:

2.1. Univariate filters:

- Fisher score [13]: is a supervised method that ranks features based on their association with the class variable, this measure prioritizes features that bring instances of the same class closer together while separating instances from different classes. Given a dataset $X \in \mathcal{R}^{n\times p}$ associated with m distinct classes, let μ_k and $(\sigma_k)^2$ represent the mean and variance, respectively, of the j th feature for class k, the Fisher score for the j th feature is defined as:

$$F\left(X^j\right) = \frac{\sum_{k=1}^{m} n_k \left(\mu_k^j - \mu^j\right)^2}{\sum_{k=1}^{m} n_k \left(\sigma_k^j\right)^2} \tag{1}$$

  where n_k is the number of samples in class k, and μ^j is the overall mean of the jth feature

- Information Gain(IG) [14], [15]:this measure can be used to quantify the relevance between variables X and Y, the higher the IG value, the stronger the discriminative power of the explanatory variable X. IG can be calculated as follows:

$$IG\left(X_k\right) = I\left(Y; X_k\right) = H(Y) - H(Y \mid X_k) \tag{2}$$

  Here $H(Y)$ is the entropy of Y: $H(Y) = -\sum_y p(y)\log_2(p(y))$, p is the probability mass function, and the conditional entropy of Y given X_k is given by: $H(Y \mid X_k) = \sum_x p(x)\left(-\sum_y p(y \mid x)\log_2(p(y \mid x))\right)$.

- Symmetric uncertainty (SU) [16]: this technique is employed to modify the Information Gain metric, mitigating the bias toward variables with numerous distinct values and scaling the IG to the range of 0 to 1, a value of 0 indicates an independent relationship between X and Y, while a value of 1 denotes a stronger dependency relationship between them. SU is calculated as follows:

$$SU(X_k, Y) = \frac{2 \times I(X_k, Y)}{H(X_k) + H(Y)} \tag{3}$$

- Chi-square (Chi2) [17]: this statistical approach quantifies the divergence between expected and observed distributions of a given feature, the larger the value of this statistical indicator, the more robust the association between the feature and the class label. The formula for calculating this metric is provided in the subsequent equation:

$$\chi^2 = \sum_{i,j} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \tag{4}$$

where O_{ij} and E_{ij} represent the observed and expected distributions respectively

2.2. *Multivariate filters:*

- Mutual Information Maximization (MIM) [18]: is a method that iteratively calculates feature scores, allowing us to rank the features based on their Mutual Information Maximization scores. We can then select the top K features for our study, K is determined as half the number of samples:

$$MIM\left(X_k\right) = \max_{j \in \{1,\ldots,p\}} I\left(Y; X_j\right) \tag{5}$$

- Joint Mutual Information (JMI) [19]: quantifies the informative value that the joint input variables X_k and X_j provide about the target variable Y, this metric is calculated by pairing the candidate X_k with each previously selected feature. The underlying premise is that if the candidate feature is complementary to the existing features, it should be included in the model:

$$JMI\left(X_k\right) = \sum_{X_j \in S} I\left(Y; X_k, X_j\right) \tag{6}$$

  S denotes the set of already chosen features
- Minimum Redundancy Maximum Relevance (MRMR) [20]: aims to identify a set of features that are highly relevant to the target variable Y while minimizing redundancy among the selected features:

$$MRMR\left(X_k\right) = I\left(Y; X_k\right) - \frac{1}{|S|} \sum_{X_j \in S} I\left(X_k; X_j\right) \tag{7}$$

- ReliefF [21], [22]: randomly selects an instance R_i, and then identifies k of its nearest neighbors from the same class H_j as well as k nearest neighbors from each of the different classes M_j, the score calculated by ReliefF for each feature is updated based on the values for R_i, H_j hits, and M_j misses. Finally, the feature score is defined as:

$$\begin{aligned} W[A] = W[A] - \sum_{j=1}^{k} \frac{\phi\left(A, R_i, H_j\right)}{m \cdot k} \\ + \sum_{C \neq c \text{ class } R_i} \left[\rho \sum_{j=1}^{k} \frac{\phi\left(A, R_i, M_j(C)\right)}{m \cdot k} \right] \\ \text{with } \rho = \frac{P(C)}{1 - P\left(class\left(R_i\right)\right)} \end{aligned} \tag{8}$$

  $\phi\left(A, R_i, H\right)$ is defined as the distance between instance R_i and its nearest hit H and m denotes the user-specified number of iterations

3. Experiments and results

The study evaluated the effectiveness of each filter method, investigating their predictive capability across the top-ranked features, which ranged from 2 to half the samples for each dataset. Additionally, a statistical analysis was performed. The experimental programs were implemented using Python, with several key packages utilized, including ”sklearn”, ”pandas”, and ”matplotlib”, all filter methods were accessible through the ”skfeature” package, the computing environment was a Microsoft Windows 11 system with an Intel Core i7-7600U CPU running at 2.80GHz and 16GB of RAM.

