Int. J. Mach. Learn. & Cyber.
DOI 10.1007/s13042-016-0594-z

ORIGINAL ARTICLE

Cardiac arrhythmia classification using multi-granulation rough set approaches

S. Senthil Kumar[1] · H. Hannah Inbarani[1]

Received: 10 November 2015 / Accepted: 10 September 2016

Abstract Cardiovascular disease is a most important reason for human death in modern society. Electrocardiogram (ECG) signal deals with valuable information about functioning of the heart. For that reason, ECG investigation signifies an efficient way to identification and treat different types of cardiac arrhythmia diseases. Nowadays various pattern classification methods has been developed for the classification of ECG signals. These classification methods helps to physician for diseases diagnosis. A multi-granulation rough set (MGRS) has become a new direction of rough set theory, which is based on multiple binary relations on the universe of discourse. In this present study, Multi-granulation rough set based classification approaches (Pessimistic Multi-Granulation Rough Set (PMGRS) and Optimistic Multi-Granulation Rough Set (OMGRS)) are applied to mine appropriate rules to explore better decision making process. The experiments were conducted on the ECG data from the Physionet arrhythmia database to classify five kinds of normal and abnormal signals. In the classification process, Feature extraction played an important role. And we have used two kinds feature extraction methods (1) Pan Tomkins (PT) feature extraction method. This method used to extract the morphological features are P, Q, R, S, T peak intervals, which is also used to determine heart rate. (2) Wavelet transform (WT) feature extraction method. This method used to extract the wavelet coefficients. Both two methods are successfully applied to (ECG signal) classification. The proposed multi-granulation rough set rule based classification methods is validated using the first 24 channel of the ECG signal records of the MIT-BITH arrhythmia database, and achieves finding high accuracies. Experimental results show that the proposed classification techniques significantly outperforms other well-known techniques.

Keywords ECG · Feature extraction · Pan tomkins (PT) · Wavelet transform (WT) · Classification · Multi-granulation rough set

1 Introduction

Electrocardiography (ECG) deals with the electrical activity of the heart. The state of cardiac health is generally reflected in the shape of the ECG waveform. An electrocardiogram can contain important peak points to the nature of diseases afflicting the heart. However, bio-signals being non-stationary, such pointers can occur at random on the time scale. Therefore, the study for effective diagnostics of ECG patterns and can be identified heart abnormalities. Sometimes the heart disorder symptoms might not be found. So ECG signals are observed for several hours and it's create a facts of massive volumes of data. Consequently the task of abnormalities detection from ECG signal becomes deadly and time consuming process for medical experts. Thus there is need for system based techniques for diagnosis of heart disease from the large volumes of ECG signal data [36]. Cardiac arrhythmia is the most common cardiac diseases and its early diagnosis is very important. Systematize the detection methods help in the diagnosis of cardiac arrhythmia in electrocardiograms (ECGs) have been proposed during the last 15 years. These techniques can be come

✉ S. Senthil Kumar
pkssenthilmca@gmail.com

H. Hannah Inbarani
hhinba@gmail.com

[1] Department of Computer Science, Periyar University, Salem 636011, Tamil Nadu, India

together depending on the computational paradigm on which they are based (rule-based methods, neural networks, fuzzy set systems, pattern recognition, signal processing, etc.). Rule-based method shows have certain advantages. Such as shortest transformation of medical knowledge to rules, low computational load and description of the diagnostic decisions. On the other hand, their diagnostic value depends on the appropriate selection and combination of the rules and the method for the extraction of feature values used in the rules. The most challenging problem faced by today's ECG examination is the huge difference in the morphologies of ECG signs. [43, 45]. To conquer those problems, many researches have been conducted to discover effective signal evaluation and pattern recognition techniques for ECG signal. ECG signal data were obtained from the physionet MIT-BIH arrhythmia database [16]. In this research work, ECG signal recognition system divided into a sequence of stages. Such as, feature extraction and classification. ECG signal characteristics such as, inconsistency on feature extraction and classification, unclassifiable beats and a strong class unbalance, Prior to recording, the ECG signals were processed to remove noise due to muscle tremors, spikes etc. In this study, the cardiac arrhythmias were classified into five categories, namely (i) Left Bundle Branch Block (LBBB) (ii) Normal Sinus Rhythm (NSR) (iii) Pre-Ventricular Contraction (PVC) (iv) Right Bundle Block (RBBB) and (v) Paced rhythm (PR). First stage with feature extraction from the going on sample, which is the conversion of the pattern to functions which might be appeared as a summarized representation. The information visible in the ECG signal is useful in detecting the disease and affecting the heart. The problem of new model established has always been of interest to the pattern recognition community. Initially, the goal of new model was to improve the efficiency of decision making by adopting rules. We are using two kinds of feature extraction methods. Such as (1) Pan-Tompkins (PT) and (2) Wavelet Transform (WT). Pan-Tompkins (PT) algorithm [15] is used for extracting morphological features from ECG signals. The morphological features are P, QRS Complex and T. This algorithm searches for the local maximum or minimum peak near the beat label to establish local points. The P-QRS-T waves to representing the cardiac function [40]. Statistical features are extracted from ECG signal using wavelet Transform (WT) method [24, 42, 47]. Nonetheless, a wavelet transform-based method is used to resolve the deficiencies of the existing methods. The extracted morphological features and Statistical features are classified into five different classes. The signal decision classes are normal sinus rhythm (NSR) and various cardiac arrhythmias including Left Bundle Branch Block (LBBB), premature ventricular contraction (PVC), Right Bundle Branch Block (RBBB), and Paced Rhythm (PR).

The second stage is ECG signal classification (cardiac arrhythmia or not) using medical knowledge in the form of features obtained in the first stage. In the everyday medical practice when a cardiologist uses ECG to diagnose cardiac arrhythmia or not, in this two types of features and rough set based methods are examined. Rough set theory proposed by Pawlak [20, 51, 52] is an extension of the classical set theory. Rough set theory [20, 46] gives a powerful method of managing uncertainties and can be take on for responsibilities together with information dependency analysis, attribute identification, dimensionality reduction, and pattern classification. Rough set concept offers an approach to approximation of units that ends in beneficial kinds of granular computing. Information granules [50] have played a significant role in human understanding, reasoning, and decision making. Such information processing is called information granulation. Zadeh proposed and discussed the issue of fuzzy information granulation in 1997 [55]. Then, the basic idea of information granulation has been applied to many fields, such as theory of rough sets. With a granular computing point of view, an equivalence relation on the universe can be regarded as a granulation, and a partition on the universe can be regarded as a granulation space [50]. Hence, the classical rough set theory is based on a single granulation (only one equivalence relation). When the rough set approach is used to unravel decision rules from a given information system, two types of decision rules may be derived. Based on the lower approximation of a decision class, certain information can be discovered and certain rules can be derived, whereas by using the upper approximation of a decision class, uncertain or partially certain information may be discovered and possible rules induced. Various approaches using rough set theory have been proposed to discover decision rules from data sets taking the form of decision [57–64]. Such methods are neighborhood rough set [69], rough set approximation based on dynamic granulation [29], Decision-theoretic rough set models [28, 67] and Composite rough sets theory offered various relation [68]. Then, researchers have proposed extended the rough set to the multi-granulation Rough sets. The multi-granulation rough set is different from Pawlak's rough set model because the former is constructed on the basis of a family of binary relations instead of a single indiscernibility relation. The concept of multi-granulation rough set theory was firstly proposed by Qian [19, 49]. However, when the rough set is based on many granulations induced from several equivalence relations. In their approach, two different models have been defined. The first one is the optimistic multi-granulation rough set [16, 18], the second one is the pessimistic multi-granulation rough set. These two methods deals with uncertainty and finds approximation space for multi-granules [17–19, 21–23]. A decision rule maps an observation to a proper action. Decision rules composition plays an important role in the theory of uncertainty mapping. Recently rough set extension methods has been used to deal with certain problems of signal processing. Informative features are extracted from ECG signal using Pan Tomkin's (PT) [15] algorithm and Wavelet Transform (WT) [30]. Rough set based classification methods only cope with discrete variables. Extracted ECG features are a continuous values in this difficult situation discretization process is necessary. Discretization is the process of transforming numerical variables into discrete ones. In this paper, a Min–Max discretization method is applied for ECG data set [32]. In this study, Multi-granulation rough set based classification methods viz., Pessimistic Multi-Granulation Rough set (PMGRS) rule based classification and Optimistic Multi-Granulation Rough set (OMGRS) rule based classification are applied for the classification of the extracted features from ECG signal. In the proposed work, two types of decision rules are derived from approximations space. Deterministic rules correspond to the lower approximation and non-Deterministic rules correspond to the upper approximation. Rough set rule based techniques is dealing with the similar rule retrieved quickly, hence has the same indiscernibility and reduce the computational time. The approximation space considers a training set to match testing set and to compute the classification accuracy. These two methods are the most appropriate methods for detecting ECG cardiac arrhythmia. These two classification methods are compared with several bench mark algorithms such as Naïve bayes, Multi-Layer Perceptron (MLP), multi-class classifier, J48, JRip and Decision Table. Our proposed methods provide high accuracy than other existing methods. The remaining part of this paper is organized as follows: motivation and contribution of our study are described in Sect. 2 and Sect. 3. Section 4 described related work of present study. The proposed methodology is described in Sect. 5. Experimental results and Discussion are described in Sect. 6. Finally, concluding statements are presented in Sect. 7.

