Motor Imagery Classification Based on Variable Precision Multigranulation Rough Set and Game Theoretic Rough Set

K. Renuga Devi and H. Hannah Inbarani

Abstract In this work classification of motor imagery BCI based on Variable Precision Multigranulation Rough Set and Game theoretic Rough Set is proposed. The efficient classification of motor imagery movements of patients can lead to accurate design of Brain Computer Interface (BCI). Data set are collected from BCI Competition III dataset 3a and BCI competition IV data set I. During acquisition there are several noises that affect classification of Electroencephalogram (EEG) Signal, so pre-processing is carried out with Chebyshev type1 filter between 4–40 Hz in order to remove the noises that may exist in signal. The Daubechies wavelet is used for extraction of features from EEG Signal. Variable Precision Multigranulation Rough Set is applied for classification of EEG Signal. Game theoretic Rough set is applied to determine best combination of α and β are based on accuracy of Variable Precision Multigranulation Rough Set. An experimental result depicts higher accuracy with Variable Precision Multigranulation Rough Set and Game Theoretic rough set compared to existing technique.

Keywords Chebyshev type 1 filter · Daubechies wavelet · Variable precision multigranulation rough set · Game theoretic rough set

1 Introduction

A Brain-Computer Interface (BCI) provides a functional interaction between the human brain and the external device [1]. A Brain-Computer Interface (BCI) helps to move artificial hand, leg, wheel chair based on motor imagery brain signals of subject [2–4]. The subject sends the signal through BCI can be considered as being the only way of communication for people affected by motor disabilities [5–7].

K.R. Devi (✉) · H.H. Inbarani (✉)
Department of Computer Science, Periyar University, Salem, India
e-mail: renuga.star@yahoo.co.in

H.H. Inbarani
e-mail: hhinba@gmail.com

N. Dey et al. (eds.), *Medical Imaging in Clinical Applications*,
Studies in Computational Intelligence 651, DOI 10.1007/978-3-319-33793-7_7

Motor imagery involves imagination of various body parts which modulates sensorimotor oscillation resulting in sensorimotor cortex activation in Electroencephalogram (EEG) [8]. The frequency of the sensorimotor rhythm is in the range of 13–15 Hz. Event Related Desynchronization (ERD) is a temporary decrease in power of the mu (8–13 Hz) and beta (13–50 Hz) brain waves [1, 6, 7]. When a subject performs or imagines movement, ERD can be recorded using Electroencephalography. Event Related Synchronization (ERS) which is an increase in power in the mu and beta bands that occurs when the subject stops imagining a movement [1, 9–12]. Contralateral hemisphere is the hemisphere on the side of the body opposite to the limb for which the motor imagery task is executed, whereas ipsilateral hemisphere on the same side of the body [1, 13, 14]. If the subject imagines a movement with his or her right hand, event-related desynchronization and synchronization occurs mostly in the left hemisphere. Similarly, if the subject imagines a movement with his or her left hand, the ERD/ERS occurs mostly in the right hemisphere. Larger areas of motor cortex are affected during Feet and arms movements so it is difficult to separate them [1, 15, 16]. The application of BCI are helping paralyzed [4], video games and virtual reality [8], creative expression [9, 17], neural prosthetics [18], wheelchairs [18], access to the internet [18] etc.

An efficient algorithm for classifying different user commands is an important part of a brain-computer interface. The goal of this paper is to classify different motor imagery tasks, left hand, right hand, both feet, or tongue. Recently, rough set theory is a technique was used for the data reduction data reduction, discovery of data dependencies, rule induction from databases and approximate set classification [19–21]. Rough-set data analysis avoids prior model assumptions such as probabilistic distribution, membership function used in fuzzy sets theory, and basic probability assignment in Dempster–Shafer theory of evidence and uses only internal knowledge, and does not depend on external parameters [19]. Rough Set can result in information loss, extensions of Rough set such as Variable Precision Rough set, Multigranulation Rough set, Game Theoretic Rough Set can handle real-valued domains [20]. Variable precision multigranulation rough set (VPMGRS) is an extension to Rough set with flexible classification of uncertain objects needs only slight modifications of the original Variable Precision Rough Set (VPRS) model [22]. Variable precision multigranulation rough set (VPMGRS) is applied to Motor imagery dataset collected from BCI Competition website. In VPMGRS thresholds values α, β are determined with game theoretic rough set. Classification accuracy is evaluated and compared with existing techniques. Section 2 discusses related work, Sect. 3 discusses the preliminaries, Sect. 4 depicts proposed algorithm, Sect. 5 depicts results and discussion and Sect. 6 concludes the proposed work.

