Applied Computer Science 2026, vol. 22, no. 2, pp. 169–179
https://doi.org/10.35784/acs_9078

Submitted: 2026-01-25 | Revised: 2026-04-29 | Accepted: 2026-05-14

Keywords: classification, data normalization, extreme learning machine, microarray data

Inggih PERMANA [1, 2], *Shir Li WANG* [1*], *Hoi Yeh LEE* [1], *Suliana SULAIMAN* [1], *Hasnatul Nazuha HASSAN* [1]
[1] Universiti Pendidikan Sultan Idris, Malaysia, inggihpermana@uin-suska.ac.id, shirli_wang@meta.upsi.edu.my, suliana@meta.upsi.edu.my, nazuha@meta.upsi.edu.my
[2] Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia, inggihpermana@uin-suska.ac.id
[*] Corresponding author: shirli_wang@meta.upsi.edu.my

Data normalization methods on microarray data

Abstract

Data normalization is a critical preprocessing step in machine learning, particularly for high-dimensional, low-sample-size datasets such as microarray data. However, the effectiveness of a normalization technique depends on both the dataset's characteristics and the classification algorithm employed. This study investigates the impact of eight data normalization methods on the classification performance of the Extreme Learning Machine (ELM) classifier on microarray data. The evaluated methods include Vector Normalization (L2 Normalization), Quantile Normalization with Gaussian and Uniform distributions, Maximum Absolute Scaling, Z-score Normalization, Min-Max Normalization, Power Transformation, and Robust Scaling. Experiments were conducted on four benchmark microarray datasets: Leukaemia 2-Class, Leukaemia 3-Class, Leukaemia 4-Class, and Small Round Blue Cell Tumour (SRBCT). The experimental framework incorporated Information Gain feature selection, SMOTE-based data balancing, and 10-fold cross-validation. To improve the reliability of the results, each experiment was repeated 30 times with different random seeds, and performance was evaluated using Accuracy and F1-score. The results demonstrate that Vector Normalization consistently achieved the best classification performance across all datasets. It attained the highest accuracy and F1-score in Leukaemia 2-Class (73.74% and 70.38%), Leukaemia 3-Class (75.83% and 70.83%), Leukaemia 4-Class (68.55% and 58.93%), and SRBCT (64.30% and 59.62%). Furthermore, Wilcoxon statistical tests confirmed that the performance improvements obtained by Vector Normalization were statistically significant at the 0.05 significance level. These findings indicate that preserving the original structure and direction of feature vectors through L2-based Normalization is more effective than distribution-based scaling methods for heterogeneous microarray data. Therefore, Vector Normalization can be considered a robust and effective preprocessing technique for ELM-based classification of microarray data.

1. INTRODUCTION

Preprocessing is a crucial step for an effective machine learning process. One task in data preprocessing is data normalization. Data normalization, also known as feature scaling or standardization, is a technique that scales features to balance their numerical range (Demircioğlu, 2024). Normalizing data can help eliminate the effects of inconsistent ranges in some machine learning techniques (Ahmed et al., 2022). The difference in value across attributes causes malfunctions in attributes with significantly lower values. Therefore, it is necessary to normalize the data to equalize the value range for each attribute on a common scale, resulting in well-normalized data (Henderi et al., 2021). In machine learning, data normalization is used to organize, enhance, and simplify scattered data within a dataset, while also reducing the impact of the attribute's value range on each record (Çelik, 2024).

Comparing normalization methods is essential because no single method is superior for all scenarios. Previous researchers have compared normalization methods across various cases. For instance, Demircioğlu (2024) compared seven normalization methods across 15 radiomics datasets to assess their impact on predictive performance. The findings indicated that normalization can yield a relatively small performance improvement, but the magnitude of this effect depends heavily on the dataset's characteristics. Another study on the impact of data normalization was conducted by Lima and Souza (2023). They conducted a comprehensive study of 10 normalization methods across 38 time-series datasets. Their findings revealed that the Z-score, which had been considered the best standard for time-series data for three decades, does not always yield optimal performance. Other alternatives, such as maximum absolute scaling and mean normalization, are also worth considering as effective normalization techniques for time-series data. Additionally, Ahsan et al. (2021) compared normalization methods in terms of their impact on machine learning algorithm performance. Ahsan et al. (2021) used six normalization methods in their study and found that data normalization does not always improve performance; rather, it depends on the combination of the algorithm and the normalization method used. However, this study had a limitation: it relied on a single dataset, making its findings context-specific and less generalizable. Based on the findings from these three studies, comparing normalization methods is considered essential to identify the most suitable method for the specific task and dataset characteristics. This process can determine whether data normalization is truly necessary and, if so, which normalization method is most relevant and effective to apply.

