Efficient ECG classification based on the probabilistic Kullback-Leibler divergence

Al-Shammary, Dhiah, Radhi, Mohammad R, Alsaeedi, Ali Hakem, Mahdiyar, Amir, Ibaida, Ayman and Ahmed, Khandakar (2024) Efficient ECG classification based on the probabilistic Kullback-Leibler divergence. Informatics in Medicine Unlocked, 47. ISSN 2352-9148

Informatics in Medicine Unlocked 47 (2024) 101510

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Efficient ECG classification based on the probabilistic Kullback-Leibler divergence

Dhiah Al-Shammary [a,*], Mohammed Radhi [a], Ali Hakem AlSaeedi [a], Ahmed M. Mahdi [a], Ayman Ibaida [b], Khandakar Ahmed [b]

[a] College of Computer Science and Information Technology, University of Al-Qadisiyah, Iraq
[b] Intelligent Technology Innovation Lab, Victoria University, Melbourne, Victoria, Australia

A R T I C L E I N F O

Keywords:
KL divergence
ECG arrhythmia
Data mining
Features optimization
Probabilistic classifier

A B S T R A C T

Diagnostic systems of cardiac arrhythmias face early and accurate detection challenges due to the overlap of electrocardiogram (ECG) patterns. Additionally, these systems must manage a huge number of features. This paper proposes a new classifier Kullback-Leibler classifier (KLC) that combines feature optimization and probabilistic Kullback-Leibler (KL) divergence. Particle swarm optimization (PSO) is used for optimizing the features of ECG data, while KL divergence counts the variance between training and testing probability distributions. The proposed framework led the new classifier to distinguish normal and abnormal rhythms accurately. MIT-BIH Standard Arrhythmia Dataset (MIT-BIH) is used to test the validity of the proposed model. The experimental results show the proposed classifier achieves results in precision (86.67%), recall (86.67%), and F1_Score (86.5%).

1. Introduction

A lot of people die from cardiovascular disease (CVD) because of arrhythmia [1]. The arrhythmia contributes to strokes, heart attacks, and sudden cardiac death. Therefore, Early detection of arrhythmia is necessary to improve treatment outcomes and reduce morbidity and mortality. One of the methods used to detect triglyceride diseases is Electrocardiography (ECG) analysis. The measures of the electrical impulses that regulate the contraction and relaxation of the heart are an essential tool for diagnosing heart problems. Electrocardiography (ECG) has developed much since its introduction in the early 20th century [1–5]. They are crucial for detecting any abnormalities in electrical signals. In addition, the ECGs are essential for screening and managing arrhythmia [2]. Most clinics caring for people with heart disease rely on ECG to identify heart problems due to its availability, low cost, and accuracy in identifying diseases. Waveforms of ECG with multiple peaks and troughs challenge clinicians in the diagnosis of heart disease [2].

Examination of these records by a human may result in many errors due to overlapping disease forms and a large number of records. Also, human decision-making was not quick, so most health institutions turned to help specialists in the field of primary care or those in intensive care using algorithms. Smart systems to deal with these cases provide early warning to specialists to alert them to take the necessary measures to treat and follow up on patients [6].

Artificial intelligence algorithms (machine learning) are based on extracting patterns from the ECG and using these patterns to detect abnormal heart rhythms [7]. In order to accurately recognize patterns using learning algorithms, two basic parameters must be defined, important attributes selected and meaningful patterns formed from these attributes in an adapted form. This leads to an improvement of the definition due to its accuracy and also increases the explanatory possibilities since it is a specific and precise number of impulses for a particular disease. There are several applications that could use the KL metrics in its statistical approaches. For example, biology, economics, natural language processing (NLP), and so on. In computational biology, the KL potential plays a role in the comparison of sequence alignments. In economics, it can be utilized to analyze variations in market distributions over time. Moreover, in natural language processing, KL divergence can assess the similarity between different textual corpora, thereby aiding in tasks such as document classification and sentiment analysis.

* Corresponding author.
E-mail addresses: d.alshammary@qu.edu.iq (D. Al-Shammary), mohammad.aziz@qu.edu.iq (M. Radhi), ali.alsaeedi@qu.edu.iq (A.H. AlSaeedi), ahmed.m.mahdi@qu.edu.iq (A.M. Mahdi), ayman.ibaida@vu.edu.au (A. Ibaida), khandakar.ahmed@vu.edu.au (K. Ahmed).

https://doi.org/10.1016/j.imu.2024.101510
Received 2 December 2023; Received in revised form 23 April 2024; Accepted 26 April 2024
Available online 26 April 2024

1.1. Motivation

The ECG is one of the most important analyses to diagnose heart disease [8]. The traditional methods professionals use to diagnose these cases are not free of errors and there are delays in reaching patients and diagnosing the injury. It is difficult for people to monitor and observe patients 24 h a day, which is the reason for the weakness of this method. It delays timely decision-making in serious medical conditions [9]. Emergency diseases including heart diseases, are among the diseases that require advanced medical personnel to make decisions in the shortest possible time. This prompted advanced medical institutions to rely on electronic technologies to follow up on patients and assist advanced medical personnel in making the right decisions quickly [5, 10].