2 Motivation

- Electrocardiogram (ECG) signals are generally contemplated in the shape of ECG waveform and heart rate. ECG, if properly analyzed, can offer records regarding diverse diseases related to heart.
- ECG signal visual analysis can not be relied upon and the possibility of the analyst missing the vital facts is high. Consequently, computer based evaluation and classification of diseases may be very beneficial in diagnosis.
- Numerous contributions had been made in literature concerning detection and classification of ECG signal. Maximum of them use time area illustration of the ECG waveforms, on the premise of which many specific capabilities are defined, permitting the popularity between the beats belonging to exclusive training.
- The maximum challenging problem faced by using nowadays ECG evaluation is the large difference in the morphologies of ECG waveforms. Accordingly our fundamental goal is to give you an easy classification technique having accurate computational evaluation without compromising with the efficiency.
- In this paper, ECG signal analysis carry out using Pan Tomkins method and wavelet transform method for feature extraction and classification has been done using Multi-granulation rough set rule based classification (PMGRS and OMGRS) methods. The motivation for this classification evaluation is to improve the effectiveness of the algorithms.

3 Contribution

- This work aims to develop a method for handling data inconsistency and inaccuracy in the analysis of Electrocardiograms (ECGs) using Multi-granulation Rough set (MGRS) based classification approaches.
- Minimized rule sets are generated using lower and upper approximations and five types of ECG heart rhythms such as Normal sinus rhythm (NSR), Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Paced rhythm(PR) and Ventricular Premature Contraction (PVC) [38] are analyzed.

3.1 Complexity

- The complexity analysis computed for various real time databases used for engineering, medical applications and etc.… are very difficult. In this work, Multi-granulation rough set methods (PMGRS and OMGRS) is applied for classification. Approximating lower and upper rough calculation the computational complexity of the algorithm since they are computed before completing of the main algorithm [41].
- Hence Optimistic Multi-Granulation Rough Set (OMGRS) based classification technique, is able to achieve lower computational complexity of O (n x m), where m is the number of attributes in $(A_1 \cap D) \cup (A_2 \cap D),\ldots \cup (A_n \cap D)$ and n is the number of entities in U. Pessimistic Multi-Granulation Rough Set (PMGRS) based classification technique, is able to achieve lower computational complexity of O ($n \times m$), where m is the number of features in $(A_1 \cap D) \cap (A_2 \cap D),\ldots \cap (A_n \cap D)$ and n is the number of objects in U and higher

classification accuracy is achieved as compared to the other methods.

4 Related work

This phase discusses a variety of strategies proposed in advance in literature for feature extraction, and categorization of ECG signal. There are various studies in feature extraction and class as stated in Table 1 focusing at the prognosis of ECG signal.

This paper proposes smart intelligent system for ECG sign cardiac arrhythmia disease diagnosis. In the following sections, the basic notions of rough set principle and multigranulation rough set methods (PMGRS and OMGRS) are provided.

5 Methodology

The technique for proposed work is shown in the waft of Fig. 1. The equal period of signal is taken from the MIT-BIH for the processing. This signal is further pre-processed and noise is removed and relevant features are extracted using Pan Tomkin's (PT) and Wavelet Transform (WT) approach. After preprocessing, discretization of the continuous attributes is performed using existing discretization method (Min–Max) and extracted features are set as input to the proposed two supervised Multi granulation rough set classification methods (1). Optimistic Multi-Granulation rough set (OMGRS) and (2). Pessimistic Multi-Granulation rough set (PGMRS)) applied for ECG signal waves. This ECG signal is as well classified using certain other data mining techniques. Finally, Compute validation measure for two proposed methods and other data mining algorithms results are compared.

5.1 Signal acquisition

This is first level of signal processing. The database gathering is one of the most major task of signal processing. The records are used on this research is the ECG signals [45] data from the MIT–BIH (Massachusetts Institute of Technology—Boston's Beth Israel Hospital) Cardiac Arrhythmia database available on physionet website (http://www.physionet.org/physiobank/database) [16]. This database carries 48 files divided into two portions first one is of 23 records decided on at random from this set, and another one includes 25 records. Each of the 48 records is barely over 30 min lengthy. Figure 2 indicates the original input ECG signal.

5.2 Feature extraction

Various signal processing feature extraction techniques that are in use for processing time series databases for extracting relevant features. A feature based approach suitable for complex signal processing, when ECG waveform signals used for diagnosis process. In this paper, we have to apply two feature extraction methods for ECG waveform signals. Such as (1) Morphological Feature Extraction and (2) Wavelet transform (WT) Feature Extraction.

5.2.1 Morphological feature extraction

ECG signals of the persons with arrhythmia do not have regular beat and waveforms which the points in P, QRS and T cannot be observed manually. For that reason, the morphological belongings of the heart beat such as PQRST width are the main rules of arrhythmia recognition used by doctors. The morphological features extracted are P, QRS and T peak points. The R-peak has the maximum amplitude in an ECG signal. Hence the R-peak is identified by using the Pan Tomkins's algorithm. The additional peaks are recognized by means of bypass thru the windowing characteristic on both feature of R peaks. The Q and S peaks are located with the aid of skip through at the left and proper aspect of the R height within the targeted window and locating the negative top values. By using bypass thru the left side of the Q peaks, the most value is located because the P peaks. Similarly by way of traversing the proper side of the S peaks, the maximum cost is placed as the T peaks within the window. The heart rate for each person is calculated by finding the distance between two peaks (R–R). It is projected that RR intervals in an ECG signal of a healthy person are very nearly the similar while RR intervals of arrhythmia beats of an unhealthy person are varying. However, RR interval can be a weak feature for the classification of arrhythmia types, because it does not contain information about the waveform complexity and other segments of the ECG signal. For this reason, another feature extraction method represents the waveform complexity of the EGC signal [5–7, 15]. The extracted morphological features (P, Q, R, S and T) are shown in Fig. 3.

5.2.2 Wavelet transform (WT)

In this paper, 5 type of different classes of feature set of function set are used belonging to the everyday ECG signal along with parameters, and the variance of wavelet transform (WT) detail coefficients for the one-of-a-kind scales. Functioning ECG signal used for diagnosis are usually labeled with the aid of a non-stationary time conduct. For such patterns, time representations are needed. The

