CAUCHY: Jurnal Matematika Murni dan Aplikasi
Volume 9(2) (2024), Pages 247-259
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Enhancing Binary Classification Performance in Biomedical Datasets: Regularized ELM with SMOTE and Quantile Transforms Focused on Breast Cancer Analysis

Brilliant Friezka Aina[1], Meta Kallista[2,*], Ig. Prasetya Dwi Wibawa[3], Ginaldi Ari Nugroho[4], Ivana Meiska[5], Syifa Melinda Naf'an[5]

[1]Electrical Engineering, Faculty of Electrical Engineering, Telkom University
[2]CoE HUMIC, Computer Engineering, Faculty of Electrical Engineering, Telkom University
[3]CoE STAR, Electrical Engineering, Faculty of Electrical Engineering, Telkom University
[4]National Research and Innovation Agency
[5]Computer Engineering, Faculty of Electrical Engineering, Telkom University

Email: metakallista@telkomuniversity.ac.id

ABSTRACT

Using microarray datasets, this research investigation addresses the problem of unbalanced data in binary classification tasks. The objective is to increase classification performance by adding Extreme Learning Machine (ELM) regularization, as well as Synthetic Minority Over-Sampling Technique (SMOTE) for data over-sampling and Quantile Transformer for data scaling. The study began with gathering important biological datasets from reputable sources such as UCI and Kaggle, including Pima Indian Diabetes, Heart Disease, and Wisconsin Breast Cancer. SMOTE was employed to solve the difficulty of data imbalance in the preparation of the dataset. The data was then separated into training (80%) and testing (20%) sets before being scaled using Quantile Transformation. To boost accuracy, ELMs were employed with an emphasis on introducing regularization techniques. Regularized ELM (R-ELM) surpasses ELM in terms of AUC, despite ELM's faster calculation time. The final selection of the regularization parameter (C) in R-ELM influences the model's performance and calculation time. Overall, R-ELM with SMOTE achieved an average accuracy of 0.964 across all datasets. This method shows promise for improving the classification of imbalanced biomedical datasets, especially in cancer detection. Further validation on additional datasets is recommended to confirm its robustness and broader applicability.

Keywords: Regularized Extreme Learning Machine (R-ELM) Algorithm, Biomedical Research, Synthetic Minority Over-Sampling Technique (SMOTE), Quantile Transformer

Copyright © 2024 by Authors, Published by CAUCHY Group. This is an open access article under the CC BY-SA License (https://creativecommons.org/licenses/by-sa/4.0/)

Submitted: August 05, 2024 Reviewed: September 23, 2024 Accepted: October 18, 2024
DOI: http://dx.doi.org/10.18860/ca.v9i2.28785

INTRODUCTION

Technology plays a vital role in modern healthcare, especially for managing critical conditions like diabetes, breast cancer, and heart disease. These technological advancements not only accelerate diagnosis but also enhance real-time monitoring and classification of patients, leading to more accurate disease prognosis. Additionally, technology facilitates more efficient storage and analysis of extensive medical data, which is crucial given the increasing number of patients with these conditions. For instance, in Indonesia around 2,789,064 individuals are affected by heart disease, underscoring the need for effective data management solutions to improve diagnostic speed and treatment accuracy [1]. Machine learning has emerged as a powerful tool to address the challenges of medical data management. By leveraging algorithms and mathematical models, machine learning can classify data, predict diseases, and support decision-making with greater speed and accuracy. However, a significant challenge is data imbalance, where some classes of data, such as breast cancer cases, are underrepresented compared to others, leading to skewed predictions and reduced model performance [2].

The aim of this study is to improve classification performance on imbalanced medical datasets by combining ELM with SMOTE and introducing Regularized ELM (R-ELM). The regularization component in R-ELM addresses overfitting by controlling model complexity and improving generalization. By using SMOTE to generate synthetic samples for minority classes, we observed a significant improvement in classification accuracy across all datasets, achieving an average accuracy of 0.964. Evaluation metrics, including Area Under the Curve (AUC) and confusion matrix, were used to assess model performance. The results show that the R-ELM with SMOTE outperforms the standard ELM, especially in identifying minority classes and reducing false positives and negatives. These adjustments are crucial for managing noisy and imbalanced data often found in medical datasets. Overall, this research aims to provide more reliable disease predictions and support healthcare professionals in making informed decisions based on accurate data analysis.

METHODS

The topic of this study focuses on the problem of unbalanced data in multiclass classification problems using microarray datasets. The target is to enhance classification performance by using Regularization Extreme Learning Machine (R-ELM) with changed regularization parameter (C) and combining SMOTE for data oversampling and Quantile Transformer for data scaling.