1.2. Proposed solution

This paper proposes a new classifier based on Kullback-Leibler (KL) divergence to improve early arrhythmia detection by analyzing ECG test data. The Kullback-Leibler (KL) divergence can provide potential detection capability and measure the statistical dissimilarity between the testing ECG and pre-knowledge (data in the training phase) datasets [11,12]. Then, the voting process is applied to the highest five selected ECG records to determine the context of the testing signal. Particle Swarm Optimization (PSO) refines classifier accuracy [13]. The proposed heuristic technique is specifically tailored to reduce the dimensionality of our complex, high-dimensional dataset, thus enhancing the model's performance metrics [14]. Evidently, the main contributions can be summarized by the proposed model's outperformance efficiency compared to other previous approaches. The new KL classification technique can also be used for long-term remote ECG monitoring.

1.3. Evaluation strategy

Arrhythmia paradigms, when analysed in the context of intelligent machine learning, require a rigorous evaluation strategy to assess the validity and performance of algorithmic interventions in decision making. To evaluate the proposed intelligent model for heartbeat diagnosis, an approach based on the confusion matrix and precision, recall, and f1 score were used, in addition to a comparison with state-of-the-art models in the field. This provides a comprehensive overview of the performance of the proposed model by reviewing its ability to correctly classify heartbeats and comparing it with the latest research and available models. Below are some important points to consider during the evaluation. Confusion matrix gives an idea of the errors made by the model during classification. By analysing positive true cases, negative true cases, positive errors and negative errors, the strengths and weaknesses of the model can be identified.

1.4. Paper organization

The rest of the paper is organized as follows: Section 2 reviews some existing research on IDS using machine learning algorithms. Section 3 describes the system model and proposed solution. Section 4 provides detailed results and analysis. Section 5 concludes this paper and presents a future research direction.

2. Related works

This section discusses related studies that utilized the same benchmark dataset we used in this experiment and with the same problem. Madan et al. [15] have proposed a hybrid deep learning approach called 2D–CNN–LSTM, which translates 1D ECG signals into 2D Scalogram images for noise filtering and feature extraction to improve arrhythmias detection and classification using electrocardiogram (ECG) signals. The proposed approach was evaluated using the MIT-BIH arrhythmia database, and it achieved approximately 98.7% accuracy for Cardiac Arrhythmias (ARR), 99% for Congestive Heart Failure (CHF), and 99% for Normal Sinus Rhythm (NSR). The study also reported an average sensitivity of 98.33% and a specificity value of 98.35% for all three arrhythmias.

Subasi et al. [16] have proposed an automated ECG signal classification method using tower graphs, hexadecimal binary patterns, and deep learning. Their approach combined minimum, maximum, and average pooling for feature extraction and selected informative features using Relief and NCA. The method achieved 97.10% classification accuracy on the MIT-BIH arrhythmia dataset.

Mian Qaisar et al. [17], have developed a hybrid technique for identifying arrhythmia in mobile healthcare systems using multi-rate processing, QRS selection, variational mode decomposition, feature mining, metaheuristic optimization-based feature selection, and machine learning algorithms. The authors tested three metaheuristic optimization algorithms with proposed model to select best one. Th metaheuristic optimization that used Manta ray foraging optimization (MRFO), Butterfly optimization algorithm (BOA), and Emperor penguin optimization (EPO). The BOA achieved best accuracy rate 99.14% when applied to the MIT-BIH Benchmark dataset.

Kim et al. [18], have proposed a new framework for automatic arrhythmia classification, particularly atrial fibrillation. They used a residual network combined with a bidirectional long short-term memory (LSTM). The proposed framework was evaluated on different benchmark datasets: MIT-BIH arrhythmia database (MITDB), MIT-BIH atrial fibrillation database (AFDB), and PhysioNet Computing in the cardiology challenge 2017 database (CinC DB).

Mathunjwa et al. [19], have proposed an arrhythmia classification algorithm based on ECG recurrence plot (RP) and two-dimensional deep residual CNN features to improve accuracy. The ResNet-18 architecture utilized for detecting ventricular fibrillation (VF) and noise. While ResNet-50 is employed for detecting normal, atrial fibrillation, premature atrial contraction, and premature ventricular contractions. The authors evaluated the proposed model by using publicly available databases from PhysioNet. It consists f the MIT-BIH Atrial Fibrillation Database, the MIT-BIH Arrhythmia Database, the MIT-BIH Malignant Ventricular Ectopy Database, and the Creighton University Ventricular Tachyarrhythmia Database. However, the limitation of the proposed model is low accurate due to merge four different arrhythmia types into a single label.

Kutluana and Türker [20] proposed new model that investigates the representation capacity of visibility graphs for ECG signals. The proposed model uses either the sequence of node weights or the diagonals of the adjacency matrices as feature sets, input to ResNet and Inception classifier models. The used PTB-XL ECG benchmark in testing the proposed model and achieved 93.44% AUC score.

Qin, Huang et al. [21] proposed SelfONN-based lightweight model for general ECG classification. To improve the classification accuracy, the proposed model employs the multi-labeled dataset PTB-XL as benchmark dataset for evaluation of classification performance. This experiment attains a peak of 93 % AUC score when using the ResNet model with the first three diagonals of the visibility graph as the feature set.

An improved deep residual convolutional neural network was proposed by Li et al. [22], for automatic arrhythmia classification, which overcame the challenge of class imbalance in ECG signal segments using overlapping segmentation to segment ECG signals into 5-s segments to overcome class imbalance. While the discrete wavelet transform (DWT) was used for denoising and focal loss function to address imbalanced classification difficulty. The proposed model showed comparable performance to classical and state-of-the-art methods, but a large database of annotated heartbeats is required, which can be challenging to obtain in the medical field. Further exploration of alternative loss functions could be beneficial in improving classification accuracy.