2 Related Work

Duan et al. [16] presents an approach to classify BCI Competition 2003 dataset Ia. To eliminate redundancy and extract high-dimensional EEG signals Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used. LDA after PDA shows an accuracy of 88.89.

Velásquez-Martínez et al. [15] applied Common Spatial Pattern for preprocessing and eigen decomposition method to identify main features to discriminate EEG Signals in Motor Imagery dataset. The proposed method obtained average accuracy of about 95.21 ± 4.21.

Nicolas-Alonso et al. [7] proposed adaptive classification framework for BCI Competition IV dataset 2a to address non-stationary in EEG classification. It is based on Filter Bank Common Spatial Pattern and comprises following stages. Multiple bandpass filtering using Finite Impulse response filters, spatial filtering using common spatial pattern algorithm. The results yields a significantly higher mean kappa of 0.62 compared to 0.58 from the baseline probabilistic generative model without adaptive processing.

Ang et al. [23] proposed three approaches of multi-class extension to the FBCSP algorithm on Dataset 2a, namely, Divide-and-Conquer (DC), Pair-Wise (PW), and One Versus-Rest (OVR).Feature selection algorithms the Mutual Information-based Best Individual Feature (MIBIF) and Mutual Information-based Rough Set Reduction (MIRSR) are used to select discriminative CSP features. Filter Bank Common Spatial Pattern (FBCSP) algorithm is used to optimize the subject-specific frequency band for CSP on Datasets 2a and 2b of the Brain-Computer Interface (BCI) Competition IV. The FBCSP algorithm yielded a 10×10-fold cross-validation classification accuracy of 90.3.

Rodríguez-Bermúdez et al. [17] proposed fast adaptive BCI system for feature extraction and classification of EEG Signal. Power spectral density, Hjorth parameters and autoregressive modeling are used for feature extraction. The most relevant features for linear discrimination are selected using a fast and robust wrapper methodology. The proposed method is evaluated using EEG signals from nine subjects during motor imagery tasks and experimental results show its advantages over the state-of-the-art methods, especially in terms of classification accuracy and computational cost.

Zhou et al. [24] applied improved support vector machine for classifying Graz BCI Competition 2003 dataset. EEG signals with Daubechies order 4 (db4) wavelets in 10 and 21 Hz at C3 channel, and in 10 and 20 Hz at C4 channel, for these frequencies are prominent in discrimination of left and right motor imagery tasks according to EEG frequency spectral. Classification error rate of the presented approach was as low as 9.29 %.

3 Preliminaries

The Proposed Methodology involves acquisition of EEG Signal from BCI Competition III website http://www.bbci.de/competition/iii/ and BCI Competition IV website http://www.bbci.de/competition/iv/. The acquired EEG Signal has noises with in it and needs to be preprocessed before proceeding to feature extraction and classification. Features are extracted with Daubechies wavelet, since it contains both time and frequency information in it. Extracted features are classified with variable precision multigranulation rough set and compared with existing approaches. Figure 1 depicts proposed methodology of motor imagery classification.

3.1 Daubechies Wavelet

Discrete Wavelet Transform (DWT) invented by Mallat in 1998 shows that it can be viewed as a multi-resolution decomposition of signal [25]. DWT provides sufficient information both for analysis and synthesis of signal with significant reduction in computation time. While Continuous Wavelet Transform (CWT) wavelet series is a sampled version of CWT provides highly redundant information DWT decomposes the signal into its components in different frequency bands [26]. A wavelet transform is multi-resolution analysis as it gives localization in both space and frequency domains [27]. DWT using filter bank decomposes input signal into high and low frequency component. It decomposes signal into several frequency band. Wavelet transformation involves convolution of ψ (t) with signal x (t) mother wavelet function [28].

