Empirical Evaluation of Signal Preprocessing in Electrocardiography Signal Classification

Robby Hoover
School of Information Technology
University of Cincinnati
Ohio, United States
hooverrp@mail.uc.edu

Nelly Elsayed
School of Information Technology
University of Cincinnati
Ohio, United States
elsayeny@ucmail.uc.edu

***Abstract*—Data quality plays a crucial role in the performance of machine learning model training. Data preprocessing can contribute to enhancing the data quality, leading to overall improvement of the model performance. Thus, selecting the appropriate preprocessing algorithms is a critical step in designing any machine learning model, which remains challenging. This paper provides an empirical study of scaling algorithms and their reflection on model performance, primarily focusing on the electrocardiography signal classification tasks that are utilized in different disease detection.**

***Index Terms*—ECG, classification, scaling, deep learning, machine learning**

I. Introduction

Data quality plays a significant role in machine learning algorithm performance [1]. With the increase in data size, data noises can cause multiple problems in training machines and deep learning models. Thus, data preprocessing and scaling are crucial to improve the machine learning model.

Different data sources have different complexity and noise assigned for specific data collection methodologies [2]. In addition, the range of the collected data may vary from one source to another even if the data represents the same concept [3], [4].

In this paper, we empirically study the effect of using different scalers as a step of data preprocessing on electrocardiography (EKG or ECG) signals classification tasks as one of the applications that are used to detect different diseases.

II. Electrocardiogram (ECG)

The electrocardiogram (ECG) or (EKG) is the recording of the electrical signal produced by the heart due to its activity [5]. The ECG signal is recorded by placing electrodes such as Ag/AgCl electrodes [6] over the skin to detect the tiny electrical changes on the skin. The ECG signal depolarizes and repolarizes during each heartbeat [5]. The ECG is considered one of the simplest and safest non-invasive methods to detect cardiac problems [7]. Thus, several machine learning models address detecting cardiac problems using machine learning and deep learning such in [8]–[10].

A. Types of Scalers

There are multiple types of data scalers. In this paper, we mainly focuse on eight different scalers as most commonly used scalers in signal processing tasks. These scalers are [**?**]:

- ***Standard Scaler (Z-Score):*** It is the most common normalization method, however, a disadvantage of this method is that all the normalized features have a unity variance [11]. Also, it known as Z-Score Normalization, Mean Normalization or just Standardizing. The Standard scaler is performed as follows: for a given speech feature X, the mean (average) of the data μ and the standard deviation σ are computed [12]. The complete standardized feature is calculated with the following equation:

$$\text{Standard scaler}(x) = \frac{x-\mu}{\sigma} \tag{1}$$

- ***Minimum and Maximum Scaler (MMS):*** This scaler is also know as "Feature Scaling" or just "Normalization" [13]. The MMS scales each feature to a given range and is calculated with the following equation:

$$\text{MMS}(x) = \frac{(x-\min(x))(b-a)}{\max(x)-\min(x)} + a \tag{2}$$

- ***Maximum Absolute Scaler (MAS):*** It scales each feature by its maximum absolute value.

$$MAS(x) = \frac{x}{\max(|x|)} \tag{3}$$

- ***Robust Scaler (RS):*** It removes the median and scales the data according to the quantile range. The quantile range is the range between the 1$^{\text{st}}$ quartile (25$^{\text{st}}$ quantile) and the 3$^{\text{rd}}$ quartile (75$^{\text{th}}$ quantile).

$$RS(x) = \frac{x-Q_1(x)}{Q_3(x)-Q_1(x)} \tag{4}$$

- ***Power Transformer (PT):*** it applies a power transform featurewise to make the data more Gaussian-like. This implementation uses the Yeo-Johnson transform.
- ***Quantile Uniform Transformer (QUT):*** This scaler uses quantile information to scale features. It applies a non-linear transformation such that the probability density function of each feature will be mapped to a uniform distribution. The QUT scaler operates in two steps. The first step is to compute the rank of each data point in the original feature, denoted as $R(X)$. It represents the percentile rank of each data point. The second step involves transforming the rank values $R(X)$ into a uniform

Algorithm 1: Power Transformer (PT) Scaler with Yeo-Johnson Transform

Data: Data $\mathbf{X}$, Transformation exponent λ
Result: Transformed data $\mathbf{X}'$
for *each element* x *in* $\mathbf{X}$ **do**
  if $x \leq 0$ **then**
    offset $= \min(\mathbf{X} + 1)$
    $\mathbf{X}' = \log(\mathbf{X} + \text{offset})$
  else
    $\mathbf{X}' = (\mathbf{X} + 1)^{\lambda} - 1$
return $\mathbf{X}'$

TABLE I
A COMPARISON BETWEEN DIFFERENT SCALERS EFFECT ON THE CLASSIFICATION OF ECG200 DATASET.

scaler	**Accuracy**	**F1-Score**	**Prec.**	**Recall**
No Scaler	0.870	0.861	0.856	0.868
Standard Scaler	0.890	0.878	0.887	0.872
Min-Max Scaler	0.850	0.836	0.838	0.834
Max-Abs Scaler	0.890	0.878	0.887	0.872
Robust Scaler	**0.910**	**0.904**	0.898	**0.911**
Power Transformer Scaler	0.900	0.889	**0.902**	0.879
QUT Scaler	0.870	0.858	0.860	0.856
QGT Scaler	0.880	0.870	0.870	0.870

distribution. This is done using the inverse cumulative distribution function (CDF) of the uniform distribution.