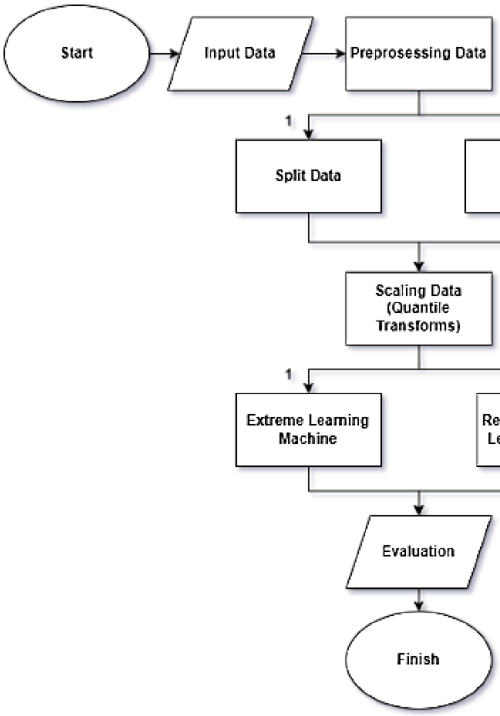


Figure 1. System Overview

The system involves several key steps: acquiring significant biological datasets (Pima Indian Diabetes, Heart Disease, and Wisconsin Breast Cancer) from reliable sources such as UCI and Kaggle. These datasets contain important information regarding diabetes diagnosis, heart conditions, and tumor characteristics, respectively. Data preprocessing using SMOTE addresses the issue of imbalance, ensuring a more representative dataset. The dataset was then divided into 80% training data and 20% testing data, followed by scaling with Quantile Transformation to improve data normalization. The Extreme Learning Machine method, with additional regularization, was used to improve accuracy. We anticipate that the application of R-ELM combined with SMOTE will result in a significant improvement in classification accuracy, ultimately helping healthcare professionals make better decisions.

SMOTE (Synthetic Minority Oversampling Technique)

The SMOTE (Synthetic Minority Over-Sampling Technique) algorithm is a technique often used to solve the issue of unbalanced data in behavioral datasets. Its primary goal is to produce supplemental data points for minority groups within the dataset. The notion of SMOTE entails producing synthetic data by duplicating or replicating samples from the minority class in a way that makes them comparable to the majority class [3-4]. To balance the class distribution, these newly produced instances are added to the original dataset. Figure 2 illustrates the process of converting in the SMOTE algorithm.

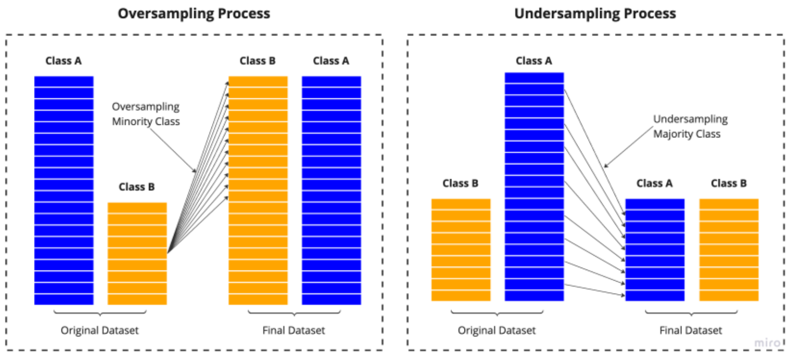


Figure 2. Systematic Conversion SMOTE

The SMOTE method operates by picking a sample from the minority class and locating its K closest neighbors inside the feature space. Interpolating between the specified sample and its neighbors generates synthetic samples. The SMOTE ratio, a user-defined parameter, controls the amount of synthetic samples generated for the minority class, ensuring balanced data [5].

Quantile Transforms

Quantile scaling is a mechanism for converting numerical input or output variables to a Gaussian or uniform probability distribution [6-7]. They give an automated method for transforming a numeric input variable into a new data distribution, which may then be utilized as input to a predictive model [6]. Quantile normalization transforms numerical input variables in the training dataset to ensure statistical equality between distributions by sorting and averaging values. It is commonly used in microarray data analysis before feeding data into machine learning models [7].

Distribution of the future is required in different real-world scenarios instead of a single point prediction. As a consequence, we consider modeling the random variable's cumulative distribution (CDF) if its value is random $Y^n_{\tau+1} \in \mathbb{R}$ complementing previous experiences. Let us designate the CDF as $F_Y(y)$, following that at the quantile level the quantile estimates $\alpha \in (0,1)$ is [8-10]:

$$Q_Y(\alpha) \coloneqq \inf\{y \in \mathbb{R} : \alpha \leq F_Y(y)\} \tag{1}$$

Which the function, Q_Y is the quantile function or the inverse CDF function. Intuitively, $\alpha \in (0,1)$ is that the possibility that Y will be smaller than $Q_Y(\alpha)$. The quantile estimate can be written as follows:

$$q^n_{\alpha,\tau+1} = Q_Y(\alpha\,|Y_{1:\tau}, X_{1:\tau}, \Theta) \tag{2}$$

The lost functionality, can be defined in the following way:

$$\begin{aligned} \rho_\alpha(Y, q_\alpha) &= (Y - q_\alpha)(\alpha - \mathbb{I}_{(Y \leq q_\alpha)}) \\ &= \begin{cases} \alpha(Y - q_\alpha), & \text{if } Y \geq q_\alpha \\ (\alpha - 1)(Y - q_\alpha), & \text{if } Y < q_\alpha \end{cases} \end{aligned} \tag{3}$$

Quantile normalization sorts distributions and sets each value to the average of corresponding values. It is commonly used in microarray data analysis [10]. It is provided in the scikit-learn toolkit and may be applied to any numerical input variable in a dataset. Quantile normalization, on the other hand is a technique used in statistics to make two distributions identical in statistical properties.