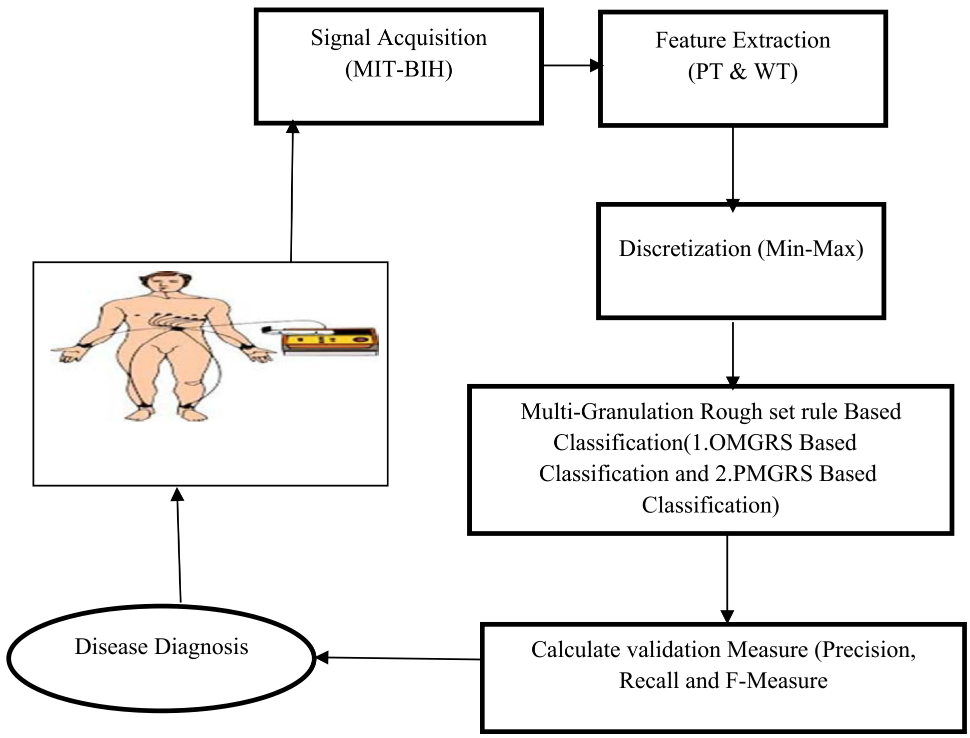
Table 1 Related work for this study

Authors	Techniques	Proposes
Korurek et al. [1], Pasolli et al. [2], Melgani et al. [3], Ubeyli [4]	Classification and feature extraction	An electrocardiographic (ECG) signal can exhibit considerable information about the functioning of a heart by analyzing the number of waves and intervals in the ECG signal. The aptitude of a detection algorithm to retrieve these components by design is thus an important factor. Several studies are available in the bio-signal literature regarding detection of cardiac arrhythmias. In general, in some of the current studies, these studies are different in look to their feature extraction, classification, and/or the number of methods adopted for ECG cardiac arrhythmias
Qibin Zhao et al. [30]	Feature extraction	In this work, wavelet transform is used to extract the wavelet coefficients features of each ECG sign and autoregressive modelling (AR) is also applied to attain the temporal systems of ECG waveforms. Then the Support Vector Machine (SVM) with Gaussian kernel is used to classify various type ECG heart rhythm
Hari Mohan Rai et al. [31]	Feature extraction	The proposed system deals with ECG signal evaluation based on Artificial Neural Network based (discrete wavelet transform and morphology) features. This study has proposed a technique to certainly classify ECG signals statistics into two classes (abnormal and normal) the usage of various neural classifiers
Kohler et al. [5], Christov et al. [6]	Feature extraction	This work compares the performance of all these QRS detectors, morphological and time–frequency ECG descriptors for heartbeat classification
Pan et al. [15]	Feature extraction	Real-time algorithm for detection of the QRS complexes of ECG signals is proposed on this work. It reliably recognizes QRS complexes based totally upon virtual analyses of slope, amplitude, and width. A special digital band pass filter out reduces false detections as a result of the various kinds of interference found in ECG signals
Giovanni Acampora et al. [14]	Classification	In this work, a fuzzy inference engine based on Type-2 Fuzzy Markup Language suitable of evaluating the ECG and heart rate variation, where the proposed approach has diagnostic framework produced good performances both in terms of precision and recall
Karpagachelvi et al. [24]	Classification	The Extreme Learning Machine (ELM) method that is proposed and compared with support vector machine (SVM) based automatic classification approach is applied for the detection of cardiac arrhythmia. The obtained results clearly confirm the superiority of the ELM approach as compared with traditional classifiers
Kumari et al. [26]	Classification	Many algorithms exist for ECG detection/classification. In this paper, RR interval based ECG classification procedures for arrhythmia beat category is investigated. Extracted RR interval data is used to classify ECG signal via a boosting algorithm and Fuzzy Unordered Rule Induction set of rules (FURIA). Finally, categorization performance is evaluated
Senthilkumar et al. [9, 10, 11], Udhaya Kumar et al. [12]	Classification	Rough set theory has been successfully applied in many real-time problems in medicine, engineering and etc.… The rough set based classification techniques are good for various medical disease diagnosis, such as heart valve diagnosis, medical diagnosis, and neo jaundice diagnosis. Rough set based classification methods are applied in this work to diagnose the medical diseases
Michał Kania et al. [7]	Classification	ECG signal is being more sensitive to real morphology variations. The proposed system deals with Multiple parameter analysis which gives whole assessment, screening time and amplitude's scaling effects of ECG signal due to electrode displacements. Proposed system results are achieved. It may help to choose alternate localities of precordial electrodes when there is essential to expose the space on the body surface for other diagnostic procedures
Ali Khazaee et al. [27]	Classification	In this work, a novel approach power spectral-based hybrid particle swarm optimization-support vector machine (SVMPSO) classification technique is proposed which significantly provides better performance over the SVM

Table 1 continued

Authors	Techniques	Proposes
Senthil Kumar et al. [11], Ahmet Mert Niyazi Kılıc et al. [25], Kumar et al. [69]	Feature extraction and classification	Cardiac arrhythmia detection and classification is critical in medical cardiology, particularly when achieved in real time. This paper investigates the detection and classification of ECG arrhythmias. In order to establish the correct diagnosis and define the proper treatment, various pattern reorganization techniques are applied for ECG cardiac arrhythmia datasets

Fig. 1 Proposed methodology



prevalence features in addition to the temporal features may be categorized with recognize to uncertainty popular. The WT of a signal f(x) is defined as: [39]

$$W_s f(x) = f(x) * \Psi s(x) = \frac{1}{s} \int_{-\infty}^{+\infty} f(t) \Psi\left(\frac{x-t}{s}\right) dt$$

where s is scale factor. $\Psi_s(x) = \frac{1}{s}\Psi\left(\frac{x}{s}\right)$ is the dilation of a basic wavelet $\Psi(x)$ by the scale factor s. Let $s = 2_j$ ($j \in Z$, Z is the integral set). The WT of a signal f(n) can be calculated with algorithm as follows:

$$S_{2i} f(n) = \sum_{k \in z} h_k S_{2j-1} f(n - 2^{j-1} k)$$

$$W_{2i} f(n) = \sum_{k \in z} g_k S_{2j-1} f(n - 2^{j-1} k)$$

where S2j is a smoothing operator. S2j f(n) = aj, aj is low frequency coefficients that is the approximation of original signals at the same time as w2j f(n) = dj, dj is high frequency coefficients that is the element of original signals. By means of the multi- resolution illustration it is viable to explain the signal shape by only some coefficients inside the wavelet area. The choice of suitable wavelet and the range of decomposition degree may be very critical in analysis of signals the use of the WT. The wide variety of decomposition ranges is selected primarily based on the dominant frequency components of the signals. The level are chosen such that those parts of the signal that correlate well with the frequencies required for category of the signal are retained inside the wavelet coefficients. The wavelet used on this work is one member of the Daubechies households. The wide variety of decomposition tiers turned into chosen to be 4. Therefore, the ECG signals had been decomposed into the information d1 − d4 and one approximation a4. Usually, exams are accomplished with extraordinary forms of wavelets and the one which offers maximum performance is selected for the precise

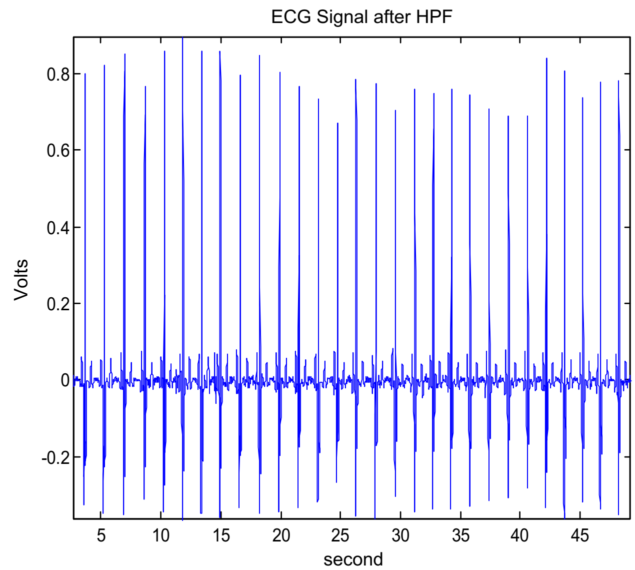


Fig. 2 Original input ECG signal

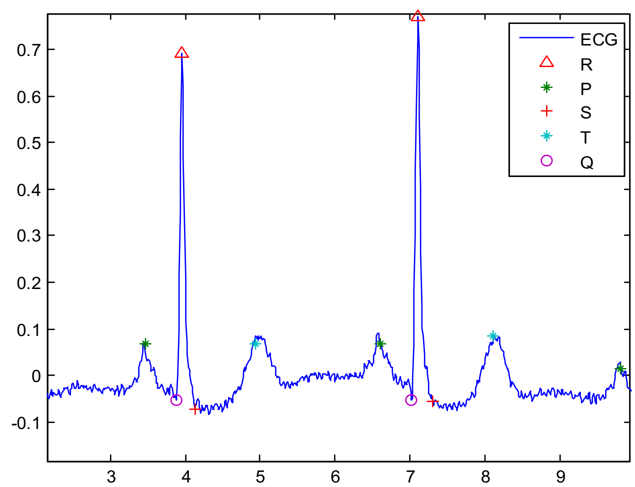


Fig. 3 Extracted P, QRS and T features

application. Consequently, the Daubechies wavelet of order 8 changed into chosen ECG data sets [30].

5.3 Discretization

Discretization is an essential preprocessing before conducting uncertainty based classification on the continuous data. Discretizing continuous-valued attributes that offers us gain in computational performance that consequences in dashing up the evaluation process for Continuous-valued attribute discretization. Biomedical data, specifically signal medical data, encompass a big number of variables sampled in abnormal way, often such as each time factor and time durations, hence offering numerous demanding situations for analysis and to carry out diverse classification obligations based totally at the ECG signal data, inclusive of for functions of sickness diagnosis [44]. In this paper, although, we first implemented a Min–Max technique that learns in a supervised way how to optimally discretize each ECG data, a good way to remodel the time-point series right into a time-interval collection, mine those time series, and carry out type after which proceed to evaluate the method within numerous extraordinary cardiac arrhythmias [32].