Furthermore, several studies have specifically compared the predictive performance of normalized data with that of unnormalized data. For instance, Borkin et al. (2019) reported findings that contradict common assumptions, demonstrating that normalization actually reduced XGBoost's classification performance. However, this study was limited to a single, relatively small dataset and a single normalization method, min-max normalization. Consequently, Borkin et al. (2019)'s research is insufficient as a universal reference on the effects of normalization; it is more appropriately viewed as a limited case study that highlights the need for empirical verification for each combination of dataset, algorithm, and data pre-processing technique. Similarly, Ahmed et al. (2022) investigated the effect of data normalization on predictive performance using various combinations of datasets and algorithms. Their findings indicated that the impact of normalization depends heavily on the type of algorithm and the dataset's characteristics. Nevertheless, this study had one limitation: it used only one normalization method, namely min-max normalization. Additionally, Çelik (2024) proved that applying data normalization can significantly improve classification accuracy. This study, however, was also limited by its use of only one normalization technique (min-max normalization) and a single dataset. This makes the research less comprehensive and less suitable as a general reference for other contexts and classification algorithms. Despite the limitations of these three studies, which use only a single normalization method, they have consistently shown that the impact of normalization is highly dependent on the specific case at hand, as it interacts with dataset characteristics, the algorithm, and the normalization method applied.

This study evaluates the effect of using various normalization methods on the performance of microarray data classification. Microarray technology allows researchers to measure the expression levels of thousands of genes simultaneously. This data has several unique and challenging characteristics. It is high-dimensional (Daoud & Mayo, 2019), has a small sample size (Bolón-Canedo et al., 2014), and often exhibits an unbalanced class distribution. The combination of these characteristics makes data normalization a crucial step before classification. This study compares eight normalization methods: Vector Normalization (L2 Normalization), Quantile Normalization (Gaussian Distribution), Quantile Normalization (Uniform Distribution), Maximum Absolute Scaling, Z-score Normalization, Min-Max Normalization, Power Transformation, and Robust Scaling. In the existing literature, min-max normalization has been the dominant choice in microarray classification research, as evidenced by studies from Baliarsingh et al. (2021), Hira and Bai (2022), Panda et al. (2025), Sucharita et al. (2024) and Tripathy et al. (2024; 2025). However, no comprehensive study has yet explicitly compared the impact of these different normalization methods on the classification performance of microarray data. Therefore, this research aims to comprehensively investigate the effect of these eight normalization methods on classification performance and identify the most effective method.

This study employs the Extreme Learning Machine (ELM) algorithm as a classifier. ELM is a variant of Single-hidden Layer Feedforward Networks (SLFNs) first proposed by Huang et al. (2006) and Wang et al. (2022). This algorithm was chosen for its significant computational advantage: extremely fast training without iterative adjustment of weight vectors (Huang et al., 2015). This makes ELM highly suitable for classifying microarray data, which are characterized by high dimensionality and a small number of data points, as it can substantially reduce computational complexity and training time. Consequently, ELM facilitates an efficient and effective analysis process on data with these characteristics. Additionally, ELM has been widely applied in recent studies for the classification of microarray data. The effectiveness of this algorithm has been demonstrated in a range of related studies, as evidenced by various works (Baliarsingh et al., 2021; Hira & Bai, 2022; Sucharita et al., 2024; Tripathy et al., 2024; 2025; Nagra et al., 2024).

This paper is structured as follows. The next section, Methods, details the experimental procedure, the datasets used, the eight data normalization methods compared, and the experimental setup. Subsequently, the Results and Discussion section presents the findings from the comparative experiments and discusses the impact of each normalization method on model performance. Finally, the Conclusions and Future Work section summarises the main findings of this study and outlines potential directions for future research.

2. METHODS

This section details the research methodology used to achieve the study's objectives. It begins by outlining the General Research Steps. The Datasets and the various Data Normalization Methods applied are then described. Finally, the Experimental Setup explains the configurations and parameters used in this study.

2.1. General research steps

The experimental procedure was conducted for each combination of a microarray dataset and a data normalization method, as illustrated in Figure 1. This method employs k-fold cross-validation, in which the dataset is divided into k subsets (folds) that are used in turn for training and testing. Subsequently, the training data undergoes a series of preprocessing steps as follows:

1. Feature selection

A filter approach was employed for feature selection, utilizing the Information Gain method. This method was previously employed by Sucharita et al. (2024) as an ensemble model in a study involving microarray data and the ELM algorithm.

2. Data balancing

Following feature selection, the training data was balanced using the Synthetic Minority Over-sampling Technique (SMOTE). The choice of SMOTE was based on the findings of Salisah et al. (2025), which showed its superior performance for microarray data classification.

3. Data normalization

The final preprocessing step involved normalizing the training data using the method under investigation. The data normalization methods employed will be discussed in more detail in the next subsection.

The ELM model was then trained on the pre-processed training data. The resulting model was subsequently evaluated using the testing data. It is important to note that the test data underwent feature selection and normalization using parameters derived from the preceding training steps. The evaluation metrics used were Accuracy and F1-score. Accuracy measures the proportion of correct predictions among all predictions, while the F1-Score is the harmonic mean of precision and recall, providing a balanced measure of performance, especially for imbalanced datasets. The formulas for accuracy and F1-score are shown in Formulas 1 and 2. To ensure the reliability and stability of the results, each dataset-normalization method combination was run 30 times with different random seeds (1, 2, ..., 30). The final performance was reported as the mean values of accuracy and F1-score across all experiments. In addition, to assess the stability of model performance, the standard deviation of each metric was computed for each dataset-method combination. Furthermore, the Wilcoxon statistical test was used to assess significant differences in accuracy between the methods, thereby providing stronger evidence of the observed performance differences.