Extreme Learning Machine

The Extreme Learning Machine is a feedforward neural network with a single hidden layer or commonly known as SLFNs. This approach has applications in classification, regression, clustering, pattern recognition, forecasting, diagnostics, and image processing. ELM is thought to solve the shortcomings of previous artificial neural networks that excel in terms of learning speed and generalization [11]. The ELM method does not train the input weights, but it was random or biased in order to maximize the value achieved in the ELM process. Figure 3 depicts the systematics in the ELM process.

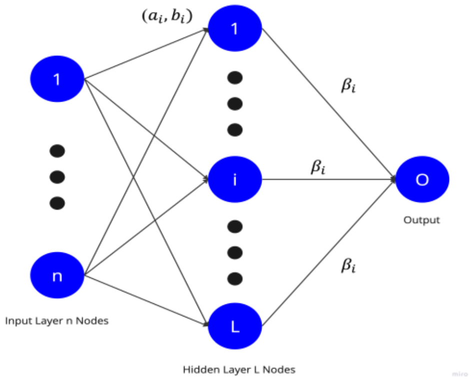


Figure 1. Extreme Learning Machine Algorithm Architecture

In Figure 3, ELM has three primary stages: data enters the input layer, then the concealed layer, and finally the output layer. ELM, unlike other artificial neural networks, does not employ the backpropagation technique in its implementation to train weights repeatedly [12, 13]. The ELM algorithm employs a mechanism for determining weights between the input layer and randomly initialized hidden units, and the weights of the hidden units and the output layer (beta matrix) are determined directly. The Hidden Layer L Node output may be estimated using the following formula:

$$T = \sum_{i=1}^{L} \beta_i . g_i(a_i x_k + b_i) = h(x).\beta \tag{4}$$

The main objective of training the ELM algorithm is to find the beta weights that may provide output that is near to or reduces the difference between the forecast output and the intended target [14]. If there is no regularization (C = 0) in the procedure, the answer for the beta weight matrix is as follows:

$$\beta = H^T T \tag{5}$$

In such a situation, R-ELM has two solutions for the beta weight matrix corresponding to the number of training samples is also an option. In this research, we will concentrate on regularization in ELM, this sample point determination is separated into two parts:
Number of training samples that are not too large (solution 1):

$$\beta = H^T \left(\frac{I}{C} + HH^T\right)^{-1} T \tag{6}$$

For the second, the number of training sample points is very large (solution 2):

$$\beta = \left(\frac{I}{C} + H^T H\right)^{-1} H^T T \tag{7}$$

ELM can categorize data into many groups. It's been utilized in a variety of applications, including image identification, text categorization, and medical diagnostics [15].

RESULTS AND DISCUSSION

Dataset

The datasets focused on in this research are biomedical datasets which are used from UCI and Kaggle as benchmark datasets. The datasets are:

1) Pima Indian Diabetes

The National Institute of Diabetes and Digestive and Kidney Diseases published this data [16]. The dataset's goal is to diagnostically predict whether or not a patient has diabetes based on specific diagnostic parameters. All patients here, in particular, are Pima Indian women over the age of 21, and information sets include a number of medical predictor factors as well as one target variable [17]. Variables include the patient's previous pregnancies, glucose, blood pressure, skin thickness, BMI, insulin level, age, and the classification result, which becomes the desired output in the form of binary classification. The dataset refers to "Outcome" attributes, which are utilized as study goals.

2) Heart Disease

Heart Disease Dataset is a multivariate dataset, which indicates that it contains or includes a number of distinct mathematical or statistical variables, as well as multivariate numerical data analysis covers 76 properties, although all published research focuses on a subset of 14 of these [18]. There are 14 attributes: age, gender, type of chest pain, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiography results, maximum achieved heart rate, exercise-induced angina, peak duration, degree of exercise-induced depression relative to rest, slope of peak exercise segment, number of major blood vessels, thalassemia, and diagnosis with a binary classification form of 0 and 1.

3) Breast Cancer Wisconsin (Diagnostic)

This database is also accessible through the UW CS FTP server: For each illustration, the mean, standard error, and "worst" or worst (average of the three largest values) of these characteristics were determined, yielding 30 features [19]. This study's diagnostic characteristics are included in the dataset. Multiple evaluations were taken on the cell nuclei from the breast mass photos in the collection. The dataset's class distribution indicates 357 cases of benign cells and 212 instances of malignant cells. This distribution offers a balanced representation of both classes, allowing for accurate categorization and analysis of breast mass pictures.