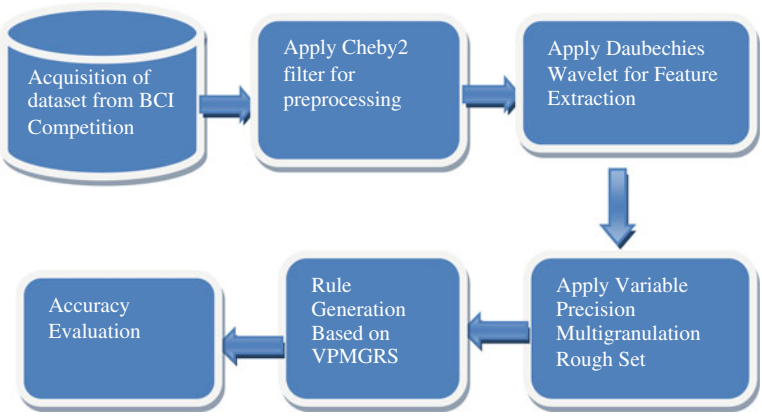

Fig. 1 Proposed methodology of motor imagery classification

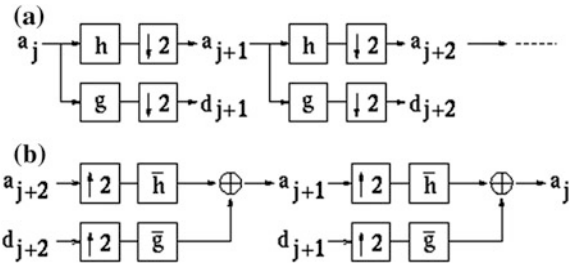

Fig. 2 **a** Decomposition. **b** Reconstruction

$$\gamma(s,\tau) = \int f(t)\Psi^{*}_{s,t}(t)dt \tag{1}$$

$$f(t) = \iint \gamma(s,\tau)\Psi_{s,t}(t)d\tau ds \tag{2}$$

$$\Psi_{s,t}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-\tau}{s}\right) \tag{3}$$

where, $\Psi_{s,t}(t)$-wavelet with scale s and time t [29], τ-Shift in time, s-change in scale, $\frac{1}{\sqrt{s}}$ is used for normalization. The original signal approximation at scale index is the combination of approximation and detail signal at the next lower scale [29, 30]. Figure 2 shows decomposition and reconstruction of signal

$$a_{j+1}(p) = \sum_{n} h(n-2p)a_j(n) \tag{4}$$

$$d_{j+1}(p) = \sum_{n} g(n-2p)a_j(n) \tag{5}$$

where the set of numbers represents $a_j(n)$ the approximation of the signal at the resolution 2^{-j} and the set of numbers d_j (n) represents the details in approximating the signal at resolution 2^{-j-1}. This is referred to as multi-resolution analysis of a signal using wavelet transform

$$\phi(x/2) = 2^{1/2}\sum_{n} h(n)\phi(x-n) \tag{6}$$

$$\Psi(x/2) = 2^{1/2}\sum_{n} g(n)\phi(x-n) \tag{7}$$

Discrete Wavelet Transform is used in this work, to decompose the signal through low-pass and high-pass filtering into low and high frequency proportion of the signal [17, 24, 31]. In real life problem, DWT is more suitable in area of biomedical applications. DWT determination examines the signal with different resolutions at

different frequency bands by decomposing the signal into detailed coefficients (cD) and Approximation coefficients (cA) [32]. In short, wavelet transform analysis of time domain at high frequency and frequency domain at low frequency [30]. Wavelet algorithm provides a way of representing a time frequency information [31]. It also transforms the Electroencephalograph (EEG) signal to allow further extraction and classification because EEG signals are non-stationary [17].

An Electroencephalograph signal can be classified into orthogonal, symmetry, compact support, and non-stationary signals, but these signals can be further classified according to their characteristics and properties. The wavelet must be suitable for analyzing EEG signal to give a better reconstruction with fewer decomposition levels.

An orthogonal signal is important because it conserves the energy of the signal throughout the wavelet transform so that no information will be lost [25, 33]. It allows wavelet transformation that can extract high and low frequency details. because its wavelets are smoother and can enhance the illustration of transients in the signal. Besides that, EEG signals that contain the features information such compact support also allows the wavelet transform to efficiently characterize [23, 25, 33]. Wavelet families such as Haar, Daubechies, Symlets, and Coiflets have sufficient properties to analyze signal. The Haar wavelet is comprised of a Daubechies order of 1 (db1). Reconstruction coefficient number that had to be more than two, So Haar wavelet was not selected. The Morlet and Mexican Hat wavelet are not selected because they are not suitable for biomedical signal processing [25].

Morlet (mor1) wavelet has no scaling function but is explicit. Mexican Hat (mexh) wavelet has no scaling function and is derived from a function that is proportional to the second derivative function of Gaussian. Meyer (meyr) Scaling function is defined in the frequency domain. Haar (haar) Discontinuous and resembles a step function. Haar represent as Daubechies db1. Symlets (symN) wavelet is modification to the db family. Coiflets (coifN) built by Ingrid Daubechies that has 2N moments. Splines biorthogonal wavelets (biorNr.Nd) wavelet needed two wavelets for signal and image reconstruction. So Daubechies wavelet is chosen because of the smoothing features, without losing any information, the down-sampling of a time domain signal can be divided into low and high filtering [25]. Down-sampling occurs when the original signal x(n), passes through a high-pass filter, g(n), (detail coefficient) and then a low-pass filter, h(n) (approximation coefficient) [25].