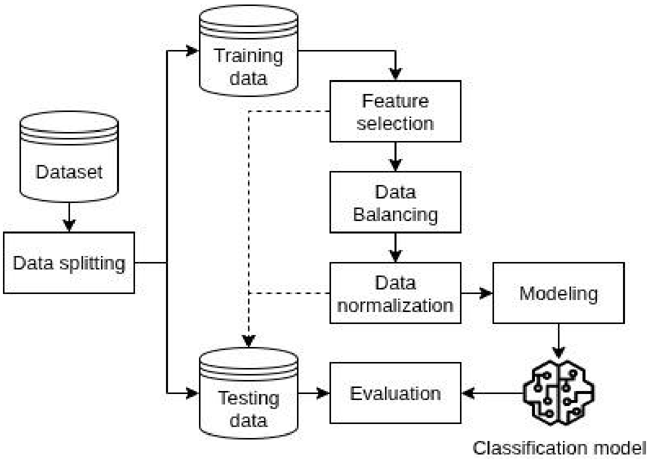


Fig. 1. Experimental procedure for each combination of a microarray dataset and data normalization method

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives} \tag{1}$$

$$F1 = 2 \times \frac{Precision + Recall}{Precision \times Recall} \tag{2}$$

To calculate the F1-Score, one must first determine the precision and Recall values, as shown in Formulas 3 and 4.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{3}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{4}$$

2.2. Data normalization methods

Based on the introduction, this study implements eight data normalization methods used in previous research on the impact of data normalization on predictive performance. A detailed explanation for each method is presented as follows:

1. Vector normalization

Vector Normalization, also known as L2 Normalization, is a method that normalizes each data sample to a unit vector. This process is accomplished by dividing each feature value in a sample by its L2-norm, which is the square root of the sum of the squares of its elements. This normalization approach has been implemented in previous studies, including that conducted by Ahsan et al. (2021). The mathematical formula for Vector Normalization is presented in Formula 5.

$$\hat{v} = \frac{v}{||v||_2} \tag{5}$$

In Formula 5, v represents the original data vector. This vector is divided by its length, or L2 normalization (denoted as $||v||_2$), to produce a new vector, $\hat{v}$, which has a magnitude of 1.

2. Quantile normalization (Gaussian distribution)

Quantile normalization with a Gaussian distribution transforms data to conform to a normal (Gaussian) distribution. The process involves converting each feature's value into its rank (quantile) and then mapping it to the corresponding values of a standard normal distribution. This technique proves highly effective for non-normally distributed data. Lima and Souza (2023) and De Amorim et al. (2023) have employed this method as one of the Normalisation methods.

3. Quantile normalization (Uniform distribution)

Similar to Gaussian quantile normalization, this method also transforms raw data to ranks, then maps these values to a uniform distribution between 0 and 1. This method is useful when the assumption of a Gaussian distribution is not applicable. This method has been employed by Lima and Souza (2023), Demircioğlu (2024), and Ahsan et al. (2021) for data normalization.

4. Maximum Absolute Scaling

Maximum Absolute Scaling is a data normalization method that scales each feature by dividing its value by the feature's maximum absolute value. Its purpose is to transform the data into a range of [-1, 1]. This method is beneficial because it does not shift the data distribution and is less affected by outliers compared to Z-score normalization. Several studies, including those by Lima and Souza (2023), Ahsan et al. (2021), and De Amorim et al. (2023), have used this normalization method in experiments. The formula for Maximum Absolute Scaling is shown in Formula 6.

$$x_{new} = \frac{x_{old}}{max(|x|)} \tag{6}$$

In Formula 6, x_{new} represents the normalized data value, while x_{old} is the original data value. The denominator, max(|x|), denotes the maximum absolute value of all data points within a single feature (column).

5. Z-Score normalisation

Z-score normalization, also known as standardization, is one of the most frequently used data normalization methods. This technique rescales data so that each feature has a mean of 0 and a standard deviation of 1. This approach is effective for models that are sensitive to data scale. This method has been widely applied, including as a data normalization technique in studies by Lima and Souza (2023), Demircioğlu (2024), Ahsan et al. (2021), and De Amorim et al. (2023). The formula for Z-score Normalization is presented in Formula 7.

$$z = \frac{x-\mu}{\sigma} \tag{7}$$

In Formula 7, z represents the normalized data value (Z-score), x is the original data value, μ is the mean of all data points within a single feature, and σ is the standard deviation of all data points within a single feature.

6. Min-Max normalisation

Min-Max Normalization is a data normalization method that scales each feature to a predefined range, typically between 0 and 1. This normalization is performed by shifting the minimum feature value to 0 and the maximum feature value to 1. Similar to Z-score Normalization, this method is very common in machine learning. Lima and Souza (2023), Demircioğlu (2024), Ahsan et al. (2021), and De Amorim et al. (2023) have used this data normalization method in their experiments. The formula for Min-Max Normalization is presented in Formula 8.

$$x_{new} = \frac{x-x_{min}}{x_{max}-x_{min}} \tag{8}$$

In Formula 8, x_{new} represents the normalized data value, while x is the original data value. The terms x_{min} and x_{max} denote the minimum and maximum values, respectively, of all data points within a single feature.