3. Kullback leibler divergence

The Kullback-Leibler (KL) divergence, often termed as relative entropy, stands as a cornerstone in the fields of information theory and statistical inference. Serving as a measure to quantify the divergence between two distinct probability distributions, the KL divergence holds universal applicability across diverse scientific disciplines. Mathematically, The KL measures the "distance" between two distributions, albeit in a non-symmetric fashion. Therefore, it can used in different statistical models.

The formal mathematical definition of KL divergence, denoted as $D_{KL}(P \parallel Q)$, for two probability distributions P and Q as expressed in Eq. (1) [12]:

$$D_{KL}(P \parallel Q) = \sum_i P(i) log \frac{P(i)}{Q(i)} \tag{1}$$

Here, i signifies each event in the sample space, while $P(i)$ and $Q(i)$ represent the respective probabilities of event i according to the distributions $P(i)$ and $Q(i)$. The logarithmic term amplifies the discrepancies between $P(i)$ and $Q(i)$. That makes KL divergence a sensitive measure for finding the differences between the two distributions [11].

It is important to acknowledge that KL divergence is not a metric in the strict mathematical sense due to its non-symmetric property, according Eq. (2) [11]:

$$D_{KL}(P \parallel Q) \neq D_{KL}(Q \parallel P) \tag{2}$$

This asymmetry calls for accurate interpretation, especially when employed in comparative or multidimensional analyses involving multiple distributions.

4. Proposed model

Kullback-Leibler Divergence is developed as a new classifier for ECG. The divergence among the input patient record (input vector) and other dataset records (dataset vectors) are computed. Resultant divergence can be a significant indication to have the right classification. Fig. 1 representation illustrates the proposed solution that endeavors to amalgamate two prominent techniques: the Kullback-Leibler (KL) divergence for classification and Particle Swarm Optimization (PSO) for feature selection. This proposed methodology seeks to harness the inherent strengths of both techniques, crafting an integrated approach aimed at bolstering classification accuracy and efficiency.

The proposed methodology starts by reading and normalizing the raw dataset. Then, split normalized dataset into two sections; testing and training sets of ECG records. Next, investigate the current testing ECG and find optimal similar class (Arrhythmia or normal) by calculating the lowest KL divergence in matching with the training ECG records. Nominate the best (lowest divergence) five training ECG records. Finally, diagnose the input testing ECG based on the highest occurrence class of similar training ECG records. Moreover, in order to investigate the proposed KL ECG classification with the presence of a potential optimizer like PSO. PSO is applied and only its feature selections are depicted for the KL classification. Empirically, accuracy is calculated in order to evaluate the proposed classification efficiency.

4.1. Preprocessing

At the outset, raw data is the foundational bedrock upon which the algorithmic process is anchored. The raw data encompass many features, some of which may be critical to classification while others might be ancillary. Preprocessing, a cardinal step in any data-driven operation, witnesses the normalization of this data. The normalization process is the first preprocessing that should be applied to the raw data to be in analytic form. The normalization process converts each numerical feature value to be between 0 and 1 by utilizing z-score technique, according to Eq. (3), where x, μ, σ reprsent the single value, mean and standard deviation in specific feature.

$$Z1_{score} = \frac{x - \mu}{\sigma} \tag{3}$$

Data normalization is necessary to ensure that all data values are within a specific range. For example, when normalization is between zero and one, it ensures that all values are between zero and one and that