3.2 Variable Precision Rough Set

Rough Set model proposed by pawlak can be extended to characterize a set in terms of uncertain information under some levels of certainty [19]. Variable Precision Rough Set (VPRS) is useful for addressing problems where data sets have lots of boundary objects. Variable Precision Rough Set (VPRS) has the additional desirable property of allowing for partial classification instead of complete classification

required by Rough Set (RST) [20]. When an object is classified using Rough set it involves complete certainty that is it goes for a correct classification [21, 34–36]. VPRS has a degree of confidence for detailed analysis of the data in classification, which is achieved through the use of a majority inclusion relation. The measure c (X, Y) of the *relative degree of misclassification* (8) of the set X with respect to set Y defined as

$$\begin{aligned} c(X,Y) &= 1 - |X \cap Y|/|X| \; if \; |X| > 0 \\ or\, c(X,Y) &= 0 \; if \; |X| = 0 \end{aligned} \tag{8}$$

The *majority inclusion relation* (9) under an admissible classification error β (which must be within the range $0 \leq \beta \leq 0.5$ is defined as

$$X \subseteq \beta \quad Y \Leftrightarrow C(X,Y) \leq \beta \tag{9}$$

Let I = (U, A) be an information system, where U is a non-empty, finite set of objects and A is a non-empty, finite set of attributes such that $a : U \rightarrow Va$ for every $a \in A$ [22]. Va is the set of values that attribute a might hold [37]. The information system assigns a value a(x) from Va to each attribute a and object x in the universe U. With any $R \subseteq A$ there is an associated equivalence relation [22, 37]

$$\mathrm{IND(R)} = \{(x, y) \in U2 | \forall a \in R, a(x) = a(y)\} \tag{10}$$

The relation IND(R) is called a *R-indiscernibility relation* (10). The partition of U is a family of all equivalence classes of IND(R) and is denoted by U/IND(R) [37–39].

Let $X \subseteq U$ using equivalence classes induced with attribute subset R By replacing the inclusion relation with majority inclusion relation in the original definition of lower approximation and upper approximation, the generalized notion of *β-lower approximation* (11) and *β-upper approximation* (12)

$$\underline{R}_\beta X = \cup \left\{ [x]_R \in \cup / R : [x]_R \subseteq_\beta X \right\} \tag{11}$$

$$\overline{R}_\beta X = \cup \left\{ [x]_R \in \cup / R : [x]_R \subseteq_{1-\beta} X \right\} \tag{12}$$

The *β-positive region* (13), *β-negative region* (14) and *β-Boundary region* (15) are defined based on upper and lower approximation [40].

$$POS_{R,\beta}(X) = \underline{R}_\beta X \tag{13}$$

$$NEG_{R,\beta}(X) = U - \underline{R}_\beta X \tag{14}$$

$$BN_{R,\beta}(X) = \overline{R}_\beta X - \underline{R}_\beta X \tag{15}$$

The quality of the classification is defined as the proportion of cardinality of positive regions of all the equivalence classes of decision based on the equivalence

classes for a subset P of the condition attributes C [37]. The quality of classification is dependent on the β min value associated with all the condition attributes [37–39].

3.3 Multigranulation Rough Set

The Multigranulation Rough Set (MGRS) is formulated on the basis of a family of binary relations instead of a single indiscernibility relation, where the set approximations are defined by using multi equivalence relations on the universe [41]. A concept of approximation reduct is introduced to describe the smallest attribute subset that preserves the upper approximation and lower approximation of all decision classes in MGRS [41, 42]. The form of decision rules in MGRS is "OR" unlike the "AND" rules from classical rough set model.

In classical Rough set, upper and lower approximations are defined under a single granulation, i.e., the concept is induced from a single relation (such as equivalence relation, reflexive relation and tolerance relation) on the universe. Let us assume P and Q are two sets from predictor features and $X \subseteq U$ is a desired target approach, then the rough set of X is derived from the quotient set U/(PUQ) [41]. The *quotient set* (16) is equivalent to the formula

$$P \cup Q = \{P_i \cap Q_j : P_i \in \cup/P,\ Q_j \in \cup/Q,\ P_i \cap P_j \neq 0\} \tag{16}$$

The above assumption cannot always be satisfied in following cases. In some data analysis issues, for the same object, there is a inconsistent relationship between its values under one attribute set P and those under another attribute set Q [21, 22, 41]. In other words, we cannot perform the intersection operations between their quotient sets and the target concept and cannot be approximated by using U/(PUQ). The decision makers may have independent view for the same project in the universe [21, 22, 41]. In this situation, the intersection operations of one quotient set will be redundant with another quotient set for decision making. The time complexity of rule extractions can be reduced, by not performing the intersection operations in all the sites of distributive information systems. In this approach two models are defined they are optimistic multigranulation rough set and pessimistic multigranulation rough set [22].