7. Power transformation

Power Transformation is a normalization method aimed at making data more closely resemble a Gaussian distribution. This data normalization technique works by applying a mathematical transformation (such as square root, log, or Box-Cox) to each feature. This transformation is particularly helpful when the data exhibits high skewness. Several researchers, such as Lima and Souza (2023) and Demircioğlu (2024), have used this method as a data normalization technique in experiments.

8. Robust scaling

Robust Scaling is a data normalization method that is resistant to outliers. This method normalizes data by using the median and quartiles instead of the mean and standard deviation, which are highly sensitive to extreme values. This makes Robust Scaling ideal for datasets known to contain numerous outliers. Demircioğlu (2024), Ahsan et al. (2021), and De Amorim et al. (2023) are examples of research that implemented this as a data normalization method. The formula for Robust Scaling is shown in Formula 9.

$$x_{new} = \frac{x-Q_2}{Q_3-Q_1} \tag{9}$$

In Formula 9, x_{new} represents the scaled data value, while x is the original data value. The numerator is determined by the difference between x and Q_2, where Q_2 is the median (second quartile) of all data points within a given feature. The median is the middle value of a sorted dataset. The denominator, Q_3-Q_1, represents the Interquartile Range (IQR), where Q_1 (the first quartile) is the median of the lower half of the data, and Q_3 (the third quartile) is the median of the upper half.

2.3. Datasets

This study utilizes four microarray datasets previously used by Zhu et al. (2007), which are available for download (X. Zhu, 2018). Table 1 describes the datasets used.

Tab. 1. Dataset descriptions

	Genes	Instances	Classes
Dataset 1: Leukaemia 2-Class	7129	72	2
Dataset 2: Leukaemia 3-Class	7129	72	3
Dataset 3: Leukaemia 4-Class	7129	72	4
Dataset 4: SRBCT	2308	83	4

Although the datasets compiled by Zhu et al. (2007) are nearly two decades old, they remain valuable resources and are still widely utilized in recent studies. For instance, Deng et al. (2023) employed four of the eleven microarray datasets originally introduced by Zhu et al. (2007). In addition, Salisah et al. (2025) utilized three microarray datasets from Zhu et al. (2007), whereas Permana et al. (2025) used the complete dataset set provided by Zhu et al. (2007).

2.4. Experimental setup

For the experiment, this study used k = 10 in k-fold cross-validation. This study used an ELM with a network architecture consisting of 64 hidden nodes and the ReLU activation function. The 64 hidden nodes were used because Permana et al. (2025) used the same number in an ELM-based study that also analyzed microarray data. The ReLU activation function was chosen because it is computationally efficient and has simple derivatives, thereby accelerating training compared to more complex activation functions (Vargas et al., 2021). The implementation was written in Python, with Google Colab as the development platform. All normalization methods used in this study were implemented using the preprocessing module in the Scikit-Learn library.

3. RESULTS AND DISCUSSION

The experimental results on the Leukaemia 2-Class, Leukaemia 3-Class, Leukaemia 4-Class, and SRBCT microarray datasets, as shown in Table 2 and Figures 2 to 5, indicate that the normalization technique significantly affects the performance of the classification models. Among all the methods tested, the Vector Normalization consistently produced the best performance across all datasets, outperforming both the non-normalization method and the other scaling techniques. This improvement is further supported by the Wilcoxon test results in Appendix A, which show that at the significance level $\alpha = 0.05$, all p-values are less than α. Therefore, it can be concluded that the Vector Normalization is statistically significantly different from the other seven normalization methods and from the unnormalized case. Given that Vector Normalization shows better results across data sets, as indicated by the Wilcoxon test, it is significantly better than the seven normalization methods and the no-normalization method for the measurements. These results indicate that this method is not only numerically better but also more stable and consistent across multiple experimental iterations.

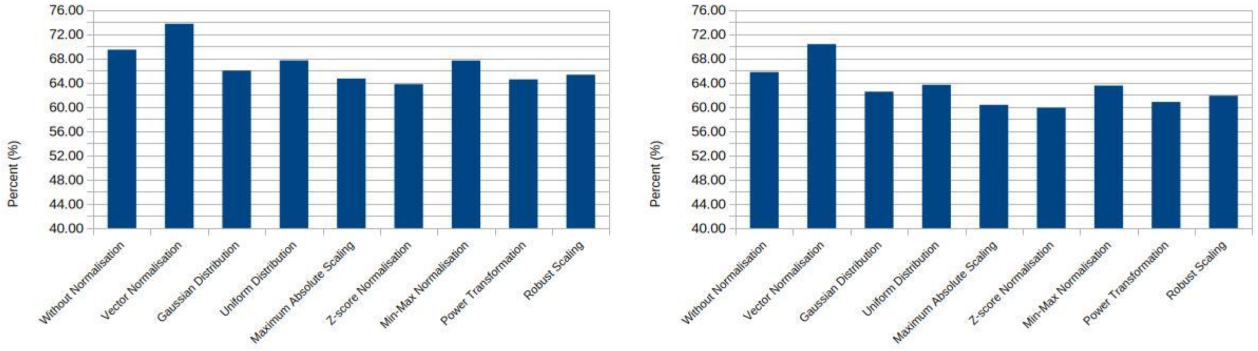
Tab. 2. Experiment results on classification performance with various datasets and data normalisation method combinations