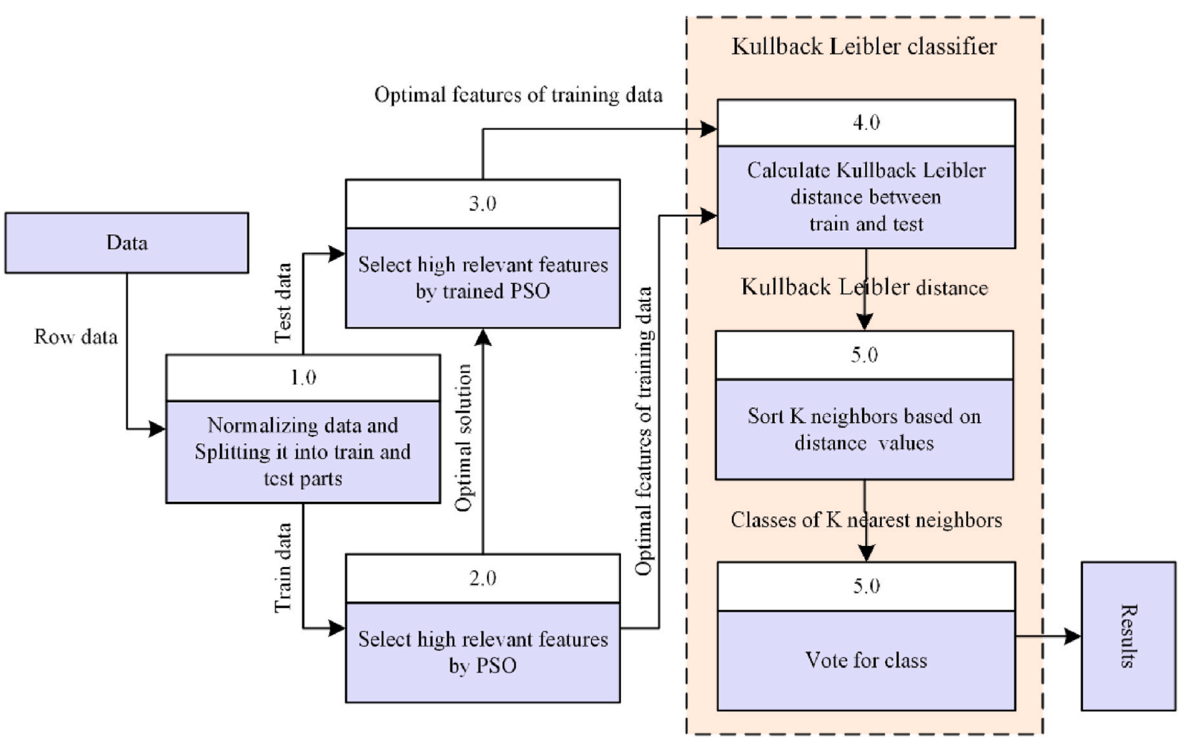

Fig. 1. Proposed solution block digram

there is no value that exceeds this range. Normalization ensures that equal values are given to all attributes whose influence on the advanced stages of processing the data and preparing it for classification is equal.

Moreover, dividing data into test_data and training_data is necessary for training and evaluating the machine learning algorithms. Technically, machine learning algorithms depend on the number of cases present or trained on in the test data. Therefore, choosing an appropriate amount of test data is necessary for the work of the machine algorithms. In addition, various models or various problems must be chosen to test the test data in order to ensure that the machine learning algorithm is trained on more than one type. To provide high accuracy in classifying machine learning algorithms of new data or test_data.

4.2. Feature selection

The feature selection process is the process of selecting particularly relevant features from the data. This paper used wrapper with PSO for selection optimal features. PSO uses natural phenomena such as flocks of birds or schools of fish to identify the most important features for classification process. The PSO algorithm is based on the determination of local and global solutions, which are assigned to each particle individually. The group of local and global solutions represents the candidate solutions.

4.3. Proposed classifier

After selecting the attributes by the optimization algorithm, this selection is applied to the test and training data. The proposed model calculates the KL measure between the test data and the training data. This measure allows finding subtle differences between structures to increase the accuracy of data separation. We arrange these values according to the smallest, and we choose the smallest appropriate values, the smallest number of values that correspond to the entirety of the data. This step is considered one of the steps that precedes the classification process. We rely on the cumulative principle to classify data. The label with the highest frequency among the K-nearest training samples that determined by KL divergence is new label of test sample.

5. Experiments and evaluation

To test the proposed model, we applied it to the classification of data related to heart disease. The classification results were evaluated by analyzing the results of the confusion matrix (CF). We extract three groups of evaluation metrics from the CF: true and false positives, accuracy, precision, recall, and f1 score. The true and false positive metrics calculate the accuracy of the model's correct and incorrect predictions. The accuracy parameter is a general measure of the accuracy of the algorithm in classifying data. The size of the sample data "unbalanced data categories" for each category is greatly affected by the evaluation criteria. Therefore, the precision and recall criterion is calculated, which is not affected by the weighting of the data. In addition to these two criteria, the f1 score metric is used to determine the relationship between precision and recall.

The confusion matrix serves as an indispensable tool in this analysis as shown in Fig. 2. This matrix, structured as a tabular juxtaposition of predicted versus actual conditions, encompasses true positive, true negative, false positive, and false negative results.

5.1. Compiled MIT-BIH-Arrhythmia dataset

The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979 [23]. This dataset is completely raw data and not compiled as the researchers need to structure data and clean it from noise before any experiments. This fact has resulted in unfair evaluation as researchers may use

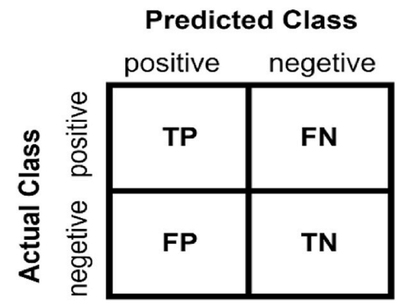

Fig. 2. Confusion matrix.

different segments with different statuses and features [24].

The compiled MIT-BIH Arrhythmia dataset is extracted from the MIT-BIH Arrhythmia Database. As the recording were digitized at 360 samples per second, 3600 samples are extracted for each segment to cover 10 s per segment. Segments are extracted with the condition of no overlapping of status, which means segments have one label one class for all signal cycles. Recodes are distributed into five classes, and each class with 100 records. Classes are described in the following [10].

According to AAMI guideline [22], Classes of ECG Compiled MIT-BIH Arrhythmia dataset: Class N: (Non-ectopic beats), and this class includes Normal Beats, Left bundle branch block, Right bundle branch block, Nodal (junctional) escape beat, and Atrial escape beat. 100 records are extracted [25]. Table 1 describes AAMI guideline.

1. Class S: (Supraventricular ectopic beats), and this class includes Aberrated atrial premature beat, Supraventricular premature beat, Atrial premature beat, and Nodal (junctional) premature beat. 100 records are extracted.
2. Class V: (Ventricular ectopic beats) and this class includes Ventricular escape beat and Premature ventricular contraction. 100 records are extracted.
3. Class F: (Fusion beats) and this class includes Fusion of ventricular and normal beat. 100 records are extracted.
4. Class Q: (Unknown beats) and this class includes Paced beat, Unclassified beat and Fusion of paced and normal beats. 100 records are extracted.