Let S = {U, A, V, f} is an information system where U represents universe contains non-empty and finite set of objects and A represents non-empty and finite set of attributes, Va is the domain of the attribute a, V = S a ∈ A Va and f: U × A → V is a function f(x, a) ∈ Va for each a ∈ A. $X \subseteq U$ and $P = \{P_i \subseteq A | P_i \cap P_j = \phi(i \neq j),\ i,j \leq l\}$ [22, 41].

The optimistic lower and upper approximation sets of X with respect to P can be defined as follows [41]

$$\underline{OM}(X) = \left\{ x\varepsilon \cup |\vee([x]_{pi} \subseteq X),\ i \leq l \right\} \tag{17}$$

$$\overline{OM}(X) = \left\{ x\varepsilon \cup |\wedge([x]_{pi} \cap X \neq \phi),\ i \leq l \right\} \tag{18}$$

The pessimistic lower and upper approximation sets of X with respect to P can be defined as follows [41]

$$\underline{PM}(X) = \left\{ x\varepsilon \cup |\wedge([x]_{pi} \subseteq X), i \leq l \right\} \tag{19}$$

$$\overline{PM}(X) = \left\{ x\varepsilon \cup |\vee([x]_{pi} \cap X \neq \phi),\ i \leq l \right\} \tag{20}$$

3.4 Variable Precision Multigranulation Rough Set

Variable Precision Multigranulation Rough Set is general notions of multi-granulation rough sets model to process data with noise and it allows for a controlled degree of misclassification (8) with majority inclusion relation (9) [43]. In the variable precision multi-granulation rough set model, the requirement of accuracy on each granulation is determined by means of parameter α, β and supporting characteristic function wi^{α}, $\mu^{x}_{pi}(x)$ [13]. Let $S = (U; A)$ be an information system, U represents universe contains non-empty and finite set of objects and A is a non-empty and finite set of attributes, $X \subseteq U$ and $P = \{P_i \subseteq A | P_i \cap P_j = \phi(i \neq j),\ i,j \leq l\}$ [43]. l is the number of partitions. Then lower and upper approximation sets of X with respect to P can be defined as follows

$$\underline{VP}(x)^{\alpha}_{\beta} = \{ x\varepsilon \cup | \sum_{i=1}^{l} w^{\alpha}_{i} \mu^{x}_{pi}(x) \geq \beta \} \tag{21}$$

$$\overline{VP}(x)^{\alpha}_{\beta} = \sim \underline{P}(\sim X)^{\alpha}_{\beta} \tag{22}$$

$$wi^{\alpha} = \begin{cases} \frac{1}{7} & \alpha \leq \mu^{X}_{pi}(x) \leq l \\ 0 & \mu^{X}_{pi}(x) < \alpha \end{cases} \tag{23}$$

The parameter α determines the precision of every granulation which are used to approximate the target concept. $\sim$ denotes complementary operation of the set [43]. Let S = (U, A) be an information system, X $\subseteq$ U and $P = \{P_i \subseteq A | P_i \cap P_j = \phi(i \neq j),\ i,j \leq l\}$. If α = 1, then

$$\begin{aligned}\underline{VP}(x)_{\beta}^{1} &= \{x\,\varepsilon\,\cup\,|\sum_{i=1}^{l} w_i^1 \mu_{pi}^x(x) \geq \beta\} \\ &= \{x\,\varepsilon\,\cup\,|\sum_{i=1}^{l} \mu_{pi}^x(x) = 1^{1/l} \geq \beta\} \\ &= \underline{P}(X)_{\beta}^{1}\end{aligned} \tag{24}$$

$$\begin{aligned}\overline{VP}(x)_{\beta}^{1} &= \sim\underline{P}(\sim X)_{\beta}^{1} \\ &= \sim\{x\,\varepsilon\,\cup\,|\sum_{i=1}^{l} \mu_{pi}^x(x) = 1^{1/l} \geq \beta\} \\ &= \sim\{x\,\varepsilon\,\cup\,|\sum_{i=1}^{l} \mu_{pi}^x(x) = 1^{1/l} \geq \beta\} \\ &= \sim\{x\,\varepsilon\,\cup\,|\sum_{i=1}^{l} \mu_{pi}^x(x) = 0^{1/l} \geq \beta\} \\ &= \sim\{x\,\varepsilon\,\cup\,|\sum_{i=1}^{l} \mu_{pi}^x(x) \neq 0^{1/l} \geq \beta\} \\ &= \overline{P}(X)_{\beta}^{1}\end{aligned} \tag{25}$$

These proposed models generalize the multigranulation rough set approach, and are helpful to enhance its capability of dealing with noisy data [43].