3.1. Datasets

The study utilized 9 high-dimensional datasets for the experimental analyses, these datasets comprised both binary classification and multiclass, the number of classes ranged from 2 to 5, while the number of features varied between 2000 and 12600. Table I provides a concise description of these datasets:

Table 1. The data utilized in the study are described herein

Datasets	**Observations**	**Features**	**Classes**
Colon [23]	62	2000	2
SRBCT [24]	83	2308	4
Leukemia [25]	72	7129	2
Lymphoma [26]	66	4026	3
CNS [27]	60	7129	2
DLBCL [28]	77	5469	2
MLL [29]	72	12582	3
Prostate [30]	102	12600	2
Lung [31]	203	12600	5

3.2. Results

To examine the impact of the number of selected features on the performance of the filter methods, Figures 1, 2, 3, and 4 present the accuracies for various subset sizes, ranging from two features to half the number of samples, with an increment of two for the univariate and multivariate filters, two popular classifiers, the Support Vector Machine, and K-Nearest Neighbor were employed to evaluate the leave-one-out cross-validation accuracy.

Based on the evidence in Figures 1 and 2 and Table 2, the study utilized an SVM classifier with a linear kernel. For the univariate filter methods, the Chi-squared technique demonstrated the highest accuracy across the evaluated datasets. In terms of the multivariate filters, ReliefF exhibited the best performance in the Lymphoma, DLBCL, MLL, and Lung datasets, while the MIM approach gave the optimal results in four datasets, namely Colon, DLBCL, Prostate, and Lung. Subsequently, the MRMR method showcased the highest accuracy in the SRBCT and Leukemia datasets, and the JMI technique performed best in the Colon and DLBCL datasets.

Figures 3, 4, and Table 3 demonstrate that the K-Nearest Neighbors classifier with a k parameter of five achieved the highest accuracy through the univariate Chi-squared method across the evaluated datasets. Regarding the multivariate filtering approaches, the MIM filter exhibited the greatest accuracy in the Colon, SRBCT, Lymphoma, DLBCL, and Lung datasets. The ReliefF algorithm yielded the optimal performance in the Colon, Leukemia, MLL, and Lymphoma datasets. Additionally, the MRMR method demonstrated the highest performance in four datasets: Colon, Lymphoma, CNS, and Prostate. The JMI method provided the greatest accuracy solely in the CNS dataset.

The results show that the Chi-square method has the highest accuracy among the univariate and multivariate filter techniques across all datasets, when using both the K-Nearest Neighbors and Support Vector Machine classifiers.

3.3. Statistical test

The Wilcoxon non-parametric statistical test [32] is also employed to assess the performance of the 8 filters, this test is implemented to determine the statistical significance of differences between the algorithms, and this test is conducted at a 5% significance level to verify whether there is a statistically significant difference in the accuracy, results obtained, Tables 4 and 5 reported the p-value of the Wilcoxon rank-sum test for the classification accuracy using the SVM and KNN classifiers, with a significance level of $\alpha = 0.05$, according to the SVM results, the Chi2