5.2. Experiments and results

This section presents a thorough comparative analysis of several machine learning classifiers, which have been evaluated based on four key metrics: Accuracy, Precision, Recall, and F1 Score. The classifiers

Table 1
Beat group and classes according to the AAMI guideline.

AAMI classes	MIT-BIH heartbeat description	Symbol
Non-ectopic beats (**N**)	Normal beats	N
	Left bundle branch block	L
	Right bundle branch block	R
	Nodal (junctional) escape beat	j
	Atrial escape beat	e
Supraventricular ectopic beats (**S**)	Aberrated atrial premature beat	a
	Supraventricular premature beat	S
	Atrial premature beat	A
	Nodal (junctional) premature beat	J
Ventricular ectopic beats (**V**)	Ventricular escape beat	E
	Premature ventricular contraction	V
Fusion beat (**F**)	Fusion of ventricular and normal beat	F
	Paced beat	/
	Unclassifiable beat	Q
Unknown beats (**Q**)	Fusion of paced and normal beats	f

being considered include the Kullback-Leibler, K-Nearest Neighbors, Support Vector Machine, Naive Bayes, Decision Tree, and Random Forest classifiers. The analysis also includes data on the performance of these algorithms both with and without the implementation of feature selection techniques.

Drawing from the experiment conducted, it can be inferred that the Kullback-Leibler classifier surpasses other established algorithms when it comes to categorizing arrhythmia conditions. Fig. 3, illustrates that the classifier's confusion matrix displays an impressive degree of accuracy in forecasting both arrhythmia and normal cases, with a substantial count of true positives (102) and true negatives (30), and relatively minimal false negatives (9) and false positives (9). This consistent performance highlights the Kullback-Leibler classifier's expertise in balancing sensitivity and specificity.

The Kullback-Leibler Classifier's confusion matrix presents a more balanced classification with a higher number of true positives and true negatives, and notably fewer false negatives and positives. This superior performance is echoed in the evaluation metrics chart, where the Kullback-Leibler Classifier outperforms the other algorithms in classification performance.

On the other side, the Gaussian Naive Bayes Classifier presents a balanced false positive and false negative rate in predicting arrhythmia, with 88 true positives and 23 false negatives, implying a slightly conservative approach, potentially missing some arrhythmic cases. For the normal predictions, it has a lower false positive rate with 13 instances but manages to correctly classify 26 normal instances, indicating a better specificity than the Decision Tree Classifier. The K-Nearest Neighbors classifier shows a high true positive rate with 107 correctly predicted arrhythmic instances and only 4 missed, which is commendable. However, it has a higher false positive rate with 21 instances, like the Decision Tree classifier, suggesting a need for parameter tuning to reduce misclassification of the normal cases. The Support Vector classifier (SVC) demonstrates a unique case where it perfectly identifies all arrhythmic cases (111 true positives) without any false negatives. The Random Forest Classifier offers a better balance with 107 true positives and 4 false negatives for arrhythmia, alongside 22 false positives and 17 true negatives.

The high performance of the proposed KLC classifier can be attributed to its basic principle of measuring the divergence between two probability distributions. It may have been successful due to its ability to discriminate between the distribution of features in normal rhythm versus different arrhythmic patterns. In the field of machine learning, classifiers such as the KLC classifier that measure divergence can provide nuanced insights into the data. Unlike other classifiers that rely on distance metrics or probability estimates, the KLC's divergence-based approach offers a unique perspective that closely aligns with the natural variability in medical data.

Fig. 4 shows the comparison result of the proposed KLC classifier without PSO with the standard machine learning algorithms in terms of evaluation metrics: accuracy, precision, recall and f1-score. It shows proposed KLC improvement in all evaluation metrics.

The proposed KLC classifier demonstrates better performance across evaluation metrics compared to its counterparts as shown Table 2. The success of proposed classifier can be ascribed to the foundational

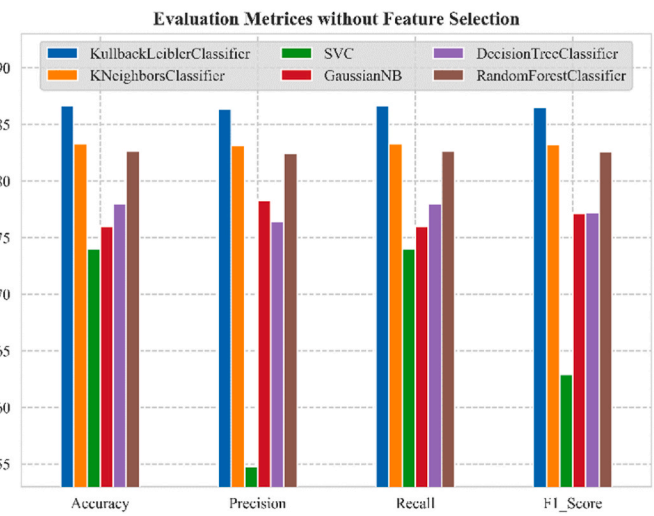

Fig. 4. Evaluation metrics without feature selection.