Datasets	Normalisation Methods	Accuracy (%)		F1-Score (%)	
		Avg	Std	Avg	Std
Leukaemia – 2 Class	Without Normalisation	69.46	17.61	65.77	19.68
	Vector Normalisation	**73.74**	**16.51**	**70.38**	**18.77**
	Quantile Normalisation (Gaussian Distribution)	66.01	18.24	62.53	19.59
	Quantile Normalisation (Uniform Distribution)	67.70	17.04	63.67	19.09
	Maximum Absolute Scaling	64.70	17.31	60.36	19.07
	Z-score Normalisation	63.77	17.35	59.88	18.98
	Min-Max Normalisation	67.68	17.59	63.54	19.59
	Power Transformation	64.56	18.92	60.84	20.44
	Robust Scaling	65.33	17.54	61.86	18.89
Leukaemia – 3 Class	Without Normalisation	67.53	19.28	61.92	22.29
	Vector Normalisation	**75.83**	**16.86**	**70.83**	**21.21**
	Quantile Normalisation (Gaussian Distribution)	61.16	19.10	54.27	21.79
	Quantile Normalisation (Uniform Distribution)	67.51	18.26	59.89	21.69
	Maximum Absolute Scaling	61.40	20.17	54.53	22.96
	Z-score Normalisation	57.98	19.36	50.44	21.54
	Min-Max Normalisation	62.55	19.30	55.75	21.89
	Power Transformation	57.10	18.63	49.22	20.32
	Robust Scaling	57.23	19.88	50.14	22.04

Tab. 2. Experiment results on classification performance with various datasets and data normalisation method combinations, continued

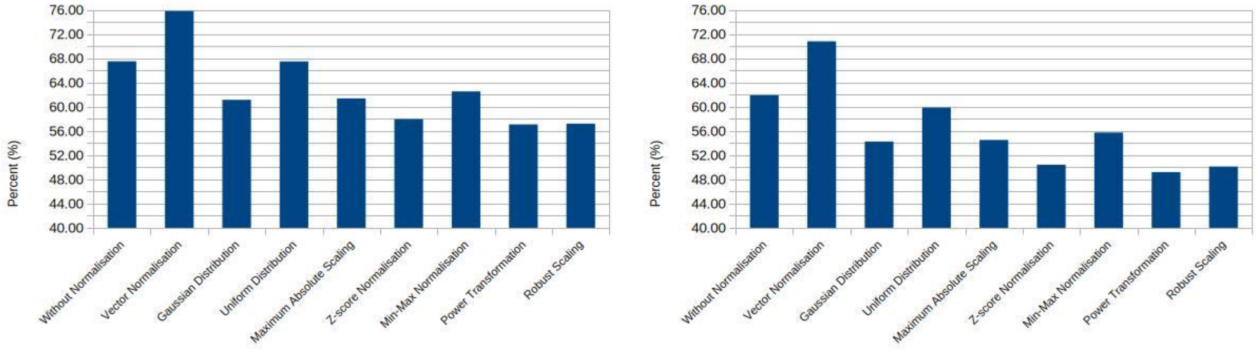
Datasets	Normalisation Methods	Accuracy (%)		F1-Score (%)	
		Avg	Std	Avg	Std
Leukaemia – 4 Class	Without Normalisation	65.04	19.03	55.06	22.63
	Vector Normalisation	**68.55**	**17.39**	**58.93**	**21.34**
	Quantile Normalisation (Gaussian Distribution)	58.89	19.18	46.95	20.72
	Quantile Normalisation (Uniform Distribution)	63.10	17.33	51.45	21.43
	Maximum Absolute Scaling	56.59	18.48	46.39	21.04
	Z-score Normalisation	55.80	18.11	43.99	20.49
	Min-Max Normalisation	60.83	18.49	49.84	21.81
	Power Transformation	55.63	17.49	43.72	19.27
	Robust Scaling	53.83	18.86	42.27	19.41
SRBCT	Without Normalisation	58.03	16.29	52.98	17.97
	Vector Normalisation	**64.30**	**17.49**	**59.62**	**19.70**
	Quantile Normalisation (Gaussian Distribution)	54.46	18.12	50.10	20.05
	Quantile Normalisation (Uniform Distribution)	61.33	16.66	57.56	19.23
	Maximum Absolute Scaling	60.27	16.35	55.81	18.71
	Z-score Normalisation	51.87	17.00	46.86	18.75
	Min-Max Normalisation	58.68	16.36	54.30	18.78
	Power Transformation	57.41	18.86	52.67	20.35
	Robust Scaling	53.28	17.38	47.46	18.67

Nevertheless, all datasets exhibited relatively high standard deviations in performance. This is a common characteristic of high-dimensional, low-sample bioinformatics data, such as microarray data, which have a very large feature-to-sample ratio, making the model sensitive to variations in data distribution and more susceptible to overfitting (Buša & Poļaka, 2021). This performance variability also tends to increase with increasing classification complexity, as in the case of Leukaemia 3-Class, Leukaemia 4-Class, and the SRBCT dataset, which have more complex class distributions. However, despite this high variability, the Vector Normalization maintained more consistent performance than other methods, indicating its greater robustness to changes in data complexity and variations in the training and testing data.