5.4 Basic concepts of rough set and multi-granulation rough sets

5.4.1 Rough set

Rough set theory was initiated by Pawlak [20, 51–54] for dealing with vagueness in information systems. This theory handles the approximation of an arbitrary subset of universe by two definable or observable subsets called lower and upper approximations. Classical definitions of lower and upper approximations, sometimes called Pawlak's rough approximations, were originally introduced with reference to an indiscernibility relation which is assumed to be an equivalence relation. This model is useful in the analysis of data presented in terms of complete information systems and complete decision tables. It has been successfully applied to machine learning, intelligent systems, inductive reasoning, pattern recognition, image processing, signal analysis, knowledge discovery, decision analysis, expert systems, and many other fields [11, 34, 46].

Definition 1 Let be an equivalence relation on U. The pair (U, R) is called a Pawlak approximation space. The equivalence relation R is often called an indiscernibility relation. Using the indiscernibility relation R, one can define the following two rough approximations:

$$R_*(X) = \{x \in U : [x]_R \subseteq X\} \quad (1)$$

$$R^*(X) = \{x \in U : [x]_R \bigcap X \neq \theta\} \quad (2)$$

$R_*(X)$ and $R^*(X)$ are called the Pawlak lower approximation and the Pawlak upper approximation of X, respectively [37].

5.4.2 Multi-granulation rough sets

Granular computing may be regarded to as a label of the family of theories. The basic ideas of granular computing have appeared in many fields, such as interval analysis, quantization, rough set theory. There are many reasons for the study of granular computing [50]. When a problem involves incomplete, uncertain, or vague information, it may be difficult to differentiate distinct elements and one is forced to consider granules. Information granules have played a significant role in human understanding, reasoning, and decision making. Zadeh proposed and discussed the issue of fuzzy information granulation in 1997 [55]. Then, the basic idea of information granulation has been applied to many fields, such as theory of rough sets. With granular computing point of view, an equivalence relation on the universe can be regarded as a granulation, and a partition on the universe can be regarded as a granulation space. Hence, the classical rough set theory is based on a single granulation (only one equivalence relation). Then, the researchers have proposed extended the rough set to the multi-granulation Rough sets. The multi-granulation rough set is different from Pawlak's rough set model because the former is constructed on the basis of a family of binary relations instead of a single indiscernibility relation. The concept of multi-granulation rough set theory was firstly proposed by Qian [19, 49]. In their approach, two different models have been defined. The first one is the optimistic multi-granulation rough set [16, 18], the second one is the pessimistic multi-granulation rough set Multi-granulation rough set based classification methods.

5.4.3 Pessimistic multi-granulation rough sets (PMGRS)

Rough set models are based on single granulation; they also called the single granulation rough sets. In the Pessimistic Multi-Granulation Rough Set (PMGRS), the target is approximated through the multiple granulations. It's different from the upper approximation of optimistic multi-granulation rough set and the upper approximation of pessimistic multi-granulation rough set is represented as a set in which objects have non-empty intersection with the target in terms of at least one granular structure [11, 17–19, 21–23, 33–35].

Definition 2 Let S = (U, AT, f) be called a complete target information system if values of all attributes AT and D for all objects from U are regular (known), where AT is called the conditional attributes and D is called the decision attribute, $\hat{P}$, $\hat{Q}$ be two partitions on the universe U, and X $\subseteq$ U. $\hat{P}(\mathrm{x})$ and $\hat{Q}(\mathrm{x})$ are subset of $\hat{P}$ and $\hat{Q}$. The lower approximation and the upper approximation of X in U are defined by the following:

$$\underline{X}\hat{P} + \hat{Q} = \{x : \hat{P}(x) \subseteq \mathrm{X} \text{ and } \mathrm{x} : \hat{Q}(x) \subseteq \mathrm{X}\} \tag{3}$$

$$\bar{X}\hat{P} + \hat{Q} = \sim \underline{(\sim X)}\hat{P} + \hat{Q} \tag{4}$$

5.4.4 Optimistic multi-granulation rough sets

In the optimistic multi-granulation rough set (OMGRS), the objective is approached through the multiple equivalence relation. In lower approximation, the word optimistic is used to precise the idea that in multiple independent granulations, we need only at least one of the equivalence to satisfy with the inclusion condition between data granule and target concept. The upper approximation of OMGRS is defined by the complement of the lower approximation [11, 17–19, 21–23, 34].

Definition 3 Let S = (U, AT, D, f) be called a complete target information system, where AT is called the conditional attributes and D is called the decision attribute, $\hat{A}$, $\hat{B}$ be two partitions on the universe U, and X $\subseteq$ U. $\hat{A}(x)$ and $\hat{B}(x)$ are subset of $\hat{A}, \hat{B}$ The lower approximation and the upper approximation of X in U are defined by the following:

$$\underline{X}\hat{A} + \hat{B} = \{x : \hat{A}(x) \subseteq \mathrm{X} \text{ or } \mathrm{x} : \hat{B}(x) \subseteq \mathrm{X}\} \tag{5}$$

$$\bar{X}\hat{A} + \hat{B} = \sim \underline{(\sim X)}\hat{A} + \hat{B} \tag{6}$$

Example 1 Table 2 depicts a sample decision system containing some information about a dataset. P (temperature) and Q (Cold) are the conditional attributes of the system one represents high, two represents medium and three represents low, whereas D (Fever) is the decision attribute one represents yes and zero represents no.

Let X = {e1, e2, e6, e8}. Three partitions are induced from Table 2 as follows:

$$\hat{P} = \{\{\mathrm{e1},\ \mathrm{e8}\},\ \{\mathrm{e2},\ \mathrm{e3},\ \mathrm{e4},\ \mathrm{e5},\ \mathrm{e6},\ \mathrm{e7}\}\}$$

$$\hat{Q} = \{\{\mathrm{e1},\ \mathrm{e2}\},\ \{\mathrm{e3},\ \mathrm{e4},\ \mathrm{e5},\ \mathrm{e7}\},\ \{\mathrm{e6},\ \mathrm{e8}\}\}$$

$$\widehat{P \cap Q} = \{\{e1\},\{e2\},\{e3,e4,e5,e7\},\{e6\},\{e8\}\}$$

5.4.4.1 Pessimistic multi-granulation lower approximation

The lower approximation of a target concept in complete information systems using multi-equivalence relation is defined as follows: For example, for Table 2.

$$\underline{X}\hat{P} + \hat{Q} = \{e1,e8\}\bigcap\{e1,e2,e6,e8\} = \{e1,e8\}$$

Table 2 A sample information system

U	P (Temperature)	Q (Cold)	D (Fever)
e1	1	1	1
e2	2	1	1
e3	2	2	0
e4	2	2	0
e5	2	2	0
e6	2	3	1
e7	2	2	0
e8	1	3	1

5.4.4.2 Pessimistic multi-granulation upper approximation Multi-granulation upper approximation is a complement of multi-granulation lower approximation. For example for Table 2,

$$\hat{X}\hat{P}+\hat{Q}=\{e1,e2,e3,e4,e5,e6,e7,e8\}\bigcap\{e1,e2,e3,e4,e5,e6,e8\}\}$$
$$=\{\mathrm{e1},\ \mathrm{e2},\ \mathrm{e3},\ \mathrm{e4},\ \mathrm{e5},\ \mathrm{e6},\ \mathrm{e8}\}$$

Example 2 Table 2 depicts a sample information system contains some data objects. P and Q are the conditional attributes of the system, whereas D is the decision attribute.