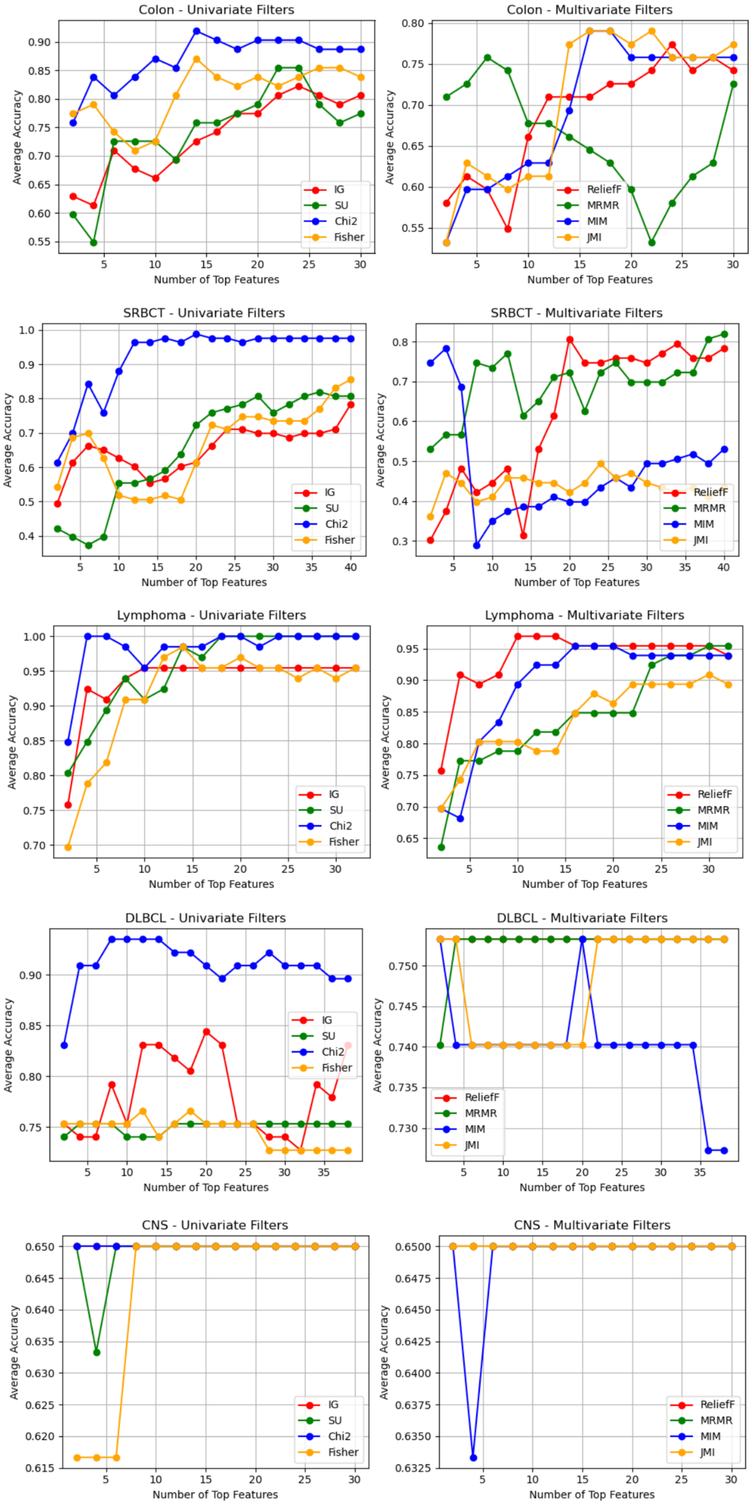

Figure 1. LOOCV accuracy related to the number of top-ranking features for the Univariate and Multivariate filters utilizing SVM classifier.

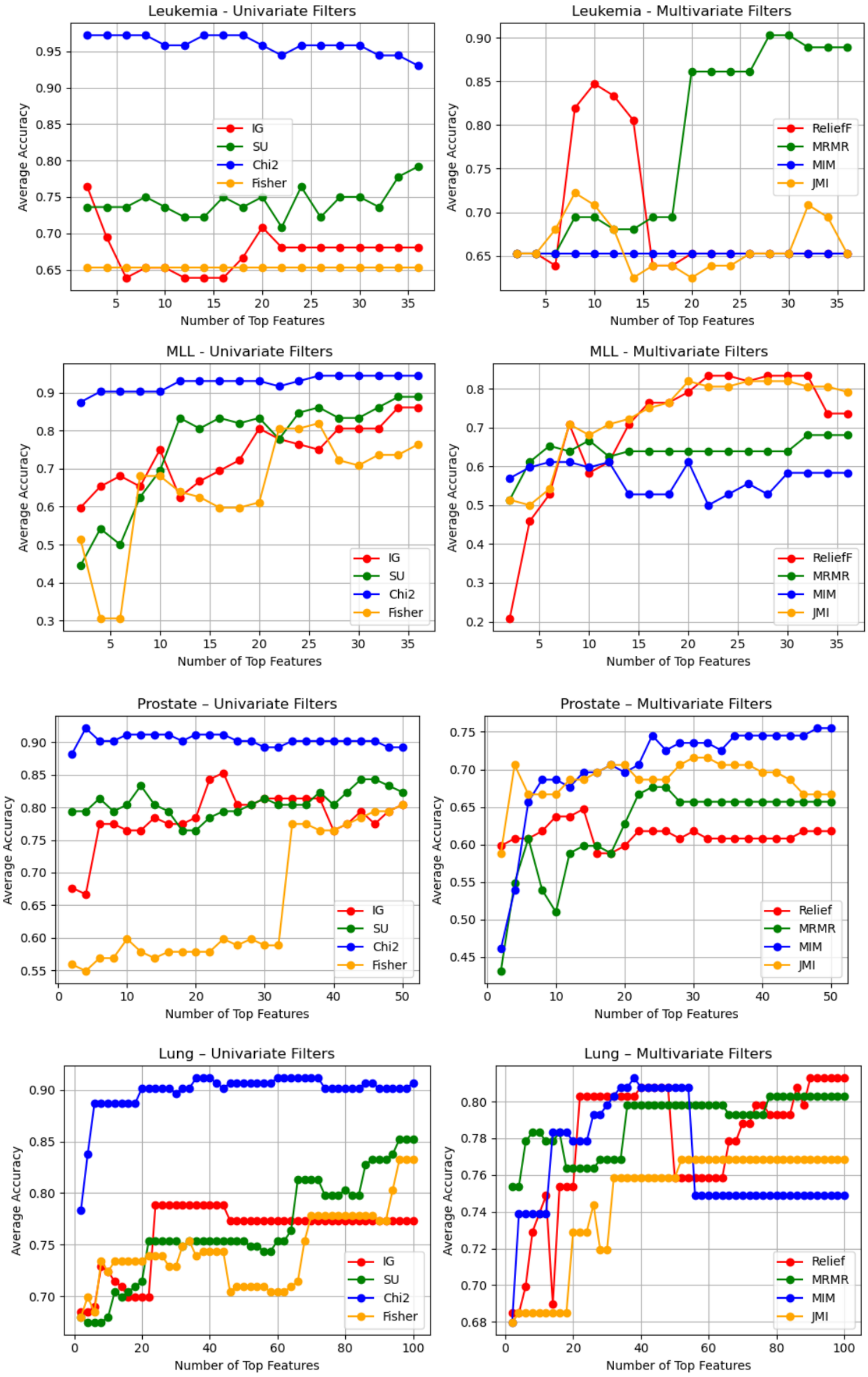

Figure 2. LOOCV accuracy related to the number of top-ranking features for the Univariate and Multivariate filters utilizing SVM classifier