Evaluation Criteria

In optimizing the Extreme Learning Machine (ELM) algorithm with SMOTE for data imbalance and Quantile Transformer for scaling, metrics such as confusion matrix provides a detailed view, including true positives, true negatives, false positives, and false negatives, from which additional metrics like G-Means, specificity, and sensitivity can be derived. The model's performance was also assessed using the ROC curve and AUC Score. The ROC curve plots the true positive rate against the false positive rate across different thresholds, while the AUC measures the area under this curve, indicating the model's ability to distinguish between positive and negative classifications [20]. The bigger the AUC value, the better the model's performance, which shows the positive rate versus the false positive rate at various levels of different categorization thresholds. The area under the ROC curve, often known as the AUC value, is a measure of the model's ability to differentiate between positive and negative classifications [21].

Result Analysis

After data preprocessing confirmed that there are no NaN or missing values, the data was deemed clean. The three datasets had imbalanced output classes, so the SMOTE approach was used to address this issue. To address this, the SMOTE approach was applied to each dataset. For instance, in the Pima Indian Diabetes Dataset, there were 500 samples for class 0 and 268 samples for class 1, as detailed in Figure 4.

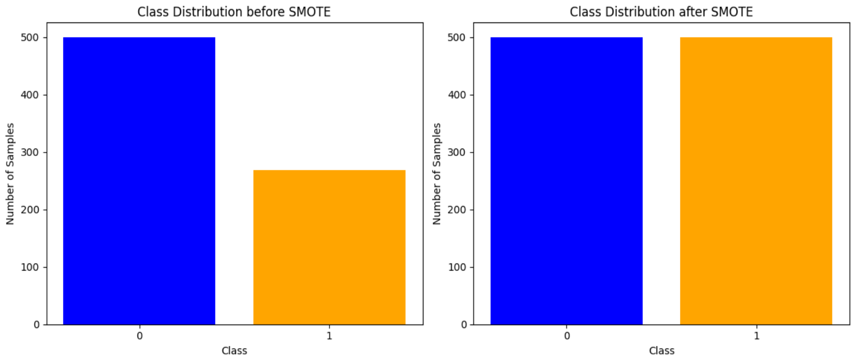


Figure 4. Comparison of The Number of Pima Indian Diabetes

Similarly, in the Heart Disease Dataset, the class distribution before SMOTE consisted of 276 samples for class 0 and 330 samples for class 1, as displayed in Figure 5, indicating a slight imbalance, which was addressed using SMOTE.

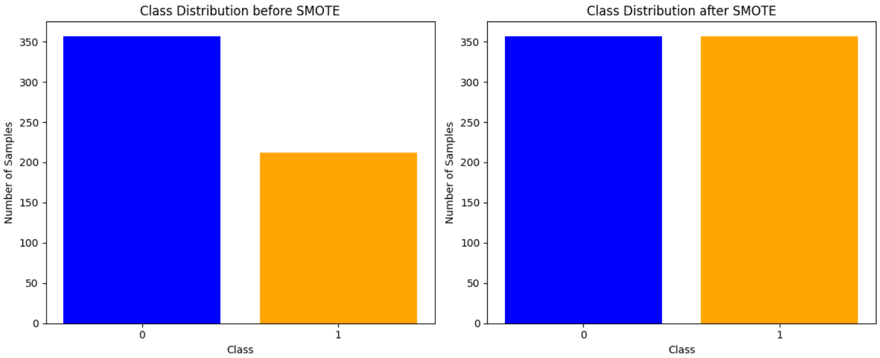


Figure 5. Comparison of The Number of Breast Cancer

In the Wisconsin Breast Cancer Dataset, the class distribution was more imbalanced, with 357 samples for class 0 and 212 samples for class 1, as presented in Figure 6, and this imbalance was also addressed using SMOTE.

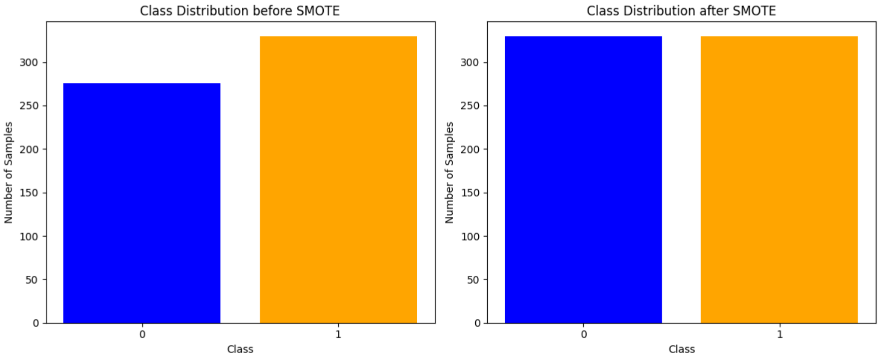


Figure 6. Comparison of The Number of Heart Disease

We compare the performance of the Regularized Extreme Learning Machine (ELM) algorithm with and without the addition of the SMOTE algorithm. To optimize accuracy, we adjust the regularization parameter or C ranging from 1 to 250 so that the results are as in the following Table 1.