Let X = {e1, e2, e6, e8}. Three partitions are induced from Table 1 as follows:

$$\hat{P}=\{\{\mathrm{e1},\ \mathrm{e8}\},\{\mathrm{e2},\ \mathrm{e3},\ \mathrm{e4},\ \mathrm{e5},\ \mathrm{e6},\ \mathrm{e7}\}\}$$

$$\hat{Q}=\{\{\mathrm{e1},\ \mathrm{e2}\},\{\mathrm{e3},\ \mathrm{e4},\ \mathrm{e5},\ \mathrm{e7}\},\{\mathrm{e6},\ \mathrm{e8}\}\}$$

$$\widehat{P\bigcap Q}=\{\{e1\},\{e2\},\{e3,e4,e5\},\{e6\},\{e7\},\{e8\}\}$$

5.4.4.3 Optimistic multi-granulation lower approximation The lower approximation of a goal notion in complete sample data objects using multiple equivalence relations is defined as follows: For example, for Table 2.

$$\underline{X}\hat{P}+\hat{Q}=\{e1,e8\}\bigcup\{e1,e2,e6,e8\}=\{e1,e2,e6,e8\}$$

5.4.4.4 Optimistic multi-granulation upper approximation Multi-granulation upper approximation is a reverse of multi-granulation lower approximation. For example for Table 2,

$$\bar{X}\hat{P}+\hat{Q}=\{e1,e2,e3,e4,e5,e6,e7,e8\}\bigcap\{e3,e4,e5,e7\}\}$$
$$=\{\mathrm{e3},\ \mathrm{e4},\ \mathrm{e5},\ \mathrm{e7}\}$$

But, the lower approximation and the upper approximation of X based on Pawlak's rough set theory are as follows:

$$\underline{X}\widehat{P\bigcup Q}=\left\{Y\in\widehat{P\bigcup Q}:Y\subseteq \mathrm{X}\right\}=\{\mathrm{e1},\mathrm{e2},\mathrm{e6},\mathrm{e8}\}$$

$$\bar{X}\widehat{P\bigcup Q}=\left\{Y\in\widehat{P\bigcup Q}:Y\bigcap \mathrm{X}\neq\theta\right\}=\{e1,e2,e6,e8\}$$

Why we need multi-granulation rough set

- The basic structure of rough set is an approximation space consisting of a universe of discourse and a binary relation imposed on it. In classical approximation space, the equivalence relation is a very restrictive condition, so the application domain of rough set model is narrowed, to some extent. Thus, various extension forms of classical rough set have been introduced over the past years. In the extension of universe, since the rough set on single universe may limit the description of decision information provided by experts, two or multiple universes can describe the real-world information more effectively and reasonably. Thus, the model of rough set over two universes has been studied extensively and applied in many real-life decision making problems [56, 65, 66].
- In some data analysis issues, for the same object, there is a contradiction or inconsistent relationship between its values under one attribute set P and those under another attribute set Q. In other words, we cannot perform the intersection operations between their quotient sets and the target concept cannot be approximated by using $\mathrm{U}/P\bigcup Q$.
- In the process of some decision making, the decision or the view of each of decision makers may be independent for the same project (or a sample, object and element) in the universe. In this situation, the intersection operations between any two quotient sets will be redundant for decision making.
- Though Multi-granulation rough set has major advantages over the other rough set methods, but Multi-granulation rough set approximation space generates too many rules than rough set and that create many difficulties while taking decisions. Therefore, it is essential to minimize the decision rules.
- To extract decision rules from distributive information systems and groups of intelligent agents through using rough set approaches, knowledge representation and rough set approximations should be investigated. For the reduction of the time complexity of rule extractions, it is unnecessary for us to perform the intersection operations in between all the sites in the context of distributive information systems [19].
- Thus, ECG database also have a certain degree of uncertainty: rules are extracted from databases are also incomplete, which suggests that rule induction method should deal with uncertain rules. According to this motivation rule induction based on rough set theory have been applied to medical databases empirically, the results of which shows that rough set based methods are useful to extract diagnostic rules. In this present study how diagnostic rules modeled by the concept of multiple relation rough set theory in more theoretical way. The key ideas are MGRS methods discusses cardiac disease diagnostic rules which is closely with rough set rule model. The important point that ECG

diagnostic reasoning is characterized by focusing mechanism, composed by screening and differential diagnosis, which corresponds to the MGRS Upper and lower approximation of a target concept. Furthermore, this paper focuses on detection of complications, which can be viewed as multiple relation between rules of different diseases.

Figures 4, 5 shows relationship and difference between rough set and multi-granulation rough sets (OMGRS and PMGRS). The rough set model cannot be used to deal with the information systems with complicated context. On the other hand, by relaxing the indiscernibility relation to more general binary relations, for that reason many extension of rough set models. Such as Multi-granulation rough set models have been successfully applied into classification systems with complicated task for decision making process. In this paper, two types of the multiple granulation rough set based classification models (OMGRS and PMGRS) have been constructed, respectively, based on multiple equivalence relations for an information system. In the two types of multiple granulations rough set model, a target concept was approximated from two different kinds of views by using the equivalence classes induced from multiple granulations. The optimistic multi-granulation rough set over two universes cannot fit all decision making problems in medical diagnosis. For instance, not all decision makers are risk-preferring. Sometimes a risk-averse decision model may be more reasonable. So, we give pessimistic multi-granulation rough set model over two universes.

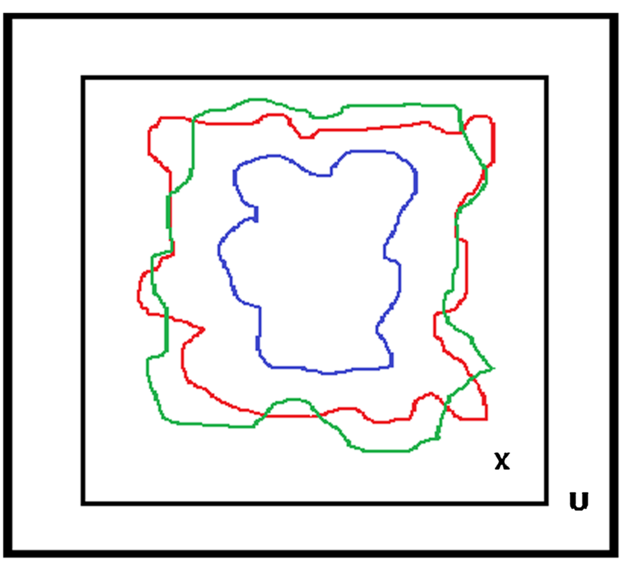


Fig. 4 Difference and relationship of lower

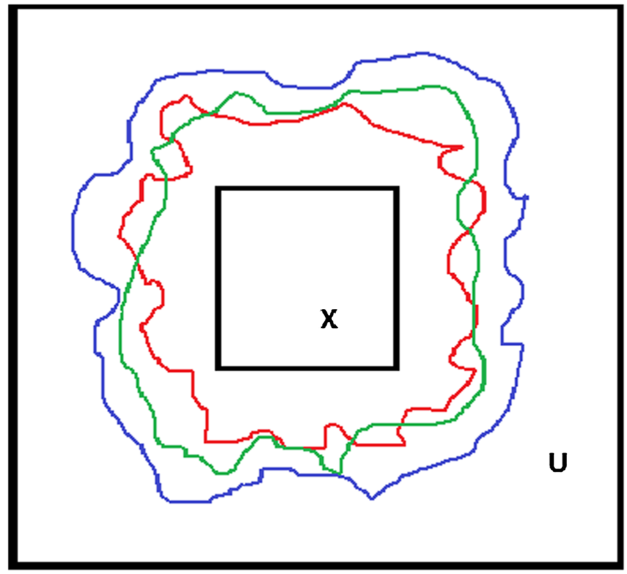


Fig. 5 Difference and relationship of Upper Approximations among the PMGRS (*blue*), the OMGRS, Approximations among the PMGRS (*blue*), the OMGRS, (*red*) and classical Rough Set (*green*). (*Red*) and classical Rough Set (*green*)

5.5 Classification

To construct a well-organized prediction system, besides discriminatory features, the choice of powerful classifier performs a key function. Consequently, several classifiers were studied. Judging a classifier generally relies upon on comparing its generalization capacity that refers to the classifier's overall performance in categorizing special cardiac arrhythmia's. As shown in Fig. 3, the Pan-Tomkins (PT) and wavelet transform (WT) approaches was utilized to extract the features of each ECG signal to consist the full feature set. Then, the discretize feature set can be achieved using the Max–Min approach. In the sequel, the Multi-granulation rough sets (PMGRS and OMGRS) classifier was exerted to classify the five-class of cardiac arrhythmia. The full frame of the proposed Multi-granulation rough set rule based classification methods (PMGRS and OMGRS) is shown in Table 3.

The decision rules generated using proposed algorithm for the example presented in Table 2 are given in Table 4.

The concept of rough-rule base is used for MIT-BIH cardiac arrhythmia to deal with the uncertainties and incompleteness as well as to gain in computation time. A Multi- granular rough rule base is formulated which proves to be effective in reducing the indiscernibility of the rule-base. This new rule-base provides more accurate results in the MIT-BIH cardiac arrhythmia classification. Three indices to evaluate the performance of classification are defined.