Step One:

$$R(X) = \frac{1}{N} \sum_{i=1}^{N} I(X_i \leq X) \tag{5}$$

where, N is the total number of data points. X_i is the value of the i-th data point. $I(\cdot)$ is the indicator function (1 if true, 0 if false).

Step Two:

$$X_{\text{QUT}} = \text{UniformInverseCDF}(R(X)) \tag{6}$$

where UniformInverseCDF(x) is the inverse CDF of the uniform distribution.

- ***Quantile Guassian Transformer (QGT):*** It is quite similar to the Quantile Uniform Transformer, however, the output distribution follows Gaussian distribution rather than a uniform distribution. The Wiener filter can be defined as:

$$\hat{S}(f) = H(f) \cdot X(f) \tag{7}$$

where: $\hat{S}(f)$ is the estimated clean signal in the frequency domain, H(f) is the frequency response of the Wiener filter, and X(f) is the observed noisy signal in the frequency domain.

III. Experimental Methodology

To investigate the performance of different scalers on the ECG signal, multiple experiments will be performed using various ECG data sources that addresses different tasks. As a pilot investigation, we started with the UCI ECG200 dataset as one of the most popular univariant timeseries datasets. To set qn equaivallent comarison between all the scalers, we used the GRU-FCN model that proposed in [14] as the classifier and in every empirical investigation, we changed the scaler. Table I shows the empirical results of our experiments on the effect of the scalers on model overall performance.

IV. Conculision and Future Work

Data quality plays a crucial role in the model performance and results. Thus, in this paper, we aim to empirically investigate the effect of scalers on the electrocardiogram (ECG) data classification as a pilot empirical evaluation. The future work includes the investigation of other ECG datasets as well as both deep learning and machine learning-based classifiers for ECG classification purposes to identify the best practice for ECG data scaling for classification purposes that can serve in multiple ECG-based disease diagnostic models.

References

[1] A. Amato and V. Di Lecce, "Data preprocessing impact on machine learning algorithm performance," *Open Computer Science*, vol. 13, no. 1, p. 20220278, 2023.

[2] D. F. Nettleton, A. Orriols-Puig, and A. Fornells, "A study of the effect of different types of noise on the precision of supervised learning techniques," *Artificial intelligence review*, vol. 33, pp. 275–306, 2010.

[3] P. Oliveira, F. Rodrigues, and P. R. Henriques, "A formal definition of data quality problems.," in *ICIQ*, 2005.

[4] Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: a big data-ai integration perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 4, pp. 1328–1347, 2019.

[5] R. J. Martis, U. R. Acharya, and H. Adeli, "Current methods in electrocardiogram characterization," *Computers in biology and medicine*, vol. 48, pp. 133–149, 2014.

[6] S. Peng, K. Xu, and W. Chen, "Comparison of active electrode materials for non-contact ecg measurement," *Sensors*, vol. 19, no. 16, p. 3585, 2019.

[7] R. Kahankova, R. Martinek, R. Jaros, K. Behbehani, A. Matonia, M. Jezewski, and J. A. Behar, "A review of signal processing techniques for non-invasive fetal electrocardiography," *IEEE reviews in biomedical engineering*, vol. 13, pp. 51–73, 2019.

[8] S. Sahoo, M. Dash, S. Behera, and S. Sabut, "Machine learning approach to detect cardiac arrhythmias in ecg signals: A survey," *Irbm*, vol. 41, no. 4, pp. 185–194, 2020.

[9] N. Elsayed, A. S. Maida, and M. Bayoumi, "An analysis of univariate and multivariate electrocardiography signal classification," in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pp. 396–399, IEEE, 2019.

[10] A. Mincholé, J. Camps, A. Lyon, and B. Rodríguez, "Machine learning in the electrocardiogram," *Journal of electrocardiology*, vol. 57, pp. S61–S64, 2019.

[11] P. Mitra, C. Murthy, and S. Pal, "Unsupervised feature selection using feature similarity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 301–312, 2002.

[12] S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *arXiv preprint arXiv:1503.06462*, 2015.

[13] R. Yao, C. Guo, W. Deng, and H. Zhao, "A novel mathematical morphology spectrum entropy based on scale-adaptive techniques," *ISA transactions*, vol. 126, pp. 691–702, 2022.

[14] N. Elsayed, A. S. Maida, and M. Bayoumi, "Deep gated recurrent and convolutional network hybrid model for univariate time series classification," *arXiv preprint arXiv:1812.07683*, 2018.