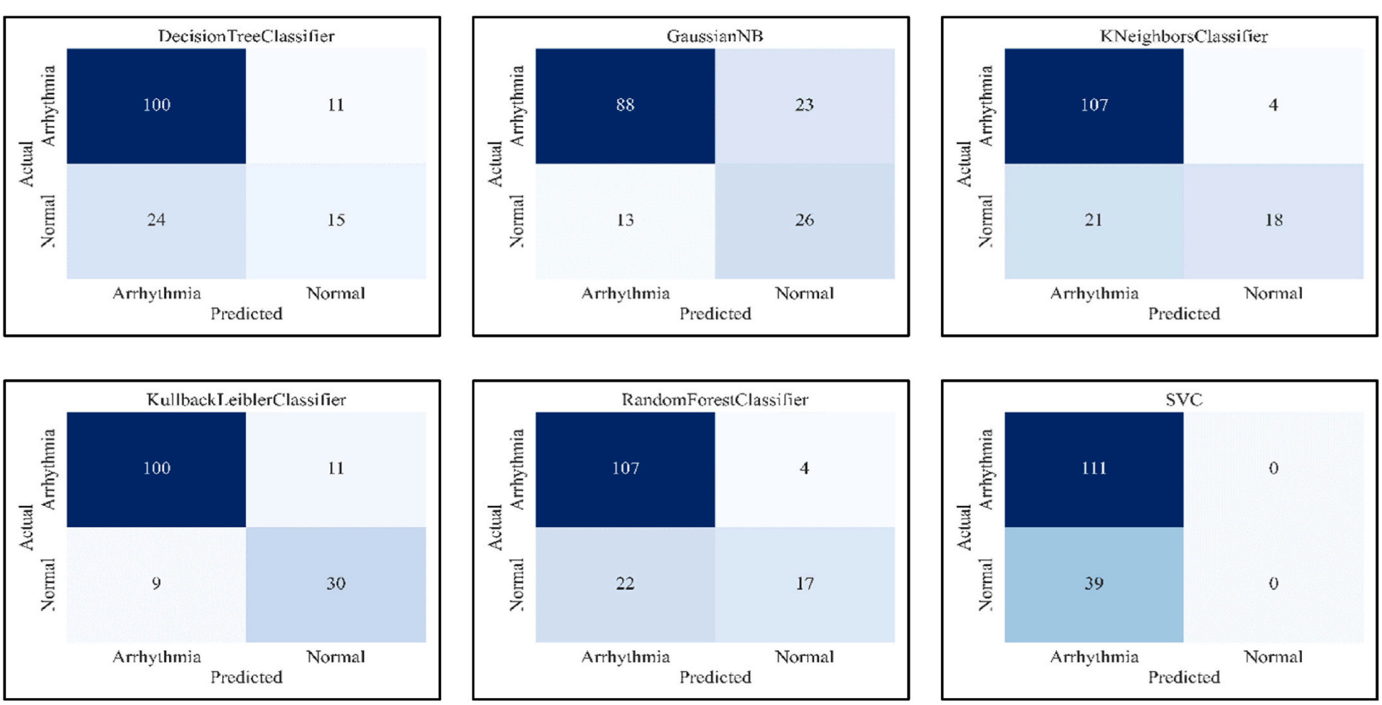

Fig. 3. Confusion matrix for the Kullback-Leibler classifier.

Table 2
Results comparesion of proposed solution without Feature selection.

Classifiers	Accuracy	Precision	Recall	F1_Score
Proposed KL Classification	86.67	86.34	86.67	86.5
KNeighborsClassifier [15]	83.33	83.13	83.33	83.23
SVC [18]	74.0	54.76	74.0	62.94
GaussianNB [20]	76.0	78.27	76.0	77.12
DecisionTreeClassifier [16]	78.0	76.44	78.0	77.21
RandomForestClassifier [22]	82.67	82.43	82.67	82.55
MLPClassifier [16]	70.67	54.1	70.67	61.28

principles upon which it operates. The Kullback-Leibler divergence measures how one probability distribution diverges from a second, expected probability distribution. In machine learning, the proposed KLC classifier leverages the divergence to effectively discriminate between classes, by quantifying the difference between the empirical distribution of the data and the expected distribution under a given model.

Fig. 5 shows the effect of the PSO feature selection on the performance of the proposed KLC and a comparison of the final results with standard machine learning.

The nature of the data and the interconnectedness and strength of the correlation between the attributes have a direct impact on the accuracy of the classification. The proposed model shows the technical strength in dealing with different data and their diverse, dynamic nature. The proposed model works to increase the consistency between the attributes before they enter the classification process, which increases the probability of their accurate distribution. the proposed KLC classifier employs a methodological estimation of probability distributions from training data, yielding a refined discernment of class demarcations.

The advantage of the proposed classifier is that it combines the process of selecting and classifying attributes at the same time. In traditional classifiers, the process of selecting attributes takes place isolated from the classifier, in a way that does not create an integrated environment for the attributes with the classification method. In this framework. The classification process in the proposed model is similar to that in deep learning models.

The classifier's success is undergirded by its operational ethos, rooted in information theory. It quantifies the divergence between predicted and actual class probability distributions. For arrhythmia classification, this methodology is particularly efficacious, as it finely distinguishes between the pattern's characteristic of normal and arrhythmic heart rhythms, by using its divergence measure that capably captures disparities in data feature distributions pertaining to each class. Table 2 compares the performance of the proposed classifier with other recent methodologies without feature selection.