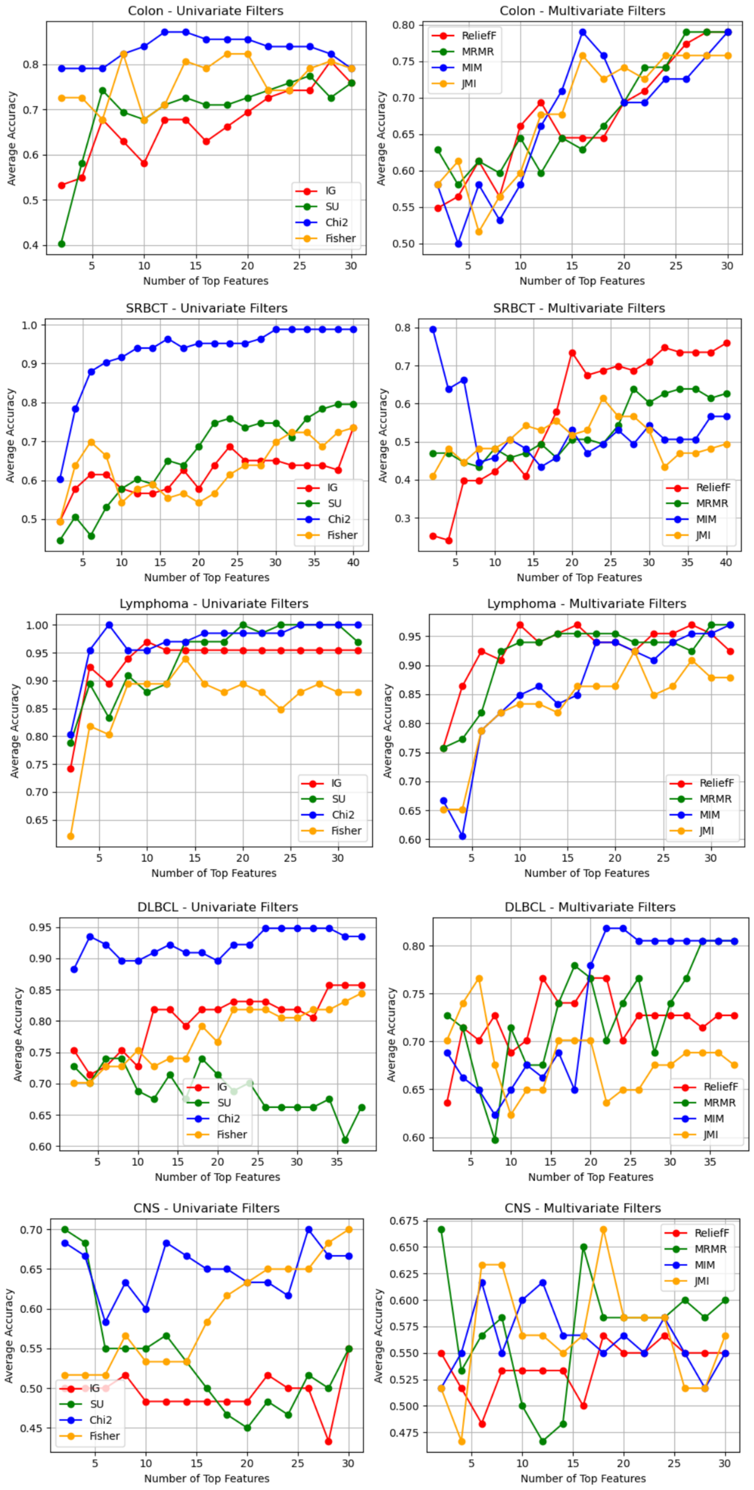

Figure 3. LOOCV accuracy related to the number of top-ranking features for the Univariate and Multivariate filters utilizing KNN classifier.