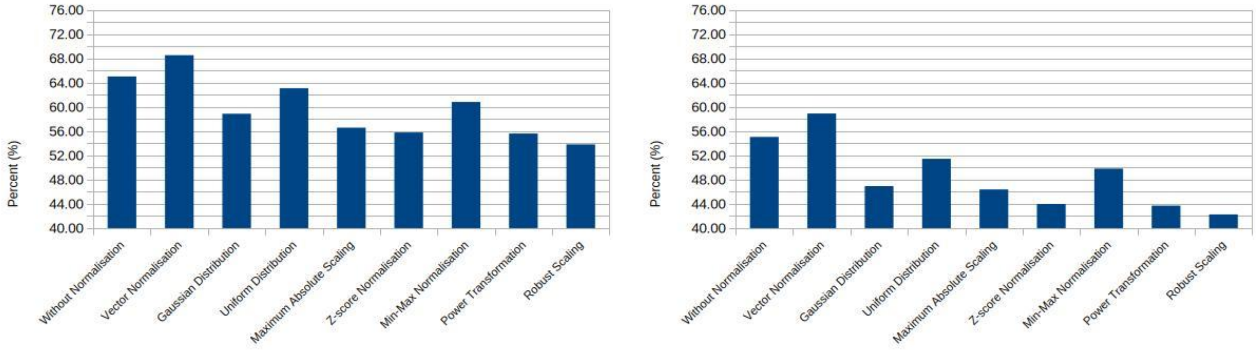
(a) Accuracy **(b) F1-Score**

Fig. 2. Comparison of the performance of normalization methods on the Leukaemia 2-class dataset



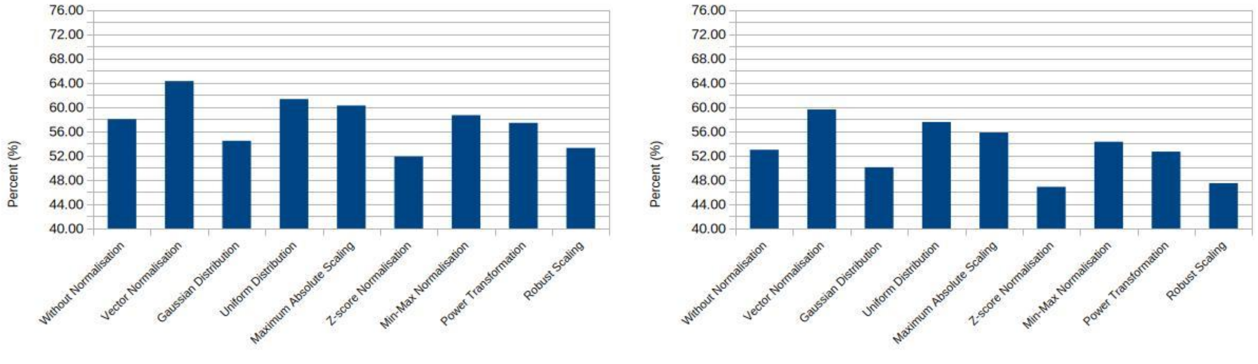
(a) Accuracy **(b) F1-Score**

Fig. 3. Comparison of the performance of normalization methods on the Leukaemia 3-class dataset



(a) Accuracy **(b) F1-Score**

Fig. 4. Comparison of the performance of normalization methods on the Leukaemia 4-class dataset



(a) Accuracy **(b) F1-Score**

Fig. 5. Comparison of the performance of normalization methods on the SRBCT dataset

The concept of L2-norm-based normalization, such as the Vector Normalization used in this study, is widely adopted in modern normalization methods because it preserves the vector's magnitude while only slightly altering its direction (Sun et al., 2020). By maintaining the vector's direction, the relative relationships between features are preserved, making this approach more suitable for high-dimensional microarray data with a limited number of samples. In contrast, methods such as Z-score normalization, Min-Max normalization, Maximum Absolute Scaling, Robust Scaling, Quantile normalization, and Power Transformation perform normalization based on a specific distribution or range of values, thereby altering the original structure of microarray data, which is generally heterogeneous and non-normally distributed. Furthermore, the method's lower performance without normalization indicates that scale differences among features still affect the model's learning. Therefore, the results of this study indicate that in high-dimensional data with a limited number of samples, methods that preserve data structure are more effective than more complex distribution-transformation methods.

4. CONCLUSIONS

Vector Normalization is confirmed as the most effective normalization method for microarray data, consistently demonstrating superior performance. This is evidenced by its highest accuracy and F1-score across all three tested datasets. The use of this method significantly improves model performance, even on more complex datasets with a higher number of classes. Consequently, Vector Normalization is recommended as the optimal data pre-processing method for microarray data analysis.

Despite Vector Normalization yielding the best results, the consistently high standard deviation (above 10 for all experiments) indicates that the outcomes are not yet fully stable and show significant fluctuations. This suggests there is still substantial room to optimize the classification model. Further optimization can be achieved by fine-tuning model parameters, exploring different algorithms, or integrating additional preprocessing methods to achieve greater stability and accuracy, thereby reducing variability in results.

Funding

This research was funded by the Ministry of Higher Education of Malaysia, grant number FRGS/1/2022/ICT02/UPSI/02/1.