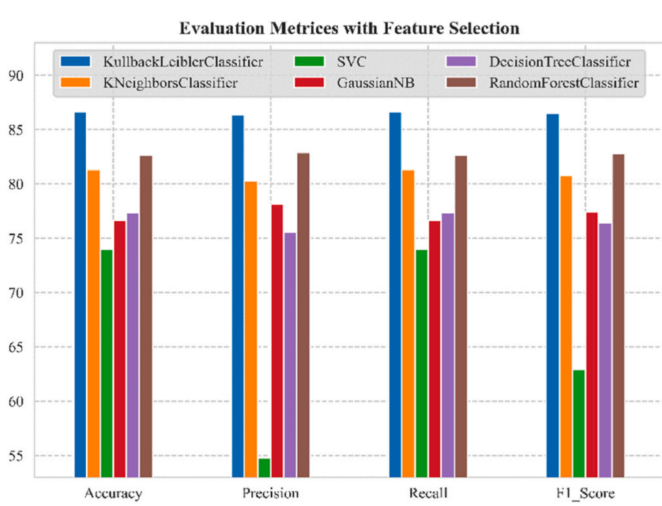

Fig. 5. Evaluation metrics with PSO feature selection.

The Kullback-Leibler classifier's deployment in medical diagnostics is meritorious, given its precision in identifying distributional nuances, which is critical in reducing misclassifications. This is highly significant in scenarios where the implications of a false negative are grave. In contrast, alternative classifiers may not consider data distributional traits with equivalent rigor, potentially leading to misdiagnoses in critical patient evaluations. Table 3 compares the performance of the proposed classifier with other recent methodologies with feature selection.

6. Conclusion

Modern medical technologies rely on integrating artificial intelligence to help doctors make quick and effective decisions in cases requiring close monitoring. The data mining algorithms suffering from significant problem when facing high-dimensional nature of data. The main challenge of these issue in extracting patterns form noisy and overlap data. Therefore, the decision made by these algorithms is not quite accurate, i.e leading to delays in decision-making and low accurate in distinguish different classes of input data. The selection of optimal features is crucial to overcome these challenges and improve the accuracy of pattern extraction. Moreover, the datamining model makes decision faster.

This paper proposes a new classification method (KLC) combines of feature optimizer and probabilistic Kullback-Leibler (KL) divergence. The first part selects the right features by using PSO as feature optimizers. The characteristics of PSO are fast and accurate in identifying the optimal features. In the second part, the Kullback-Leibler divergence (KL) method is used for data classification, which improves classification accuracy. The MIT-BIH data contains a set of heart attack readings, characterized by a high number of characteristics, more than 10,000 in each reading. This represents a real challenge to simulate the problems of the ICU and to monitor patients who suffer from heart disorders. The results show that the proposed classifier classified the MIT-BIH data with accuracy rate more than 86%. In future work, we suggest developing a deep learning model by combines KLC principles with deep learning.

Ethical statement

This to confirm that the proposed study has not involved with animals or humans.

Therefore, not Ethical approval is needed.

CRediT authorship contribution statement

Dhiah Al-Shammary: Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Mohammed Radhi:** Writing – original draft, Software, Formal analysis, Conceptualization. **Ali Hakem AlSaeedi:** Visualization, Investigation. **Ahmed M. Mahdi:** Supervision, Investigation, Formal analysis, Conceptualization. **Ayman Ibaida:** Writing – review & editing, Validation, Formal analysis. **Khandakar Ahmed:** Writing – review & editing, Validation, Formal analysis.

Table 3
Results comparesion of proposed solution with feature selection.

Classifiers	Accuracy	Precision	Recall	F1_Score
Proposed KL Classification	86.67	86.34	86.67	86.5
KNeighborsClassifier [15]	81.33	80.29	81.33	80.81
SVC [18]	74.0	54.76	74.0	62.94
GaussianNB [20]	76.67	78.17	76.67	77.41
DecisionTreeClassifier [16]	77.33	75.57	77.33	76.44
RandomForestClassifier [22]	82.67	82.9	82.67	82.78
MLPClassifier [16]	73.33	54.63	73.33	62.62

Declaration of competing interest

No conflict of interest exists.

Acknowledgement

This research was supported/partially both the University of Al-Qadisiyah (Iraq) and Victoria University (Australia). We are thankful to both Universities who provided labs, devices, and expertise that greatly assisted the research.

References

[1] Singhal S, Kumar M. A systematic review on artificial intelligence-based techniques for diagnosis of cardiovascular arrhythmia diseases: challenges and opportunities. Arch Comput Methods Eng 2023;30(2):865–88. https://doi.org/10.1007/s11831-022-09823-7.

[2] Roland G, Padhi S, Kayalvili S, Cloudin S, Kumar A. An automated system for arrhythmia detection using ECG records from MITDB. In: 2022 International Conference on automation, computing and renewable systems (ICACRS). IEEE; 2022. p. 26–33. https://doi.org/10.1109/ICACRS55517.2022.10029289.

[3] Seng KP, et al. Atrial fibrillation detection based on electrocardiogram features using modified windowing algorithm. In: Asia simulation conference. Springer; 2023. p. 483–95. https://doi.org/10.1007/978-981-99-7240-1_38.

[4] Udawat AS, Singh P. An automated detection of atrial fibrillation from single-lead ECG using HRV features and machine learning. J Electrocardiol 2022;75:70–81. https://doi.org/10.1016/j.jelectrocard.2022.07.069.