Table 1. Result of Comparison Performance Algorithm

Algorithm	Evaluation	Diabetes	Heart Disease	Breast Cancer
ELM	Accuracy	0.675	0.696	0.947
	Running Time (s)	0.017	0.039	0.018
	AUC	0.777	0.847	0.996
ELM + SMOTE	Accuracy	0.630	0.667	0.930
	Running Time (s)	0.014	0.031	0.014
	AUC	0.847	0.935	0.995
R-ELM	Accuracy	0.785	0.868	0.991

		C = 220	C = 5	C = 1
	Running Time (s)	0.005	0.021	0.013
	AUC	0.761	0.914	0.997
R-ELM + SMOTE	Accuracy	**0.795**	**0.871**	**0.991**
		C = 247	C = 1	C = 182
	Running Time (s)	0.007	0.008	0.005
	AUC	0.797	0.935	0.995

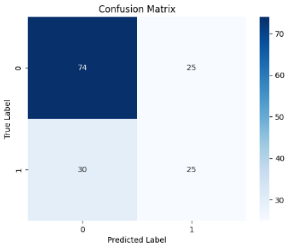
Table 1 shows that the Regularized Extreme Learning Machine (R-ELM), with or without SMOTE, consistently outperforms ELM in terms of accuracy and AUC across all datasets. Despite a slightly longer computation time, R-ELM proves more effective in handling imbalanced data, making it a better option for classification tasks. In our study, we optimized the regularization parameter C by performing a grid search over a range of values from 1 to 250 for each dataset. The selected values C=247 for the Diabetes dataset, C=1 for the Heart Disease dataset, and C=182 for the Breast Cancer dataset represent the best-performing parameters based on the highest accuracy and AUC achieved during cross-validation. The selected values represent the best performance based on accuracy and AUC. These values were chosen to minimize overfitting while maximizing generalization across the training and testing sets.

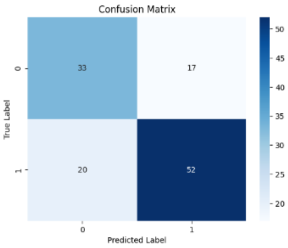
The regularization parameter C plays a crucial role in the Regularized Extreme Learning Machine (R-ELM) algorithm. It controls the trade-off between maximizing the margin of the classifier and minimizing the classification error on the training data. Specifically, a smaller value of C emphasizes maximizing the margin, which can help to prevent overfitting but may result in a higher classification error. Conversely, a larger value of C allows the model to focus more on reducing the training error, potentially leading to overfitting if the model becomes too complex. Thus, selecting an appropriate value for C is essential to achieving a balance that enhances model generalization and performance on unseen data.

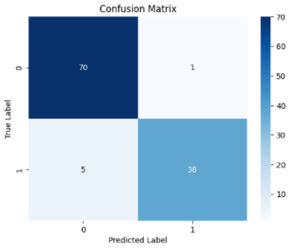
As previously mentioned, this research will be evaluated using confusion matrices derived from the classification results. These matrices provide a comprehensive set of metrics for assessing the performance of the classification algorithms. The classification performance is detailed in Tables 2 to 5, which present the evaluation results for each algorithm. Specifically, Table 2 outlines the performance metrics for the ELM algorithm, Table 3 for the ELM with SMOTE, Table 4 for the R-ELM algorithm, and Table 5 for the R-ELM with SMOTE. By analyzing these confusion matrices, we gain valuable insights into the effectiveness of each algorithm in accurately classifying the datasets.

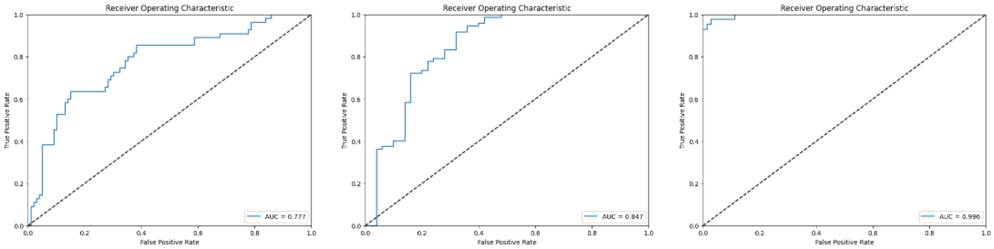
Table 2. Performance Confusion Matrix ELM

Algorithm	Evaluation	Diabetes	Heart Disease	Breast Cancer
ELM	TP	25	52	38
	TN	74	33	70
	FP	25	17	1
	FN	30	20	5