Table 3 Proposed Multi-Granulation Rough set rule based Classification (PMGRS and OMGRS) algorithms

Proposed Algorithm: Pessimistic Multi-granulation Rough set (PMGRS) based classification [11, 34]	Proposed Algorithm: Optimistic Multi-granulation Rough set (OMGRS) based classification [11, 34]
Input: Given Dataset with conditional attributes 1, 2, …, n-1 and the Decision attribute n	*Input:* Given Dataset with conditional attributes 1, 2, …, n-1 and the Decision attribute n
Output: Generated Decision Rules	*Output:* Generated Decision Rules
Step 1: Construct the Pessimistic multi-granulation rough set based lower approximation for the given data set	Step 1: Construct the Optimistic multi-granulation rough set based lower approximation for the given data set
$\underline{X}\hat{P} + \hat{Q} = \{x : \hat{P}(x) \subseteq X \text{ and } x : \hat{Q}(x) \subseteq X\}$ (11)	$\underline{X}\hat{A} + \hat{B} = \{x : \hat{A}(x) \subseteq X \text{ and } x : \hat{B}(x) \subseteq X\}$ (13)
Step 2: Construct the Pessimistic Multi-granulation rough set based upper approximation for the given dataset	Step 2: Construct the Optimistic Multi-granulation rough set based upper approximation for the given dataset
$\bar{X}\hat{P} + \hat{Q} = \sim \underline{(\sim X)}\hat{P} + \hat{Q}$ (12)	$\bar{X}\hat{A} + \hat{B} = \sim \underline{(\sim X)}\hat{A} + \hat{B}$ (14)
Step 3: Generate the certain rules using Pessimistic Multi-granulation rough set based lower approximation	Step 3: Generate the certain rules using Optimistic Multi-granulation rough set based lower approximation
Step 4: Generate the possible rules using Pessimistic Multi-granulation rough set based upper approximation	Step 4: Generate the possible rules using Optimistic Multi-granulation rough set based upper approximation
Step 5: Compute the validate measure value for each non-deterministic rule	Step 5: Compute the validate measure value for each non-deterministic rule

6 Experimental analysis

The main objective of the assessment was to compare the Multi-granulation rough set methods to the other bench mark methods, for the purpose of classifying ECG signal data sets. This is the basis of the cardiac signal examination, which is commonly applied to 24 ECG recordings to detect the cardiac arrhythmia diseases. In order to evaluate our proposed system, random selection from the mixture of data of these five classes (NSR, LBBB, RBBB, PVC and PR) may result in biased training and testing sets, i.e., the sample number of five classes are unbalanced. To deal this problem, for every trial, we do the random partition for all five classes separately. Multi-granulation approximation space of the trials in all five classes will go into the training set, while the remaining trials will be used for testing set [35]. The averaged output values (precision, recall, and F-measure) are illustrious. The evaluation of testing performance of the proposed PMGRS and OMGRS was performed by the assessment of classification outcomes. In the classification, the aim is to assign the input patterns to one of five classes, usually represented by outputs restricted to lie in the range from 0 to 1, so that they represent the probability of class. Rendering to the multiple signed ECG signal, the proposed Multi-granulation rough set methods (PMGRS and OMGRS) is comparable with the existing best methods, Multi-granulation rough set methods (PMGRS and OMGRS) and significantly better than JRip, J48, Naïve bayes, MLP and Decision table. Proposed classification trained based on test at results with five distinctive classes have been evolved with the aid of a confusion matrix. The confusion matrices displaying the category effects of the PMGRS and OMGRS are given in Table 5. From those matrices you'll be able to inform the frequency with which an ECG beat is misclassified as every other. The test overall performance of the PMGRS and OMGRS might be determined with the aid of the computation of precision, recall and F-measure [8, 12, 13]. The precision, recall and F-measure are defined as:

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}}$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}$$

$$\text{F} - \text{Measure (Czekanowski} - \text{Dice index)} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

A true negative result occurs when both the classifier and the physician recommended the time off of a positive finding. A true positive result occurs when the positive recognition of the classifier concurred with a positive finding of the general practitioner. In order to define the performance of the PMGRS and OMGRS, the classification accuracies (precision, recall, F-Measure) on the test sets are presented in Table 5.

6.1 Discussion

To evaluate the prognosis efficiency of the proposed technique, ECG cardiac arrhythmia datasets had been downloaded from the MIT-BIH database [48]. Those data sets were frequently used as benchmarks to evaluate the performance of various categorization techniques in the

Table 4 Example for proposed works

Proposed Algorithm: Pessimistic Multi-granulation Rough set (PMGRS) based classification [11, 34]	Proposed Algorithm: Optimistic Multi-granulation Rough set (OMGRS) based classification [11, 34]
Input: Given Dataset with conditional attributes P,Q and the Decision attribute n	*Input:* Given Dataset with conditional attributes P,Q and the Decision attribute n
Output: Generated Decision Rules	*Output:* Generated Decision Rules
Step 1: Construct the Pessimistic multi-granulation rough set based lower approximation for the given data set $\underline{X}\hat{P} + \hat{Q} = \{e1, e8\} \bigcap \{e1, e2, e6, e8\} = \{e1, e8\}$	Step 1: Construct the Optimistic multi-granulation rough set based lower approximation for the given data set $\underline{X}\hat{P} + \hat{Q} = \{e1, e8\} \bigcup \{e1, e2, e6, e8\} = \{e1, e2, e6, e8\}$
Step 2: Construct the Pessimistic Multi-granulation rough set based upper approximation for the given dataset $\bar{X}\hat{P} + \hat{Q} = \{e1, e2, e3, e4, e5, e6, e7, e8\} \bigcap \{e1, e2, e3, e4, e5, e6, e8\}\} =$ {e1, e2, e3, e4, e5, e6, e8}	Step 2: Construct the Optimistic Multi-granulation rough set based upper approximation for the given dataset $\bar{X}\hat{P} + \hat{Q} =$ $\{e1, e2, e3, e4, e5, e6, e7, e8\} \bigcap \{e3, e4, e5, e7\}\}$
Step 3: Generate the certain rules using Pessimistic Multi-granulation rough set based lower approximation	= {e3, e4, e5, e7}
If P = 1 and Q = 1 ≥ D=1	Step 3: Generate the certain rules using Optimistic Multi-granulation rough set based lower approximation
If P = 1 and Q = 3 ≥ D=1	If P = 1 and Q = 1 ≥ D=1
Step 4: Generate the possible rules using Pessimistic Multi-granulation rough set based upper approximation	If P = 2 and Q = 1 ≥ D=1
If P = 1 and Q = 1 ≥ D=1	If P = 2 and Q = 3 ≥ D=1
If P = 2 and Q = 1 ≥ D=1	If P = 1 and Q = 3 ≥ D=1
If P = 2 and Q = 2 ≥ D=0	Step 4: Generate the possible rules using Optimistic Multi-granulation rough set based upper approximation
If P = 2 and Q = 2 =>D = 0	If P = 2 and Q = 2 =>D = 0
If P = 2 and Q = 2 =>D = 0	If P = 2 and Q = 2 =>D = 0
If P = 2 and Q = 3 =>D = 1	If P = 2 and Q = 2 =>D = 0
If P = 1 and Q = 3 =>D = 1	If P = 3 and Q = 2 =>D = 1
Step 5: Compute the validate measure value for each non-deterministic rule	Step 5: Compute the validate measure value for each non-deterministic rule

Table 5 Results for the MIT-BIH cardiac arrhythmia data set and detailed performance comparison of proposed method with other classifiers

Algorithms	Classes	Precision		Recall		F-Measure	
		PT	WT	PT	WT	PT	WT
PMGRS	NSR	0.961	0.975	1	0.998	0.986	0.989
	LBBB	1	1	0.989	0.991	0.993	0.997
	PVC	0.995	0.993	0.984	0.988	0.984	0.989
	RBBB	0.982	0.987	0.999	1	1	0.999
	PR	0.986	1	0.994	0.994	0.989	0.991
	Over all	0.9858	0.991	0.9932	0.9942	0.9904	0.993
OMGRS	NSR	0.976	0.976	0.945	0.943	0.962	0.974
	LBBB	0.993	1	0.982	0.987	0.974	0.981
	PVC	0.962	0.972	0.973	0.971	0.966	0.964
	RBBB	0.974	0.978	0.985	0.983	0.987	0.985
	PR	0.968	0.967	0.984	0.975	0.975	0.974
	Over all	0.9746	0.9786	0.9738	0.9718	0.9728	0.9756
Decision table	NSR	0.67	0.696	0.968	0.972	0.792	0.815
	LBBB	1	0.976	0.64	0.699	0.78	0.806
	PVC	0.927	0.943	0.685	0.734	0.788	0.808
	RBBB	0.977	0.935	0.697	0.689	0.813	0.795
	PR	0.833	0.857	0.638	0.658	0.723	0.701
	Over all	0.8814	0.8814	0.7256	0.7504	0.7792	0.785
MLP	NSR	0.756	0.854	0.836	0.881	0.794	0.848
	LBBB	0.741	0.798	0.8	0.853	0.769	0.798
	PVC	0.771	0.823	0.757	0.796	0.764	0.776
	RBBB	0.821	0.868	0.754	0.795	0.786	0.831
	PR	0.784	0.827	0.617	0.664	0.69	0.753
	Over all	0.7746	0.834	0.7528	0.7978	0.7606	0.8012
Naïve Bayes	NSR	0.491	0.552	0.698	0.736	0.576	0.617
	LBBB	0.742	0.767	0.92	0.943	0.821	0.835
	PVC	0.483	0.521	0.387	0.428	0.43	0.446
	RBBB	0.742	0.766	0.92	0.930	0.821	0.837
	PR	0.432	0.449	0.404	0.456	0.418	0.434
	Over all	0.578	0.611	0.6658	0.6986	0.6132	0.6338
Multiclass classifier	NSR	0.517	0.558	0.868	0.893	0.648	0.668
	LBBB	0.85	0.895	0.68	0.716	0.756	0.77
	PVC	0.514	0.548	0.171	0.222	0.257	0.289
	RBBB	0.515	0.546	0.41	0.432	0.457	0.486
	PR	0.826	0.851	0.404	0.446	0.543	0.557
	Over all	0.6444	0.6796	0.5066	0.5418	0.5322	0.554
JRip	NSR	0.814	0.821	0.952	0.943	0.878	0.864
	LBBB	0.885	0.866	0.92	0.923	0.902	0.90
	PVC	0.901	0.910	0.82	0.811	0.858	0.838
	RBBB	0.963	0.954	0.844	0.835	0.9	0.895
	PR	0.846	0.832	0.702	0.7	0.767	0.759
	Over all	0.8818	0.8766	0.8476	0.8424	0.861	0.8512
J48	NSR	0.896	0.925	0.91	0.948	0.903	0.932
	LBBB	0.889	0.917	0.96	1	0.923	0.957
	PVC	0.843	0.896	0.874	0.906	0.858	0.894
	RBBB	0.934	0.964	0.926	0.952	0.93	0.960
	PR	1	0.987	0.83	0.885	0.907	0.924
	Over all	0.9124	0.9378	0.9	0.9382	0.9042	0.9334