[5] Oleiwi ZC, AlShemmary EN, Al-augby S. Efficient ECG beats classification techniques for the cardiac arrhythmia detection based on wavelet transformation. International Journal of Intelligent Engineering & Systems 2023;16(2). https://doi.org/10.22266/ijies2023.0430.16.

[6] chaffat Oleiwi Z, AlShemmary EN, Al-augby S. Identify best learning method for heart diseases prediction under impact of different datasets characteristics. Journal of Kufa for Mathematics and Computer 2023;10(1):27–41. https://doi.org/10.31642/JoKMC/2018/100104.

[7] Oleiwi ZC, AlShemmary EN, Al-augby S. Adaptive features selection technique for efficient heart disease prediction. Journal of Al-Qadisiyah for computer science and mathematics 2023;15(1):1–18. https://doi.org/10.29304/jqcm.2023.15.1.1137.

[8] Ojha MK, Wadhwani S, Wadhwani AK, Shukla A. Automatic detection of arrhythmias from an ECG signal using an auto-encoder and SVM classifier. Physical and Engineering Sciences in Medicine 2022;45(2):665–74. https://doi.org/10.21203/rs.3.rs-981164/v1.

[9] Düzyel O, Kuntalp M. Data augmentation with GAN increases the performance of arrhythmia classification for an unbalanced dataset. arXiv preprint arXiv:2302.13855 2023. https://doi.org/10.48550/arXiv.2302.13855.

[10] Wang L-H, et al. Three-heartbeat multilead ECG recognition method for arrhythmia classification. IEEE Access 2022;10:44046–61. https://doi.org/10.1109/ACCESS.2022.3169893.

[11] Granero-Belinchón C, Roux SG, Garnier NB. Kullback-Leibler divergence measure of intermittency: application to turbulence. Phys Rev 2018;97(1):013107. https://doi.org/10.1103/PhysRevE.97.013107.

[12] Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat 1951;22(1):79–86. https://doi.org/10.1214/aoms/1177729694.

[13] Aziz MR, Alfoudi AS. Feature selection of the anomaly network intrusion detection based on restoration particle swarm optimization. International Journal of Intelligent Engineering & Systems 2022;15(5). https://doi.org/10.22266/ijies2022.1031.51.

[14] Al-Shammary D, Albukhnefis AL, Alsaeedi AH, Al-Asfoor M. Extended particle swarm optimization for feature selection of high-dimensional biomedical data. Concurrency Comput Pract Ex 2022;34(10):e6776. https://doi.org/10.1002/cpe.6776.

[15] Madan P, Singh V, Singh DP, Diwakar M, Pant B, Kishor A. A hybrid deep learning approach for ECG-based arrhythmia classification. Bioengineering (Basel) Apr 2 2022;9(4). https://doi.org/10.3390/bioengineering9040152.

[16] Subasi A, Dogan S, Tuncer T. A novel automated tower graph based ECG signal classification method with hexadecimal local adaptive binary pattern and deep learning. J Ambient Intell Hum Comput 2021;14(2):711–25. https://doi.org/10.1007/s12652-021-03324-4.

[17] Mian Qaisar S, Khan SI, Srinivasan K, Krichen M. Arrhythmia classification using multirate processing metaheuristic optimization and variational mode decomposition. Journal of King Saud University - Computer and Information Sciences 2023;35(1):26–37. https://doi.org/10.1016/j.jksuci.2022.05.009.

[18] Kim YK, Lee M, Song HS, Lee S-W. Automatic cardiac arrhythmia classification using residual network combined with long short-term memory. IEEE Trans Instrum Meas 2022;71:1–17. https://doi.org/10.1109/tim.2022.3181276.

[19] Mathunjwa BM, Lin YT, Lin CH, Abbod MF, Sadrawi M, Shieh JS. ECG recurrence plot-based arrhythmia classification using two-dimensional deep residual CNN features. Sensors Feb 20 2022;22(4). https://doi.org/10.3390/s22041660.

[20] Kutluana G, Türker İ. Classification of cardiac disorders using weighted visibility graph features from ECG signals. Biomed Signal Process Control 2024/01/01/2024;87:105420. https://doi.org/10.1016/j.bspc.2023.105420.

[21] Qin K, Huang W, Zhang T, Zhang H, Cheng X. A lightweight SelfONN model for general ECG classification with pretraining. Biomed Signal Process Control 2024;89:105780. https://doi.org/10.1016/j.bspc.2023.105780. 2024/03/01/.

[22] Li Y, Qian R, Li K. Inter-patient arrhythmia classification with improved deep residual convolutional neural network. Comput Methods Progr Biomed Feb 2022;214:106582. https://doi.org/10.1016/j.cmpb.2021.106582.

[23] Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Eng Med Biol Mag 2001;20(3):45–50. https://doi.org/10.1109/51.932724.

[24] Kuila S, Dhanda N, Joardar S. ECG signal classification and arrhythmia detection using ELM-RNN. Multimed Tool Appl 2022;81(18):25233–49. https://doi.org/10.1007/s11042-022-11957-6.

[25] Rahhal MMA, et al. Classification of AAMI heartbeat classes with an interactive ELM ensemble learning approach. Biomed Signal Process Control 2015/05/01/2015;19:56–67. https://doi.org/10.1016/j.bspc.2015.03.010.