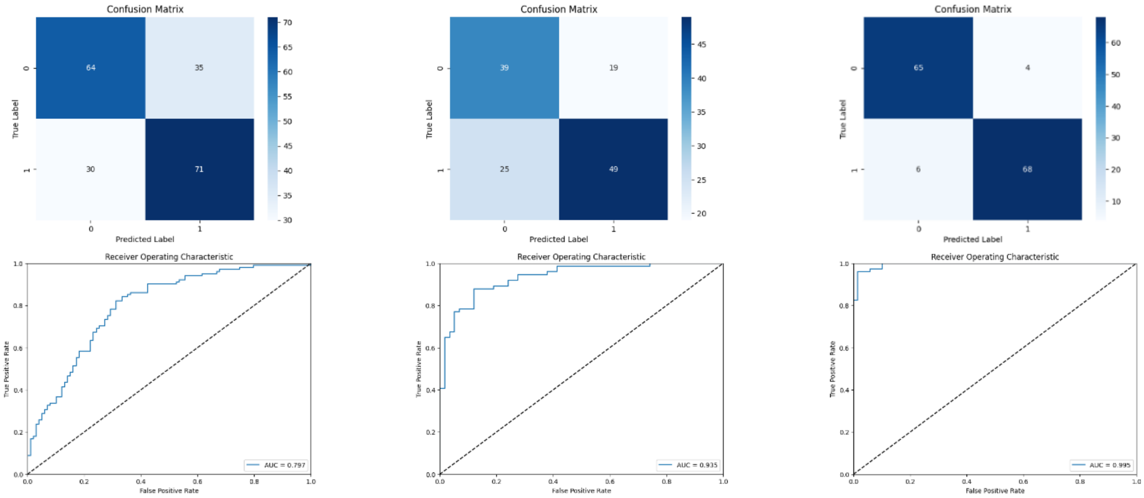
(a) (b) (c)

Figure 7. Confusion Matrix and ROC Curve ELM Algorithm of (a) Pima Indian Diabetes, (b) Heart Disease, and (c) Breast Cancer

On three separate datasets that can be seen in Figure 7, the ELM method performed inconsistently. On the Heart Disease dataset, ELM has a very high TP count but also produces considerable FP and FN. On the Diabetes dataset, ELM produces more FPs than TPs, but has a good amount of TNs as well.

Table 3. Performance Confusion Matrix ELM with SMOTE Algorithm

Algorithm	Evaluation	Diabetes	Heart Disease	Breast Cancer
ELM +SMOTE	TP	71	49	68
	TN	64	39	65
	FP	35	19	4
	FN	30	25	6



(a) (b) (c)

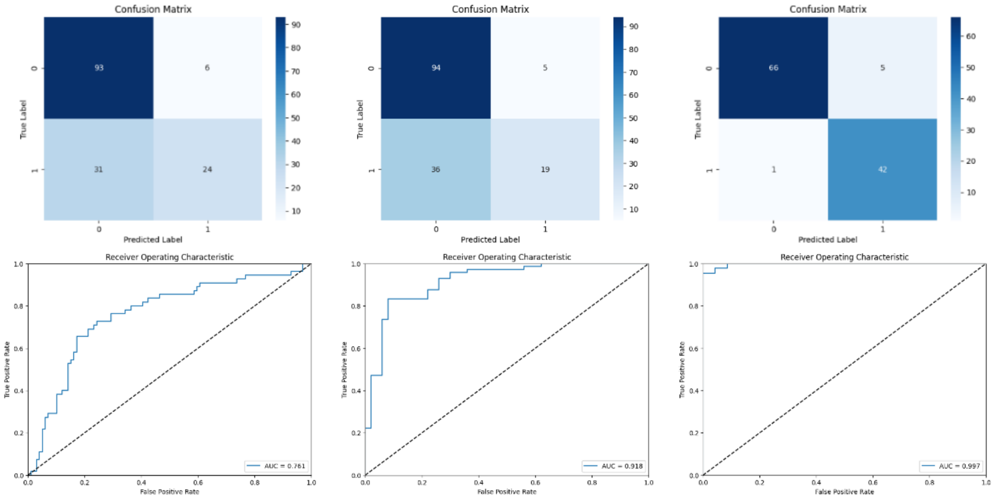
Figure 8. Confusion Matrix and ROC Curve ELM + SMOTE Algorithm of (a) Pima Indian Diabetes, (b) Heart Disease, and (c) Breast Cancer

By applying the ELM with SMOTE method, the number of True Positives (TPs) increased in both the Pima Indian Diabetes and Wisconsin Breast Cancer datasets, as shown in Figure 8. However, for the Heart Disease dataset, a slight decrease in TPs was observed, indicating that while SMOTE improved overall performance, the effect on each dataset varied. In the Heart Disease and Diabetes datasets, meanwhile, a rise in TP was accompanied by an increase in FP and FN. On the Breast Cancer dataset, ELM with SMOTE performed well, with a high TP count and a low FP and FN count.