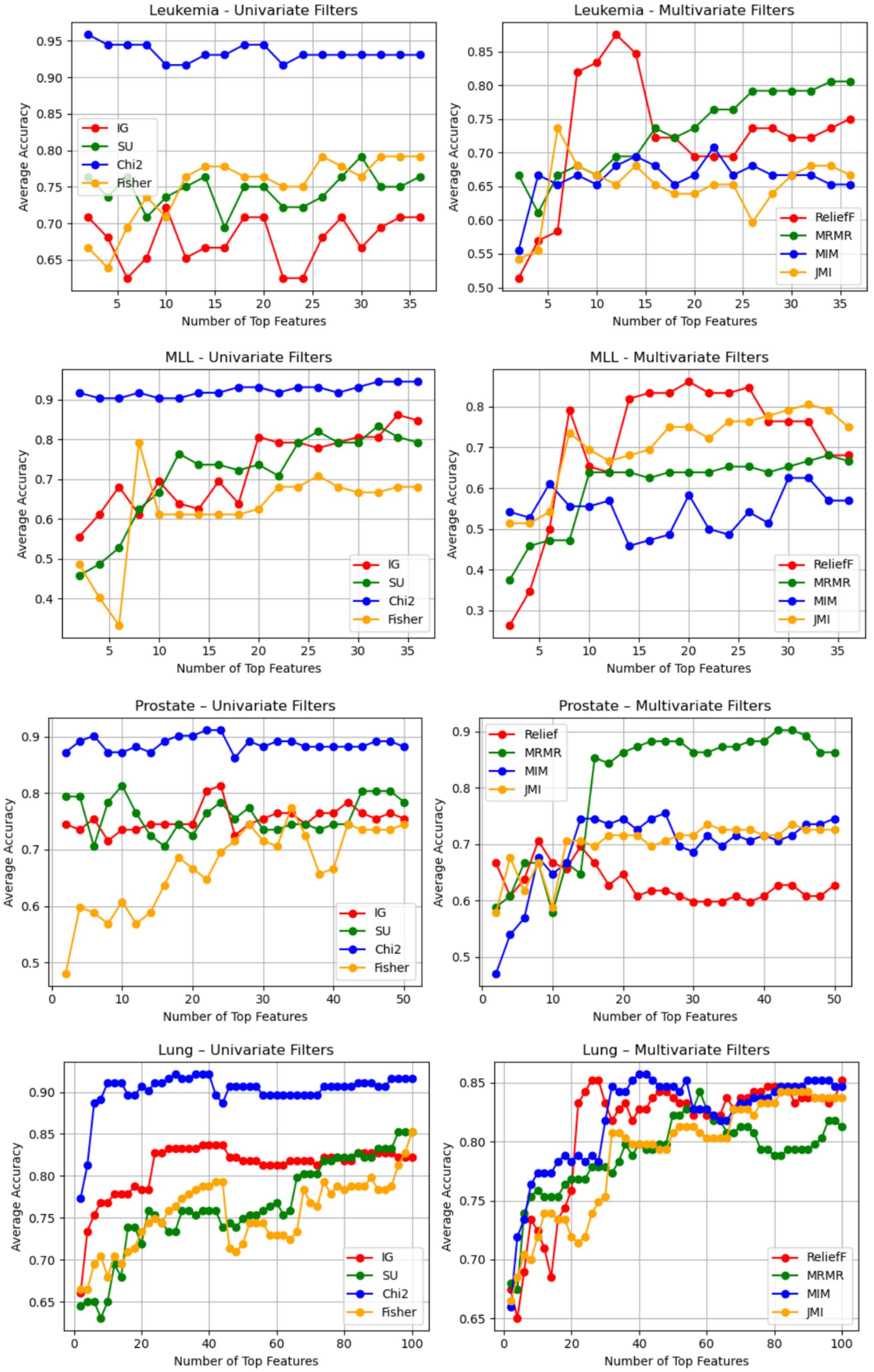

Figure 4. LOOCV accuracy related to the number of top-ranking features for the Univariate and Multivariate filters utilizing KNN classifier

method demonstrated a statistically significant difference in accuracy compared to the other filters, as the $p - value$ was less than 0.05, The results indicate that there is a notable distinction among the Fisher method, MIM, JMI, and MRMR.

Table 2. LOOCV-based analysis of the best performances obtained by the Univariate and Multivariate filters applying the SVM Classifier

Data	**Univariate Filters**								**Multivariate Filters**							
	IG		**SU**		**Chi2**		**Fisher**		**ReliefF**		**MRMR**		**MIM**		**JMI**	
	NF	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**
Colon	24	82	22	85	14	**92**	14	**87**	24	77	6	76	16	79	16	79
SRBCT	40	78	36	82	20	**99**	40	**86**	20	81	40	82	4	78	24	79
Lymphoma	10	95	18	**100**	4	**100**	14	98	10	97	30	95	16	95	30	91
DLBCL	20	**84**	4	75	8	**94**	12	77	2	75	4	75	2	75	2	75
CNS	2	**65**	2	**65**	2	**65**	8	**65**	2	**65**	2	**65**	2	**65**	2	**65**
Leukemia	2	76	36	79	2	**97**	2	65	10	84	28	**90**	2	65	8	72
MLL	34	86	34	**89**	26	**94**	26	82	22	83	32	68	6	54	20	82
Prostate	24	**85**	44	84	4	**92**	50	80	14	65	24	68	48	75	30	72
Lung	24	79	96	**85**	36	**91**	96	83	90	81	78	80	38	81	52	77

Table 3. LOOCV-based analysis of the best performances obtained by the Univariate and Multivariate filters applying the KNN Classifier