3.5 *Game Theoretic Rough Set*

The conventional Pawlak rough set theory does not allow any errors in the positive and negative regions [44]. Researchers argued that the intolerance to errors (or) qualitative absoluteness can lead to limitations in practical applications [45]. Extensions of rough sets were introduced through some measures and thresholds to introduce error tolerance [40]. The probabilistic rough set models is an extension of rough set and include the decision-theoretic rough set model, the Bayesian rough set model, the variable precision rough set model, the information-theoretic rough set model and the game-theoretic rough set model [40, 45]. In these models, a pair of thresholds α, β is used to define the rough set approximations and the resulting three regions [44]. The determination of thresholds value is an important issues in the Variable Precision Multigranulation rough sets [40]. So Game Theoretic Rough Set is applied for determination of α, β pair.

Game theory is a mathematical structure to govern competition in games between two or more parties [44]. It can be used to analyze the classification ability based on different threshold values [44]. It is used to observe the trade-off between accuracy and

precision and the relationships between this trade-off and threshold values [44]. Each player has a strategies (set of actions) with expected payoffs (benefits) as a result of taking that action [40]. A payoff function returns a payoff value by determining the utility of a chosen action [40]. An Assumption is made that all players are rational. Rational players choose strategies (set of actions) that improve their position in the game. Rational players choose strategies (set of actions) that maximize their wining ability while minimizing the other players' ability to do the same. Games are formulated into payoff tables, indicating players involved, possible strategies, and expected payoffs. A single game is defined as G = {O, S, F} O represents a set of players, S represents set of strategies, F represents Action payoffs [40].

Algorithm Input: Dataset in the form of an Information table, Initial values of α^{-}, α^{--}, β^{+}, β^{++}

Output: Thresholds (α, β)

1. Initialize $\alpha = 1.0$, $\beta = 0$
2. Repeat
3. Calculate utilities of player based on Eqs. (26) and (27)
4. Populate the pay off table with calculated values
5. Calculate Equilibrium in a payoff table using Eqs. (28) and (29)
6. Determine Selected Strategies and corresponding Thresholds (α', β')
7. Calculate α^{-}, α^{--}, β^{+}, β^{++} based on Eqs. (30)–(33)
8. (α, β) = (α', β') Until P(BND(α, β)(C)) = 0 or P(POS(α, β)(C)) > P(C) (or) $\alpha \prec 0.5$ *or* $\beta \geq 0.5$

The goal of our work is to improve classification, therefore, each player will represent a measure to achieve a maximum value. Two players are chosen as Accuracy and Dependency. Accuracy is determined by number of objects that are correctly classified as $Pos_{(\alpha,\beta)}$ and $Neg_{(\alpha,\beta)}$. Dependency is determined by number of objects that are in Positive Region divided by total objects in universe.

$$Accuracy(\alpha, \beta) = \frac{\left|(Pos_{(\alpha,\beta)}(C) \cap C) \cup Neg_{(\alpha,\beta)}(C) \cap C^{c}\right|}{\left|Pos_{(\alpha,\beta)}(C) \cup Neg_{(\alpha,\beta)}(C)\right|} \tag{26}$$

$$Dependency(\alpha, \beta) = \frac{\left|(Pos_{(\alpha,\beta)}(C)\right|}{|U|} \tag{27}$$

C^{C} is the set complement of C, containing all objects in U that are not in C. A particular player would prefer a strategy (set of action) over another strategy (set of action) if it provides higher payoff during the game. A strategy profile (s_m, s_n) would be the game solution (or) Nash equilibrium if the following conditions holds.

$$\text{For player A} : \forall s'_m \in S_1, u_A(s_m, s_n) \geq u_A(s'_m \neq s_m) \; with \; (s'_m \neq s_m) \tag{28}$$

$$\text{For player D} : \forall s'_m \in S_1, u_D(s_m, s_n) \geq u_D(s'_m \neq s_m) \; with \; (s'_n \neq s_n) \tag{29}$$

α^- occurs when Single Player Suggest to decrease α, α^- occurs when both players suggest to decrease α, β^+ occurs when single player suggest to increase β, β^{++} occurs when both players suggest to increase β [44].

$$\alpha^{--} = \alpha - (\alpha \times \text{Accuracy}(\alpha', \beta') - \text{Accuracy}(\alpha, \beta)) \tag{30}$$

$$\alpha^{--} = \alpha - \text{C}(\alpha \times \text{Accuracy}(\alpha', \beta') - \text{Accuracy}(\alpha, \beta)) \tag{31}$$

$$\beta^{+} = \beta - (\beta \times \text{Accuracy}(\alpha', \beta') - \text{Accuracy}(\alpha, \beta)) \tag{32}$$

$$\beta^{++} = \beta - \text{c}(\beta \times \text{Accuracy}(\alpha', \beta') - \text{Accuracy}(\alpha, \beta)) \tag{33}$$