literature. A feature extraction and discretization techniques (Pan–Tomkins's, Wavelet transform and Max–Min discretization approaches) are applied for ECG signals. This techniques are used as pre-processing technique for the analysis of the ECG signal. Concerning relevant statistical evaluation methods precision, recall and F-Measure, we can say that a number of mentioned algorithms outcome depend on the preprocessing stage. Because the same classifier may have various performances on different applications, its miles cautioned to investigate the wishes of the gadget before deciding on a proper classifier. In this ECG data classification system, every patient is represented as a characteristic attribute of morphological measures in cardiac rhythm [35]. At the same time as Pan-Tomkins's primarily based morphological capabilities offer powerful discrimination capability among ordinary and a few bizarre heartbeats morphological features make a contribution to the discriminating power of signal peek points (P, Q, R, S, and T). Wavelet transform since they comprehensively analyze the center as well as excessive frequency indicators, which boom the accuracy within the time–frequency signals. Throughout the development of rough set model, only the most idealist rule generation method, and the speed of the researching is very fast. The calculation in rough set based rule generation methods is very simple. Compared with the other methods, it occupies the bigger reliability and it can be completed easily. The classification accuracy of the methods is given in Table 5. The Table 5 is the number of Multi-granulation rough set methods (PMGRS and OMGRS) wins with higher classification accuracy than other methods (JRip, J48, Naïve bayes, MLP, Multiclass classifier and Decision table). The high classification accuracy of Multi-granulation Rough set methods (PMGRS and OMGRS) reflects that and maintains the vital information of the records inside the classification system. The general consistency among our consequences and the consistency with the conventional assessment give a strong indication that our algorithm is capable system of extract reliable ECG signals. Serving as extensively used statistical measures for multiclass classification. In this paper, choice is described as advantageous if the affected person outcome is cardiac arrhythmia, while bad case refers to no longer regarded. The high overlap in results is noteworthy rather because it shows that our algorithm performs well for as much as ECG signal data. Need to mention, Pessimistic Multi-granulation rough set methods (PMGRS) achieves the highest classification accuracy than Optimistic Multi-granulation rough set methods (OMGRS), while the other classification methods performance is poor. PMGRS is a better diagnostic tool than OMGRS. This shows Multi-granulation rough set methods (PMGRS and OMGRS) have the ability to solve the difficult problems. It is a decent belongings in the real-world applications.

7 Conclusion

In this study, Multi-granulation rough set rule based approaches (PMGRS and OMGRS) are proposed a new methods for ECG cardiac arrhythmia categorization. The Pan-Tomkins's and wavelet transform methods have been decided on to gain a compact set of features. This combination of features captures all the statistical and peak point aspects of beats to classify ECG signal in five different type of classes. In the experiment, classification was performed by proposed Multi-granulation rough set rule based approaches (PMGRS and OMGRS). These experimental results showed that the classification accuracy of Multi-granulation rough set rule based approach (PMGRS and OMGRS) for ECG cardiac arrhythmia classification is higher to those of other classification methods, whose objects were randomly decided on. The very best accuracy acquired with the aid of proposed method. Finally the classification undertaking achieved high-quality ever for all cardiac arrhythmia rhythms in the MIT-BIH database. Generally, the results indicate that Multi-granulation rough set rule based approaches (PMGRS and OMGRS) can be used for various classification process. The final experiment demonstrates the power of proposed methods over other algorithms (JRip, J48, Naïve bayes, MLP, multiclass classifier and Decision table).

Acknowledgments The second author would like to thank UGC (University Grant Commission), New Delhi, India for the financial support received under UGC Major Research Project No. F-41-650/2012 (SR). This article does not contain any studies with human participants or animals performed by any of the authors. The data used in this research is the ECG signals from the MIT–BIH (Massachusetts Institute of Technology—Boston's Beth Israel Hospital) Arrhythmia database available on physionet website (http://www.physionet.org/physiobank/database).

References

1. Korürek M, Doğan B (2010) ECG beat classification using particle swarm optimization and radial basis function neural network. Expert Syst Appl 37(12):7563–7569
2. Pasolli E, Melgani F (2010) Active learning methods for electrocardiographic signal classification. IEEE Trans Inf Technol Biomed 14(6):1405–1416
3. Melgani F, Bazi Y (2008) Classification of electrocardiogram signals with support vector machines and swarm particle optimization. IEEE Trans Inf Technol Biomed 12(5):667–677
4. Ubeyli ED (2009) Combining recurrent neural networks with eigenvector methods for classification of ECG beats. Digit. Signal Process. 19(2):320–329
5. Kohler BU, Henning C, Orglmeister R (2002) The principles of software QRS detection. IEEE Eng Med Biol 21(1):42–57
6. Christov I, Gómez-Herrero G, Krasteva V, Jekova I, Gotchev A, Egiazarian K (2006) Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. Med Eng Phys 28(9):876–887