Data	**Univariate Filters**								**Multivariate Filters**							
	IG		**SU**		**Chi2**		**Fisher**		**ReliefF**		**MRMR**		**MIM**		**JMI**	
	NF	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**	**NF**	**ACC (%)**
Colon	28	80	26	77	12	**87**	8	**82**	28	79	26	79	16	79	16	76
SRBCT	40	73	38	80	30	**99**	40	73	40	76	28	64	2	**80**	24	61
Lymphoma	10	97	20	**100**	6	**100**	14	94	10	97	30	97	32	97	22	92
DLBCL	34	**86**	6	74	26	**95**	38	84	14	77	34	81	22	82	6	77
CNS	30	55	2	**70**	26	**70**	30	70	18	57	2	67	6	62	18	67
Leukemia	10	72	30	79	2	**97**	26	79	12	**88**	34	81	22	71	6	74
MLL	34	86	32	83	32	**94**	8	79	20	**86**	34	68	30	63	32	81
Prostate	24	81	10	81	22	**91**	34	77	8	71	42	**90**	26	75	32	74
Lung	38	84	96	85	30	**92**	100	85	26	85	58	84	40	**86**	82	84

Table 4. The p-value of the Wilcoxon rank-sum test for the classification accuracy using the SVM with a significance level of $\alpha = 0.05$

	IG	MRMR	ReliefF	SU	MIM	JMI	Fisher	chi2
IG	1	0.0001	0.533	0.348	0.0005	1.9×10^{-6}	0.419	1.9×10^{-6}
MRMR	-	1	0.177	0.011	0.026	1.9×10^{-6}	0.0008	1.9×10^{-6}
ReliefF	-	-	1	0.015	0.026	0.0007	0.277	1.9×10^{-6}
SU	-	-	-	1	0.026	0.0003	0.409	1.9×10^{-6}
MIM	-	-	-	-	1	0.354	8.2×10^{-6}	9.5×10^{-6}
JMI	-	-	-	-	-	1	1.9×10^{-6}	1.9×10^{-6}
Fisher	-	-	-	-	-	-	1	1.9×10^{-6}
Chi2	-	-	-	-	-	-	-	1

Table 5. The p-value of the Wilcoxon rank-sum test for the classification accuracy using the KNN with a significance level of $\alpha = 0.05$

	IG	MRMR	ReliefF	SU	MIM	JMI	Fisher	chi2
IG	1	0.001	0.295	0.025	0.002	1.9×10^{-6}	0.285	1.9×10^{-6}
MRMR	-	1	0.956	0.001	0.015	0.0002	0.002	1.9×10^{-6}
ReliefF	-	-	1	8.2×10^{-5}	0.177	0.082	0.475	1.9×10^{-6}
SU	-	-	-	1	0.008	1.9×10^{-6}	0.107	1.91×10^{-6}
MIM	-	-	-	-	1	0.825	0.002	5.7×10^{-6}
JMI	-	-	-	-	-	1	0.0001	1.9×10^{-6}
Fisher	-	-	-	-	-	-	1	1.9×10^{-6}
Chi2	-	-	-	-	-	-	-	1

However, the Fisher method exhibits similarities to IG, ReliefF, and SU in terms of accuracy, as the obtained p-value was greater than 0.05, the JMI approach differs considerably from the other filter in terms of accuracy, yet it exhibits substantial similarity to the MIM filter, for the K-Nearest Neighbors classifier, we obtained comparable results to the Support Vector Machine classifier for Chi2 and Fisher filters, however, in this case, the JMI filters yielded similar performance to the SU method in terms of accuracy.