Table 4. Performance Confusion Matrix R-ELM Algorithm

Algorithm	Evaluation	Diabetes	Heart Disease	Breast Cancer
Regularized ELM	TP	24	19	42
	TN	93	96	66
	FP	6	3	5

	FN	31	36	1



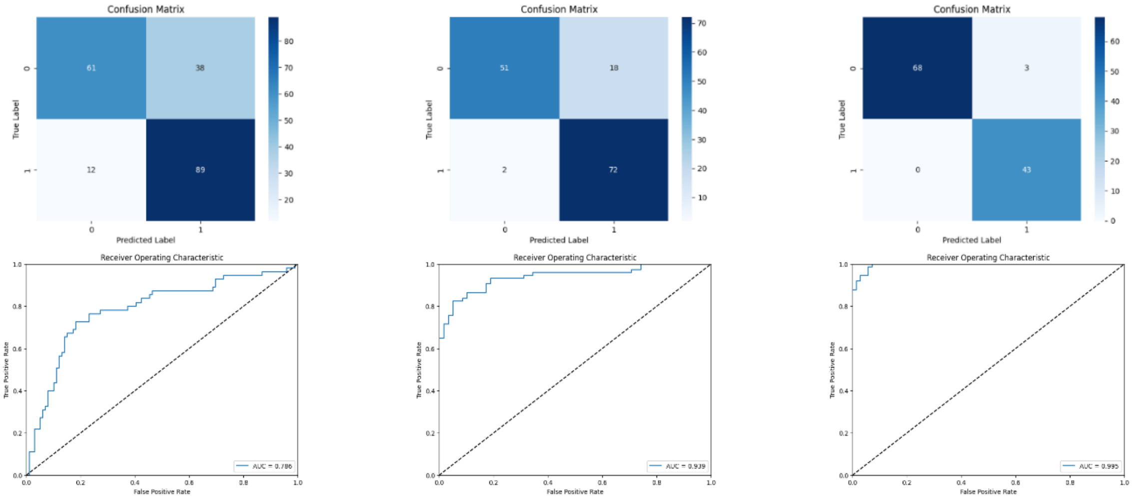
(a) (b) (c)

Figure 9. Confusion Matrix and ROC Curve R-ELM Algorithm of (a) Pima Indian Diabetes, (b) Heart Disease, and (c) Breast Cancer

On two out of three datasets that can be seen in Figure 7 and Figure 9, the R-ELM method outperformed the ELM algorithm. The R-ELM produced a higher TP count but a lower FP and FN count on the Heart Disease dataset. On the Diabetes dataset, Regularized ELM significantly reduced the number of FPs and FNs, resulting in superior performance to ELM.

Table 5. Performance Confusion Matrix R-ELM with SMOTE Algorithm

Algorithm	Evaluation	Diabetes	Heart Disease	Breast Cancer
Regularized ELM + SMOTE	TP	89	72	43
	TN	61	51	68
	FP	38	18	3
	FN	12	2	0



(a) (b) (c)

Figure 10. Confusion Matrix and ROC Curve R-ELM + SMOTE Algorithm of (a) Pima Indian Diabetes, (b) Heart Disease, and (c) Breast Cancer

On all three datasets, the R-ELM with SMOTE method performed exceptionally well. Using SMOTE, the model was able to greatly increase the number of TPs while decreasing the number of FPs and FNs. R-ELM with SMOTE performed best on the Heart Disease and Breast Cancer datasets, with the greatest TP count and the lowest FP and FN counts. In this research, the use of SMOTE to ELM or R-ELM can be an effective approach in enhancing classification accuracy and lowering mistakes in data prediction for application instances requiring imbalanced data management. Even so, the proper method must be chosen in accordance with the features of the dataset and the requirements of certain applications. Another promising area for future exploration is the application of the R-ELM with SMOTE framework to real-world problems, such as medical imaging or fraud detection, where class imbalance is a common issue. Finally, the model's robustness could be tested using real-time data, and its scalability could be assessed for larger datasets in different industries. Developing methods to automate parameter selection, such as using hyperparameter optimization techniques (e.g., Bayesian optimization), could also significantly improve the model's adaptability and performance across various domains.

CONCLUSIONS

In all datasets, R-ELM with SMOTE, outperformed ELM and ELM with SMOTE with an impressive average accuracy gain of around 0.964. This underscores its consistent ability to significantly improve accuracy in overcoming the challenges of imbalanced data across three different biological datasets. Utilizing Quantile Transforms improves the symmetry and normality of the data, thus enhancing the model's ability to detect complex patterns and correlations among features and target labels. In particular, on the breast cancer dataset R-ELM with SMOTE proved to be a superior method and in a relatively fast time produced the highest number of True Positives while reducing False Positives and False Negatives. Its superiority in identifying minority classes places it as the preferred choice for applications that require reliable classification in skewed data scenarios. Expanding on the conclusion, our model emerges as a robust solution, particularly suitable for cases such as breast cancer analysis. Its consistently superior performance and ability to handle imbalanced biomedical datasets make it valuable for reliable classification, especially in critical scenarios like cancer detection. The incorporation of Quantile Transforms further refines its ability to discern patterns, contributing not only to the broader field of imbalanced multiclass classification but also highlighting its specific suitability for optimizing breast cancer diagnostic processes.

REFERENCES

[1] K. R. I. P2PT, “Hari Jantung Sedunia (World Heart Day): Your Heart is Our Heart Too,” https://p2ptm.kemkes.go.id/, Sep. 26, 2019.