Acknowledgments

This project is supported by the Fundamental Research Grant Scheme (FRGS/1/2022/ICT02/UPSI/02/1) under the Ministry of Higher Education of Malaysia. The authors would like to thank Universiti Pendidikan Sultan Idris for its invaluable support in managing the grant.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

Ahmed, H. A., Ali, P. J. M., Faeq, A. K., & Abdullah, S. M. (2022). An investigation on disparity responds of machine learning algorithms to data normalization method. *ARO-The Scientific Journal of Koya University, 10*(2), 29–37. https://doi.org/10.14500/aro.10970

Ahsan, M. M., Mahmud, M. P., Saha, P. K., Gupta, K. D., & Siddique, Z. (2021). Effect of data scaling methods on machine learning algorithms and model performance. *Technologies, 9*(3), Article 52. https://doi.org/10.3390/technologies9030052

Baliarsingh, S. K., Dora, C., & Vipsita, S. (2021). Jaya optimized extreme learning machine for breast cancer data classification. In D. Mishra, R. Buyya, P. Mohapatra, & S. Patnaik (Eds.), *Intelligent and Cloud Computing. Smart Innovation, Systems and Technologies* (Vol. 153, pp. 459–467). Springer. https://doi.org/10.1007/978-981-15-6202-0_47

Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A., Benítez, J. M., & Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences, 282*, 111–135. https://doi.org/10.1016/j.ins.2014.05.042

Borkin, D., Némethová, A., Michaľčonok, G., & Maiorov, K. (2019). Impact of data normalization on classification model accuracy. *Vedecké Práce Materiálovotechnologickej Fakulty Slovenskej Technickej Univerzity v Bratislave so Sídlom v Trnave, 27*(45), 79–84. https://doi.org/10.2478/rput-2019-0029

Buša, J., & Poļaka, I. (2021). Variability of classification results in data with high dimensionality and small sample size. *Information Technology & Management Science, 24*, 28–34. https://doi.org/10.7250/itms-2021-0007

Çelik, A. (2024). Evaluating the impact of data normalization on rice classification using machine learning algorithms. *Acadlore Transactions on AI and Machine Learning, 3*(3), 162–171. https://doi.org/10.56578/ataiml030303

Daoud, M., & Mayo, M. (2019). A survey of neural network-based cancer prediction models from microarray data. *Artificial Intelligence in Medicine, 97*, 204–214. https://doi.org/10.1016/j.artmed.2019.01.006

De Amorim, L. B., Cavalcanti, G. D., & Cruz, R. M. (2023). The choice of scaling technique matters for classification performance. *Applied Soft Computing, 133*, Article 109924. https://doi.org/10.1016/j.asoc.2022.109924

Demircioğlu, A. (2024). The effect of feature normalization methods in radiomics. *Insights Into Imaging, 15*(1), Article 2. https://doi.org/10.1186/s13244-023-01575-7

Deng, S., Li, Y., Wang, J., Cao, R., & Li, M. (2023). A feature-thresholds guided genetic algorithm based on a multi-objective feature scoring method for high-dimensional feature selection. *Applied Soft Computing, 148*, Article 110765. https://doi.org/10.1016/j.asoc.2023.110765

Henderi, H., Wahyuningsih, T., & Rahwanto, E. (2021). Comparison of Min-Max normalization and Z-Score normalization in the K-nearest neighbor (kNN) algorithm to test the accuracy of types of breast cancer. *International Journal of Informatics and Information Systems, 4*(1), 13–20. https://doi.org/10.47738/ijiis.v4i1.73

Hira, S., & Bai, A. (2022). A novel map reduced based parallel feature selection and extreme learning for micro array cancer data classification. *Wireless Personal Communications, 123*(2), 1483–1505. https://doi.org/10.1007/s11277-021-09196-3

Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing, 70*(1-3), 489–501. https://doi.org/10.1016/j.neucom.2005.12.126

Huang, G., Huang, G.-B., Song, S., & You, K. (2015). Trends in extreme learning machines: A review. *Neural Networks, 61*, 32–48. https://doi.org/10.1016/j.neunet.2014.10.001

Lima, F. T., & Souza, V. M. (2023). A large comparison of normalization methods on time series. *Big Data Research, 34*, Article 100407. https://doi.org/10.1016/j.bdr.2023.100407

Nagra, A. A., Khan, A. H., Abubakar, M., Faheem, M., Rasool, A., Masood, K., & Hussain, M. (2024). A gene selection algorithm for microarray cancer classification using an improved particle swarm optimization. *Scientific Reports, 14*(1), Article 19613. https://doi.org/10.1038/s41598-024-68744-6

Panda, P., Bisoy, S. K., Panigrahi, A., Pati, A., Sahu, B., Guo, Z., & Jain, P. (2025). BIMSSA: Enhancing cancer prediction with salp swarm optimization and ensemble machine learning approaches. *Frontiers in Genetics, 15*, Article 1491602. https://doi.org/10.3389/fgene.2024.1491602

Permana, I., Wang, S. L., Salisah, F. N., Sanusi, & Yanto, F. (2025). A benchmark of activation functions in extreme learning machine for high-dimensional low-sample-size microarray classification. *Jurnal Inotera, 10*(2), 502–515. https://doi.org/10.31572/inotera.Vol10.Iss2.2025.ID593