7. Kania Michał, Rix Hervé, Fereniec Małgorzata, Zavala-Fernandez Heriberto, Janusek Dariusz, Mroczka Tomasz, Stix Günter, Maniewski Roman (2014) The effect of precordial lead displacement on ECG morphology. Med Biol Eng Comput 52:109–119
8. Homaeinezhad MR, Sabetian P, Feizollahi A, Ghaffari A, Rahmani R (2012) Parametric modelling of cardiac system multiple measurement signals: an open-source computer framework for performance evaluation of ECG, PCG and ABP event detectors. J Med Eng Technol 36:117–134
9. Kumar S, Senthil H Hannah, Inbarani (2015) Modified soft rough set based ecg signal classification for cardiac arrhythmias big data in complex systems. Springer International Publishing, Berlin, pp 445–470
10. Senthilkumar S, Hannah Inbarani H, Udhayakumar S (2014) Modified soft rough set for multiclass classification. Adv Intell Syst Comput 246:379–384
11. Senthil Kumar S, Hannah Inbarani H (2015) Optimistic multi-granulation rough set based classification for medical diagnosis. Procedia Comput Sci 47:374–382
12. Udhaya Kumar S, Hannah Inbarani H, Senthilkumar S (2013) Bijective soft set based classification of medical data, pattern recognition, informatics and medical engineering (PRIME), international conference, pp 517–521
13. Inbarani HH, Kumar SS, Azar AT, Hassanien AE (2014) Soft rough sets for heart valve disease diagnosis. In: Advanced conference on advanced machine learning technologies and applications. Proceedings 2nd international conference, AMLTA 2014, Cairo, Egypt, Nov 28–30 2014. Springer, pp 347–356. doi:10.1007/978-3-319-13461-1_33
14. Acampora Giovanni, Lee Chang-Shing, Wang Autilia Vitiello Mei-Hui (2012) Evaluating cardiac health through semantic soft computing techniques. Soft Comput 16:1165–1181
15. Pan J, Tompkins W (1985) A real-time QRS detection algorithm. IEEE Trans Biomed Eng 3:230–236
16. Moody GB, Mark RG (2001) The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 20(3):45–50
17. Xu W, Zhang X (2013) Zhang W (2013) Two new types of multiple granulation rough set. ISRN Appl Mathemat 2013:16. doi:10.1155/2013/791356
18. Xu W, Xiantao Zhang, Qiaorong Wang (2011) A generalized multi-granulation rough set approach. International Conference on Intelligent Computing. Springer, Berlin Heidelberg
19. Qian Y, Liang J, Yao Y, Dang C (2010) MGRS: a multi-granulation rough set. Inf Sci 180:949–970
20. Pawlak Zdzisław (1982) Rough sets. Int J Comput Inform Sci 11(5):341–356
21. Raghavan R (2013) Validation over basic set operations of internal structure of multi granular rough sets. Int J Latest Res Eng Comput (IJLREC) 1:34–42
22. Raghavan R, Tripathy BK (2011) On some topological properties of multi granular rough sets. Adv Appl Sci Res 2:536–543
23. Raghavan R, Tripathy BK (2013) On some comparison properties of rough sets based on multi granulations and types of multi granular approximations of classifications. Int J Intell Syst Appl 06:70–77
24. Karpagachelvi S, Arthanari M, Sivakumar M (2012) Classification of electrocardiogram signals with support vector machines and extreme learning machine. Neural Comput Appl 21:1331–1339
25. Mert A, Kılıc N, Akan A (2014) Evaluation of bagging ensemble method with time-domain feature extraction for diagnosing of arrhythmia beats. Neural Comput Appl 24:317–326
26. Kumari VSR, Rajesh Kumar P (2013) Fuzzy Unordered Rule Induction for Evaluating Cardiac Arrhythmia. Biomed Eng Lett 3:74–79
27. Ali Khazaee AE, Zadeh (2014) ECG beat classification using particle swarm optimization and support vector machine. Front Comput Sci. 8(2):217–231
28. Yuhua Qian A, Zhangb Hu, Sangb Yanli, Liang Jiye (2014) Multi granulation decision-theoretic rough sets. Int J Approximate Reasoning 55:225–237
29. Jiye Liang, Yuhua Qian, Chengyuan Chu, Deyu Li, Junhong Wang, Rough Set (2005) Approximation Based on Dynamic Granulation. In: Rough sets, fuzzy sets, data mining, and granular computing, lecture notes in computer science, 3641:701–708
30. Qibin Zhao, Liqing Zhang (2005) ECG feature extraction and classification using wavelet transform and support vector machines, international conference on neural networks and Brain, 2005. ICNN&B '05, 2:1089–1092
31. Rai HM, Trivedi A, Shukla S (2013) ECG signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier. Measurement 46(9):3238–3246
32. Sathishkumar EN, Thangavel K, Nishama A (2014) Comparative analysis of discretization methods for gene selection of breast cancer gene expression data. In: Computational intelligence, cyber security and computational models. Springer India, pp 373–378
33. Tripathy BK, Panda GK, Mitra A (2012) Incomplete multi granulation based on rough intuitionistic fuzzy sets. UNIASCIT 2(1):118–124
34. Kumar SS, Inbarani HH, Azar AT, Own HS, Balas VE, Olariu T (2016) Optimistic multi-granulation rough set-based classification for neonatal jaundice diagnosis. In: Soft computing applications, vol. 356. Springer, pp 307–317
35. Liu N, Lin Z, Koh Z, Huang GB, Ser W, Ong MEH (2011) Patient outcome prediction with heart rate variability and vital signs. J Signal Proc Syst 64(2):265–278
36. Jadhav Shivajirao, Nalbalwar Sanjay, Ghatol Ashok (2014) Feature elimination based random subspace ensembles learning for ECG arrhythmia diagnosis. Int J Soft Comput 18(3):579–587
37. Tong-Jun Li, Rough sets and general basic set assignments, lecture notes in computer science, 2011
38. Karaye IA., Saminu S, Özkurt N (2014) Analysis of cardiac beats using higher order spectra. In: IEEE 6th International conference on adaptive science & technology (ICAST), 29–31 Oct 2014, pp 1–8. doi:10.1109/ICASTECH.2014.7068145
39. Badiezadegan S, Soltanian-Zadeh H (2008) Design and evaluation of matched wavelets with maximum coding gain and minimum approximation error criteria for r peak detection in ECG. Int J Wavelets Multiresolut Inf Process 6(6):799–825
40. Wieben O, Afonso VX, Tompkins WJ (1999) Classification of premature ventricular complexes using filter bank features, induction of decision trees and a fuzzy rule-based system. Med Biol Eng Comput 37(5):560–565
41. Gogoi P, Bhattacharyya DK, Kalita JK (2013) A rough set–based effective rule generation method for classification with an application in intrusion detection. Int J Secur Netw 8(2):61–71
42. Engin M, Fedakar M, Engin EZ, Korürek M (2007) Feature measurements of ECG beats based on statistical classifiers. Measurement 40(9):904–912
43. Li X, Shu L, Hu H (2009) Kernel-based nonlinear dimensionality reduction for electrocardiogram recognition. Neural Comput Appl 18(8):1013–1020
44. Moskovitch R, Shahar Y (2015) Classification-driven temporal discretization of multivariate time series. Data Min Knowl Discov 29(4):871–913
45. Alajlan N, Bazi Y, Melgani F, Malek S, Bencherif MA (2014) Detection of premature ventricular contraction arrhythmias in electrocardiogram signals with kernel methods. Signal Image Video Process 8(5):931–942

46. Hassanien AE, Abraham A, Peters JF, Schaefer G (2009) Rough sets in medical informatics applications. In: Applications of soft computing. Springer Berlin Heidelberg, pp 23–30
47. Yang P, Li Q (2014) Wavelet transform-based feature extraction for ultrasonic flaw signal classification. Neural Comput Appl 24(3–4):817–826
48. El-Dahshan ESA (2011) Genetic algorithm and wavelet hybrid scheme for ECG signal denoising. Telecommun Syst 46(3):209–215
49. Qian YH, Liang JY, Dang CY (2010) Incomplete multi granulation rough set. IEEE Trans System Man Cy A 20:420–431
50. Yao YY (2001) Information Granulation and Rough Set Approximation. Int J Intell Syst 16(1):87–104
51. Pawlak Z, Skowron A (2007) Rudiments of rough sets. Inf Sci 177(1):3–27
52. Swiniarski RW, Skowron A (2003) Rough set methods in feature selection and recognition. Pattern Recogn Lett 24(6):833–849
53. Polkowski L, Skowron A (1998a). Rough sets in knowledge discovery, Vol. 1/2. Studies in Fuzziness and Soft Computing series, Physica–Verlag
54. Polkowski L, Skowron A (1998b). Rough sets and current trends in computing, LNAI 1424, Springer
55. Zadeh LA (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 90(2):111–127
56. Pei D, Xu Z-B (2004) Rough set models on two universes. Int J Gen Syst 33(5):569–581
57. Wang Xizhao, Zhexue Huang Joshua (2015) Editorial uncertainty in learning from big data. Fuzzy Sets Syst 258(1):1–4
58. Yao Yiyu, She Yanhong (2016) Rough set models in multi granulation spaces. Inf Sci 327:40–56
59. Luo Chuan, Li Tianrui, Yi Zhang, Fujita Hamido (2016) Matrix approach to decision-theoretic rough sets for evolving data. Knowl Based Syst 99(1):123–134
60. Zhang Junbo, Zhu Yun, Pan Yi, Li Tianrui (2016) Efficient parallel boolean matrix based algorithms for computing composite rough set approximations. Inf Sci 329(1):287–302
61. Li Shaoyong, Li Tianrui, Jie Hu (2015) Update of approximations in composite information systems. Knowl Based Syst 83:138–148
62. Chen Hongmei, Li Tianrui, Luo Chuan, Horng Shi-Jinn, Wang Guoyin (2015) A decision-theoretic rough set approach for dynamic data mining. IEEE Trans Fuzzy Syst 23(6):1958–1970
63. Chen Hongmei, Li Tianrui, Luo Chuan, Horng Shi-Jinn, Wang Guoyin (2014) A rough set-based method for updating decision rules on attribute values coarsening and refining. IEEE Trans Knowl Data Eng 26(12):2886–2899
64. Weihua Xu, Guo Yanting (2016) Generalized multi granulation double-quantitative decision-theoretic rough set. Knowl Based Syst 105(1):190–205
65. Yang Hai-Long, Liao Xiuwu, Wang Shouyang, Wang Jue (2013) Fuzzy probabilistic rough set model on two universes and its applications. Int J Approx Reason 54(9):1410–1420
66. Jie Hu, Li Tianrui, Chen Hongmei, Zeng Anping (2015) An incremental learning approach for updating approximations in rough set model over dual universes. Int J Intell Syst 30(8):923–947
67. Yiyu Yao (2007) Decision-theoretic rough set models, Rough Sets and Knowledge Technology, Second International Conference, RSKT 2007, Proceedings, LNCS(LNAI) 4481, pp. 1–12
68. Zhang Junbo, Li Tianrui, Chen Hongmei (2014) Composite rough sets for dynamic data mining. Inf Sci 257:81–100
69. Kumar SU, Inbarani HH (2016) Neighborhood rough set based ECG signal classification for diagnosis of cardiac diseases. Soft Comput 20:1–13