[2] R. Rustogi and A. Prasad, "Swift Imbalance Data Classification using SMOTE and Extreme Learning Machine," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862112.

[3] Hairani, N. A. Setiawan, & T. B. Adji. 2013. Metode Klasifikasi Data Mining dan Teknik Sampling Smote.Seminar Nasional Sains dan Teknologi, 168–172.

[4] A. A. Dharmasaputro, N. M. Fauzan, M. Kallista, I. P. D. Wibawa and P. D. Kusuma, "Handling Missing and Imbalanced Data to Improve Generalization Performance of Machine Learning Classifier," 2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE), Jakarta, Indonesia, 2022, pp. 140-145, doi: 10.1109/ISMODE53584.2022.9743022.

[5] Wang Y, Wu X, Chen Z, Ren F, Feng L, Du Q. Optimizing the Predictive Ability of Machine Learning Methods for Landslide Susceptibility Mapping Using SMOTE for

Lishui City in Zhejiang Province, China. International Journal of Environmental Research and Public Health. 2019; 16(3):368. https://doi.org/10.3390/ijerph16030368.

[6] J. Peng, R. Gao, L. Nguyen, Y. Liang, S. Thng and Z. Lin, "Classification of Non-Tumorous Facial Pigmentation Disorders Using Improved Smote and Transfer Learning," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 220-224, doi: 10.1109/ICIP.2019.8802993.

[7] Amaratunga, D.; Cabrera, J. (2001). "Analysis of Data from Viral DNA Microchips". Journal of the American Statistical Association. 96 (456): 1161. doi:10.1198/016214501753381814.

[8] Bolstad, B. M.; Irizarry, R. A.; Astrand, M.; Speed, T. P. (2003). "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias". Bioinformatics. 19 (2): 185–193. doi:10.1093/bio informatics/19.2.185.

[9] Jan Gasthaus, Konstantinos Benidis, Yuyang Wang, Syama Sundar Rangapuram, David Salinas, Valentin Flunkert, and Tim Januschowski. 2019. Probabilistic forecasting with spline quantile function RNNs. In International conference on Artificial Intelligence and Statistics. PMLR.

[10] Roger Koenker and Kevin F Hallock. 2001. Quantile regression. Journal of economic perspectives 15, 4 (2001), 143–156.

[11] Youngsuk Park, Danielle Maddix, François-Xavier Aubet, Kelvin Kan, Jan Gasthaus, and Yuyang Wang. 2021. Learning Quantile Functions without Quantile Crossing for Distribution-free Time Series Forecasting. arXiv preprint arXiv:2111.06581 (2021).

[12] G. Bin Huang, Q. Y. Zhu, and C. K. Siew, “Extreme learning machine: Theory and applications,” Neurocomputing, vol. 70, no. 1–3, pp. 489–501, Dec. 2006, doi: 10.1016/j.neucom.2005.12.126.

[13] I. P. D. Wibawa, C. Machbub, A. S. Rohman, and E. Hidayat, “Modified online sequential extreme learning machine algorithm using model predictive control approach,” Intelligent Systems with Applications, vol. 18, May 2023, doi: 10.1016/j.iswa.2023.200191.

[14] G. Bin Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 42, no. 2, pp. 513–529, Apr. 2012, doi: 10.1109/TSMCB.2011.2168604.

[15] C. Chen, K. Li, M. Duan, and K. Li, “Extreme Learning Machine and Its Applications in Big Data Processing,” in Big Data Analytics for Sensor-Network Collected Intelligence, Elsevier Inc., 2017, pp. 117–150. doi: 10.1016/B978-0-12-809393-1.00006-4.

[16] UCI Machine Learning, “Pima Indians Diabetes Database,” https://www.kaggle.com, 2016.

[17] Smith, J.W., Everhart, J.E., Dickson, W.C., Knowler, W.C., & Johannes, R.S. (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In Proceedings of the Symposium on Computer Applications and Medical Care (pp. 261-265). IEEE Computer Society Press.

[18] Janosi,Andras, Steinbrunn,William, Pfisterer,Matthias, and Detrano,Robert. (1988). Heart Disease. UCI Machine Learning Repository. https://doi.org/10.24432/C52P4X.

[19] Wolberg,William, Mangasarian,Olvi, Street,Nick, and Street,W.. (1995). Breast Cancer Wisconsin (Diagnostic). UCI Machine Learning Repository. https://doi.org/10.24432/C5DW2B.

[20] I.H. Witten, E. Frank, and M.A. Hall, "Data Mining Practical Machine Learning Tools and Techniques", Third Edition, Elsevier Publisher, USA, 2011.

[21] T. Y. Hadiwandra, "SATIN-Sains dan Teknologi Informasi Perbandingan Kinerja Model Klasifikasi Decission Tree, Bayesian Classifier, Instance Base, Linear Function Base, Rule Base pada 4 Dataset Berbeda," vol. 5, no. 1, 2019, [Online]. Available: http://jurnal.stmik-amik-riau.ac.id.