3.4. *Time Complexity*

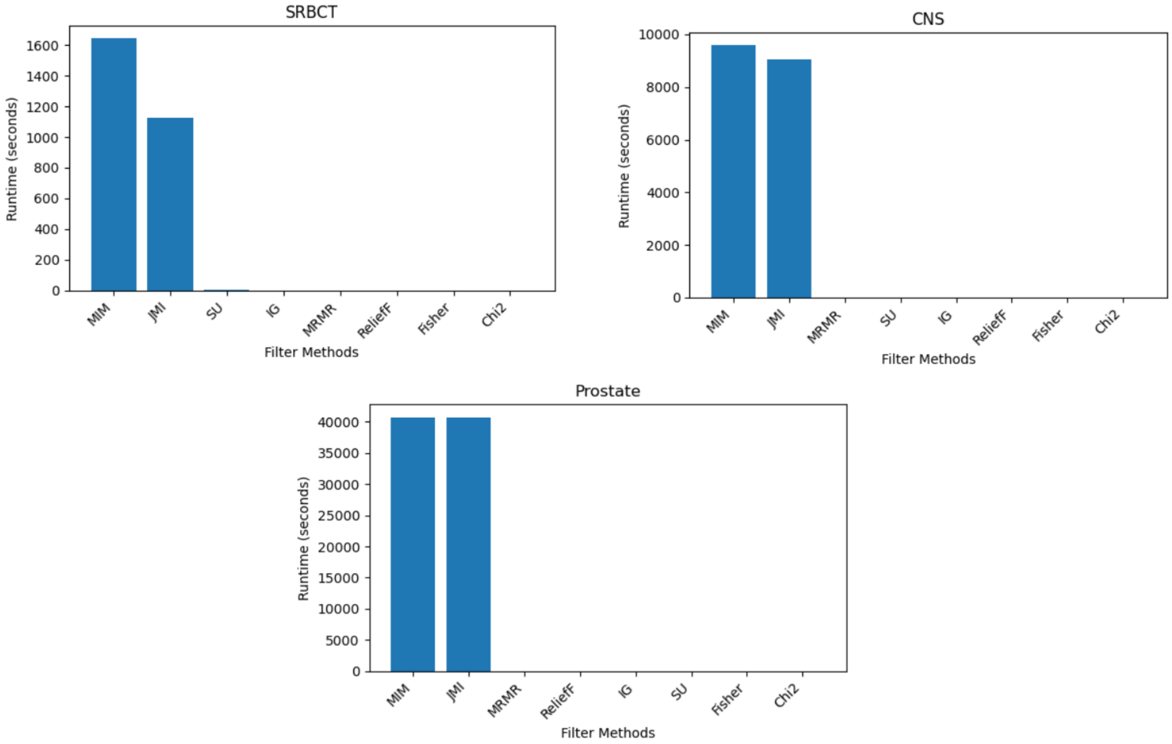

Figure 5. Time Complexity for each filter method

As shown in Figure 5 and Table 6, the runtime of the various filter methods on the SRBCT, CNS, and Prostate datasets reveals that the Chi-square and Fisher methods are the most efficient, whereas the JMI and MIM methods are the slowest. Furthermore, it can be observed that the time complexity of each filter increases as the number of features in the high-dimensional datasets grows.

Table 6. Filters runtimes on SRBCT, CNS, and Prostate datasets

SRBCT		**CNS**		**Prostate**	
Method	Runtime (s)	Method	Runtime (s)	Method	Runtime (s)
MIM	1642.302 663	MIM	9581.541 956	MIM	40 755.687 815
JMI	1125.761 858	JMI	9064.390 909	JMI	40 751.800 790
SU	1.682 174	MRMR	2.597 293	MRMR	7.053 448
IG	1.190 196	SU	1.565 046	ReliefF	6.810 991
MRMR	1.120 888	IG	1.481 679	IG	4.537 986
ReliefF	0.625 770	ReliefF	1.405 622	SU	4.156 924
Fisher	0.599 950	Fisher	1.403 822	Fisher	2.555 962
Chi2	0.013 952	Chi2	0.032 838	Chi2	0.065 533

4. Conclusion

This study investigated eight feature selection filters using nine high-dimensional datasets, the performance of these filters was evaluated based on the leave-one-out cross-validation accuracy. The Wilcoxon test was used as a statistical test to ensure the reliability of the results, our study generally found that the Chi-squared method outperformed all other approaches using the SVM and KNN classifiers, In perspective, we will assess the performance of those feature selection filters on larger, high-dimensional datasets such as GLI and SMK, which have over 20,000 features, in the following work, we focus on proposing a filter based on a new relevance measure that can be used even for high-dimensional heterogeneous data.

REFERENCES

1. MEHDI, Bazzi, HASNA, Chamlal, TAYEB, Ouaderhman, et al. Intelligent credit scoring system using knowledge management. IAES International Journal of Artificial Intelligence, 2019, vol. 8, no 4, p. 391.
2. AABOUB, Fadwa, CHAMLAL, Hasna, et OUADERHMAN, Tayeb. Analysis of the prediction performance of decision tree-based algorithms. In : 2023 International Conference on Decision Aid Sciences and Applications (DASA). IEEE, 2023. p. 7-11.
3. CHAMLAL, Hasna, AABOUB, Fadwa, et OUADERHMAN, Tayeb. A preordonance-based decision tree method and its parallel implementation in the framework of Map-Reduce. Applied Soft Computing, 2024, vol. 167, p. 112261.
4. Aaboub F, Chamlal H, Ouaderhman T. Statistical analysis of various splitting criteria for decision trees. Journal of Algorithms & Computational Technology. 2023;17. doi:10.1177/17483026231198181.
5. OUADERHMAN, Tayeb, CHAMLAL, Hasna, et JANANE, Fatima Zahra. A new filter-based gene selection approach in the DNA microarray domain. Expert Systems with Applications, 2024, vol. 240, p. 122504.
6. O. A. Alomari, A. T. Khader, M. A. Al-Betar, and L. M. Abualigah, “Gene selection for cancer classification by combining minimum redundancy maximum relevancy and bat-inspired algorithm”, 2017, vol. 19, no 1, p. 32-51.
7. A. Dabba, A. Tari, S. Meftali, and R. Mokhtari, “Gene selection and classification of high dimensional data method based on mutual information and moth flame algorithm.,” Expert Systems with Applications, vol. 166, p. 114012, Mar. 2021, doi: 10.1016/j.eswa.2020.114012.
8. R. Tibshirani, “Regression Shrinkage and Selection Via the Lasso,” Journal of the Royal Statistical Society Series B: Statistical Methodology, vol. 58, no. 1, pp. 267–288, Jan. 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.
9. H. Chamlal, A. Benzmane, and T. Ouaderhman, “Elastic net-based high dimensional data selection for regression,” Expert Systems with Applications, vol. 244, p. 122958, Jun. 2024, doi: 10.1016/j.eswa.2023.122958.
10. CHAMLAL, Hasna, OUADERHMAN, Tayeb, et EL MOURTJI, Basma. Feature selection in high dimensional data: a specific preordonnances-based memetic algorithm. Knowledge-based systems, 2023, vol. 266, p. 110420.
11. CHAMLAL, Hasna, BENZMANE, Asmaa, et OUADERHMAN, Tayeb. Maximal cliques-based hybrid high-dimensional feature selection with interaction screening for regression. Neurocomputing, 2024, vol. 607, p. 128361.
12. OUBAOUZINE, Abdelali, OUADERHMAN, Tayeb, et CHAMLAL, Hasna. A graph partitioning-based hybrid feature selection method in microarray datasets. Knowledge and Information Systems, 2025, vol. 67, no 1, p. 633-660.
13. L. Sun, X.-Y. Zhang, Y.-H. Qian, J.-C. Xu, S.-G. Zhang, and Y. Tian, “Joint neighborhood entropy-based gene selection method with fisher score for tumor classification,” Appl Intell, vol. 49, no. 4, pp. 1245–1259, Apr. 2019, doi: 10.1007/s10489-018-1320-1.
14. J. R. Quinlan, “Induction of decision trees,” Mach Learn, vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1007/BF00116251.
15. LEE, Junghye, CHOI, In Young, et JUN, Chi-Hyuck. An efficient multivariate feature ranking method for gene selection in high-dimensional microarray data. Expert Systems with Applications, 2021, vol. 166, p. 113971.