4 Proposed Methodology

4.1 Acquisition of Dataset

Motor Imagery data are collected from BCI Competition IV dataset 1. The recording was made using 59 Channel BrainAmp MR plus amplifiers and a Ag/AgCl electrode cap of 10-20 System. These data sets were recorded from healthy subjects. The dataset epoch size is about 190594 × 59 of which 200 positions are marked in training data. For each subject two classes (Hand, Foot) of motor imagery were selected where the class Hand may be either Left Hand (or) Right Hand. Signals from 59 EEG positions were measured that were most densely distributed over sensorimotor areas as shown in Fig. 2. Seven calibration data has been used for training and eight evaluation data has been used for testing. Band-pass filtered are applied for signal between 0.05 and 200 Hz and then digitized at 1000 Hz with 16 bit (0.1 μV) accuracy. Another Motor imagery data is collected from BCI Competition III dataset 3a which is Multiclass motor imagery data set. EEG amplifier 64-channel was used for recording from Neuroscan cap of 10-20 System. Sixty EEG channels were recorded. The dataset epoch size is about 986780 × 60 of which 300 positions are marked in k3b dataset and 240 position are marked in k6b and l1b dataset. Data was collected from three subjects (ranging from quite good to Fair Performance).

4.1.1 BCI Competition IV Dataset 1

Calibration data

In the first two trial runs (or) test, arrows pointing left, right, or down were presented as visual cues on a screen. Cued motor imagery task involves displaying

cues for a period of 4 s. After these periods 2 s of blank screen are displayed and then displays fixation cross in the center of the screen for period of 2 s. The fixation cross was superimposed with cues, i.e. it was displayed for 6 s. These data sets are provided to user with complete marker information.

4.1.2 BCI Competition III Dataset 3a

In BCI Competition III dataset 3a training data and test data were recorded with the same task and from the same subject, but on two different days with about 1 week in between. EEG was sampled with 250 Hz, it was filtered between 1 and 50 Hz. The task was to perform imagery left hand, right hand, foot or tongue movements according to a cue. The random cues was displayed to subject. The recording of BCI data consists of several runs (at least 6) with 40 trials each after trial begin, the first 2 s were displayed nothing, at t = 2 s an cross "+" is displayed indicating the beginning of the trial, then from t = 3 s an arrow to the left, right, up or down was shown for 1 s; at the same time the subject was asked to imagine a, tongue or foot movement, left hand, right hand, until the cross disappeared at t = 7 s.

4.2 Chebyshev Type 2 Filter for Pre-processing

EEG signals has very small amplitudes and they can be easily contaminated by noise [26, 27, 31]. The electrode noise can be generated from the body itself. Artifacts are noises in the EEG signals and need to be removed from the original signal for the analysis of the EEG signals [17, 28]. The various noises that can occur in the signals during recordings are the power line Interference, baseline movement, electrode noise, EMG disturbance and so on [24].

Chebyshev filters is Infinite Impulse Response (IIR) Filter that have steeper roll off and more pass band ripple than other filters [23]. Chebyshev filters have the property that they minimize the error between the actual and the idealized filter characteristic over the range of the filter with ripples in the pass band. Because of the passband ripple inherent in Chebyshev filters, they have a smoother response in the passband. [23, 30, 32]. They are used to separate one band of frequencies from another. It performs faster than butter worth and elliptic filter since they are carried out by recursion rather than convolution. Sensorimotor rhythm occurs in the range of mu (6.25–12.5) and beta (12.5–25) rhythm. In the BCI competition data set, since our goal is to classify sensorimotor rhythm (12.5–15.5). Passband cutoff frequency is a scalar vector with values between 0 and 1, with 1 corresponding to the normalized Nyquist frequency, π radians per sample. Hence the frequency band range is set between 4–40 Hz with cutoff frequency 0.7 Hz [30, 32].

Table 1 Decomposition level with BCI competition dataset

Decomposition levels	Frequency range
D_1 (Theta, Alpha)	4–13
D_2 (Alpha, Beta)	13–22
D_3 (Beta)	22–31
D_4 (Gamma)	31–40

4.3 Daubechies Wavelet for Feature Extraction

The Discrete Wavelet Transform is used to decompose EEG signal at resolution levels of the components of the EEG signal (Delta (δ)—0.5–3 Hz, Theta (θ)—4–7 Hz, Alpha (α)—8–13 Hz, Beta (β)—13–31 Hz and Gamma (γ)—above 31 Hz) with the Multi-Resolution Analysis (MRA) [29]. The object of wavelet analysis is to decompose signals into several frequency bands [25, 33]. Selection of appropriate wavelet and the number of decomposition levels are very important for the analysis of signals using DWT [29]. The number of decomposition levels is chosen based on the dominant frequency components of the signal.