Salisah, F. N., Permana, I., & Wang, S. L. (2025). Evaluating the impact of data balancing techniques on the k-nearest neighbors algorithm for microarray data classification. *Jurnal Inotera, 10*(2), 261–271. https://doi.org/10.31572/inotera.Vol10.Iss2.2025.ID497

Sucharita, S., Sahu, B., Swarnkar, T., & Meher, S. K. (2024). Classification of cancer microarray data using a two-step feature selection framework with moth-flame optimization and extreme learning machine. *Multimedia Tools and Applications, 83*(7), 21319–21346. https://doi.org/10.1007/s11042-023-16353-2

Sun, J., Cao, X., Liang, H., Huang, W., Chen, Z., & Li, Z. (2020). New interpretations of normalization methods in deep learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 4, pp. 5875–5882). https://doi.org/10.1609/aaai.v34i04.6046

Tripathy, J., Dash, R., & Pattanayak, B. K. (2024). Unleashing the power of machine learning in cancer analysis: A novel gene selection and classifier ensemble strategy. *Research on Biomedical Engineering, 40*(1), 125–137. https://doi.org/10.1007/s42600-023-00335-2

Tripathy, J., Dash, R., Pattanayak, B. K., & Mishra, S. K. (2025). An integrated ELM based feature reduction combination detection for gene expression data analysis. *SN Computer Science, 6*(1), Article 65. https://doi.org/10.1007/s42979-024-03612-8

Vargas, V. M., Guijo-Rubio, D., Gutiérrez, P. A., & Hervás-Martínez, C. (2021). Relu-based activations: Analysis and experimental study for deep learning. In *Advances in Artificial Intelligence: 19th Conference of the Spanish Association for Artificial Intelligence (CAEPIA)* (pp. 33–43). Springer. https://doi.org/10.1007/978-3-030-85713-4_4

Wang, J., Lu, S., Wang, S.-H., & Zhang, Y.-D. (2022). A review on extreme learning machine. *Multimedia Tools and Applications, 81*(29), 41611–41660. https://doi.org/10.1007/s11042-021-11007-7

Zhu, X. (2018). *Microarray datasets for feature selection* [Data set]. Shenzhen University, College of Computer Science and Software Engineering. http://csse.szu.edu.cn/staff/zhuzx/Datasets.html

Zhu, Z., Ong, Y.-S., & Dash, M. (2007). Markov blanket-embedded genetic algorithm for gene selection. *Pattern Recognition, 40*(11), 3236–3248. https://doi.org/10.1016/j.patcog.2007.02.007

APPENDIX A

Tab. A.1. Wilcoxon test based on accuracy between vector normalization and other normalization methods on the Leukaemia 2-class dataset

Method 1	Method 2	p-value
Without Normalisation	Vector Normalisation	2.06254597477504E-03
Z-score Normalisation		4.59894069935937E-11
Min-Max Normalisation		2.21504937481307E-05
Maximum Absolute Scaling		5.09435312356487E-10
Robust Scaling		1.66117908654419E-08
Quantile Normalization (Uniform Distribution)		1.26011316386561E-05
Quantile Normalization (Gaussian Distribution)		3.67067327174253E-07
Power Transformation		9.05056465591692E-09

Tab. A.2. Wilcoxon test based on accuracy between vector normalization and other normalization methods on the Leukaemia 3-class dataset

Method 1	Method 2	p-value
Without Normalisation	Vector Normalisation	3.49397571607174E-09
Z-score Normalisation		5.20984805089850E-27
Min-Max Normalisation		1.42604649367244E-17
Maximum Absolute Scaling		7.65591510399544E-18
Robust Scaling		3.60642808262647E-25
Quantile Normalization (Uniform Distribution)		1.61635486283375E-08
Quantile Normalization (Gaussian Distribution)		7.18557313350906E-19
Power Transformation		5.53457756095756E-28

Tab. A.3. Wilcoxon test based on accuracy between vector normalization and other normalization methods on the Leukaemia 4-class dataset

Method 1	Method 2	p-value
Without Normalisation	Vector Normalisation	4.93748337310352E-03
Z-score Normalisation		4.77570012995736E-16
Min-Max Normalisation		1.16826004263723E-07
Maximum Absolute Scaling		1.17419225464744E-15
Robust Scaling		1.33646096659908E-19
Quantile Normalization (Uniform Distribution)		2.71802012791589E-05
Quantile Normalization (Gaussian Distribution)		2.75882167388431E-10
Power Transformation		7.10958065555358E-17

Tab. A.4. Wilcoxon test based on accuracy between vector normalization and other normalization methods on the SRBCT dataset

Method 1	Method 2	p-value
Without Normalisation	Vector Normalisation	6.46634854460976E-07
Z-score Normalisation		2.83797507270013E-18
Min-Max Normalisation		7.35333280063716E-06
Maximum Absolute Scaling		1.53189094402830E-03
Robust Scaling		8.55458880278199E-16
Quantile Normalization (Uniform Distribution)		2.42936679704951E-02
Quantile Normalization (Gaussian Distribution)		1.15275796239145E-12
Power Transformation		3.26981972968950E-06