16. X. Lin, C. Li, W. Ren, X. Luo, and Y. Qi, “A new feature selection method based on symmetrical uncertainty and interaction gain,” Computational Biology and Chemistry, vol. 83, p. 107149, Dec. 2019, doi: 10.1016/j.compbiolchem.2019.107149.
17. I. Sumaiya Thaseen and C. Aswani Kumar, “Intrusion detection model using fusion of chi-square feature selection and multi class SVM,” Journal of King Saud University - Computer and Information Sciences, vol. 29, no. 4, pp. 462–472, Oct. 2017, doi: 10.1016/j.jksuci.2015.12.004.
18. D. D. Lewis, “Feature selection and feature extraction for text categorization,” in Proceedings of the workshop on Speech and Natural Language - HLT ’91, Harriman, New York: Association for Computational Linguistics, 1992, p. 212. doi: 10.3115/1075527.1075574.
19. H. H. Yang and J. Moody, “Data Visualization and Feature Selection: New Algorithms for Nongaussian Data”.
20. C. Ding and H. Peng, “MINIMUM REDUNDANCY FEATURE SELECTION FROM high dimensional GENE EXPRESSION DATA,” 2005.
21. M. R.-S. Ikonja, M. Robnik, and I. Kononenko, “Theoretical and Empirical Analysis of ReliefF and RReliefF”.
22. JANANE, Fatima Zahra, OUADERHMAN, Tayeb, et CHAMLAL, Hasna. A filter feature selection for high-dimensional data. Journal of Algorithms & Computational Technology, 2023, vol. 17, p. 17483026231184171.
23. U. Alon et al., “Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays,” Proc. Natl. Acad. Sci. U.S.A., vol. 96, no. 12, pp. 6745–6750, Jun. 1999, doi: 10.1073/pnas.96.12.6745.
24. J. Khan et al., “Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks,” Nat Med, vol. 7, no. 6, pp. 673–679, Jun. 2001, doi: 10.1038/89044.
25. T. R. Golub et al., “Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring,” Science, vol. 286, no. 5439, pp. 531–537, Oct. 1999, doi: 10.1126/science.286.5439.531.
26. A. A. Alizadeh et al., “Distinct types of diffuse large B-cell lymphoma identi®ed by gene expression pro®ling,” vol. 403, 2000.
27. S. L. Pomeroy et al., “Prediction of central nervous system embryonal tumour outcome based on gene expression,” Nature, vol. 415, no. 6870, pp. 436–442, Jan. 2002, doi: 10.1038/415436a.
28. Shippet et al., “ Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning,“ Nature medicine, 8(1), 68-74, 2002.
29. S. A. Armstrong et al., “MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia,” Nat Genet, vol. 30, no. 1, pp. 41–47, Jan. 2002, doi: 10.1038/ng765.
30. SINGH, Dinesh, FEBBO, Phillip G., ROSS, Kenneth, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer cell, 2002, vol. 1, no 2, p. 203-209.
31. BHATTACHARJEE, Arindam, RICHARDS, William G., STAUNTON, Jane, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proceedings of the National Academy of Sciences, 2001, vol. 98, no 24, p. 13790-13795.
32. J. Hamidzadeh and M. Kelidari, “Robust Feature Selection by Filled Function and Fisher Score,” Feb. 07, 2022. doi: 10.21203/rs.3.rs-1102788/v1.