In the wavelet coefficient, the levels are chosen such that parts of signal that correlate well with the frequencies for classification of the signal. In this work, Daubechies 4 (db4) is selected because its smoothing feature can detect changes of the EEG signal. The frequency band [fm/2 : fm] of each detail scale of the Discrete Wavelet Transform is related to the original signal sampling rate, which is given by fm = fs/2 l + 1, where fs is the sampling frequency, and l is the level of decomposition [25]. In this study, the sampling frequency is 100 Hz of the EEG signal. Nyquist theorem suggest that the highest frequency of signal would be fs/2. Among this decomposition levels D1 and D2 contains sensorimotor rhythm (12.5–15.5) which is passed as input. Frequency bands corresponding to five decomposition levels for wavelet db4 were listed in Table 1. The signals were decomposed into details D1–D4.

4.4 Classification Based on VPMGRS

Algorithm for Variable Precision multigranulation Rough set is as follows. Input are decomposed frequency bands D_1, D_2 (alpha, Beta) of EEG Signal.

Algorithm 1: Variable Precision Multigranulation Roughset An information system is a 4-tuple S = {U, A, V, f} where U is a non-empty and finite set of objects, called a universe, and A is a non-empty and finite set of attributes, Va is the domain of the attribute a, and f: $U \times A \rightarrow V$ is a function $f(x, a) \in Va$ for each $a \in A$. An

indiscernibility relation RB = {(x, y) ∈ U × U | a(x) = a(y), ∀a ∈ B} was determined by a non-empty subset B ⊆ A. Assign P_i to P. Let m represents number of elements [10, 43].

(1) $P = \{P_i \subseteq A | P_i \cap P_j = \phi(i \neq j), i, j \leq l\}$. Partition Generated with OR rules (Multigranulation Rough Set) instead of AND rules (Pawlak Rough Set) l is the number of partitions [10].

(2) Find the Supporting Characteristic Function wi^{α} and $\mu_{pi}^{X}(x)$

For i = 1:l
For j = 1:m

$\mu_{pi}^{X}(x_j)$ = Number of elements that match Decision and Subpartition/Total number of elements in subpartition

$$wt_j^{\alpha} = \begin{cases} \frac{1}{l} & \alpha \leq \mu_{pi}^{X}(x_j) \leq l \\ 0 & \mu_{pi}^{X}(x_j) < \alpha \end{cases}$$

End For

End For

(3) Then lower and upper approximation sets of X with respect to P can be defined as follows and explained in detail in Sect. 3.4

$$\underline{VP}(x)_{\beta}^{\alpha} = \{x \varepsilon \cup | \sum_{i=1}^{l} w_i^{\alpha} \mu_{pi}^{x}(x) \geq \beta\}$$
$$\overline{VP}(x)_{\beta}^{\alpha} = \sim \underline{P}(\sim X)_{\beta}^{\alpha}$$

Let us consider the Sample data

U = {x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}} and X denotes one of the decision variable X = {x_1, x_2, x_6, x_8, x_9} [10]

Let p1, p2 be the two partitions under consideration computed with Multigranulation Rough Set [10]

$$U/p_1 = \{\{x_1, x_7\}\{x_2, x_3, x_4, x_6, x_8, x_9\}\{x_5\}\{x_{10}\}\}$$
$$U/P_2 = \{\{x_1\}\{x_2, x_3, x_6, x_8, x_9\}\{x_4, x_5\}\{x_7\}\{x_{10}\}\}$$

Step 2

$X = \{x_1, x_2, x_6, x_8, x_9\}\, U/p_1 = \{\{x_1, x_7\}\{x_2, x_3, x_4, x_6, x_8, x_9\}\{x_5\}\{x_{10}\}\}$
$\mu_{P1}^{X}(x_1) = 0.5\, P1 = \{x1, x7\}$

Here Total Number of elements are 2
Only x1 are in X so Supporting Characteristic Function is ½ = 0.5.

$\mu^X_{P1}(x_2) = 0.6$ In P1 = {x2, x3, x4, x6, x8, x9}

Here Total Number of elements are 6

x_2, x_6, x_8, x_9 are in X so Supporting Characteristic Function is

4/6 = 0.6. Likewise all the elements supporting characteristic function for each partitions are found.

$$\begin{aligned}
\mu^X_{P1}(x_1) &= 0.5 \quad \mu^X_{P2}(x_1) = 1 \\
\mu^X_{P1}(x_2) &= 0.6 \quad \mu^X_{P2}(x_2) = 0.6 \\
\mu^X_{P1}(x_3) &= 0.6 \quad \mu^X_{P2}(x_3) = 0.6 \\
\mu^X_{P1}(x_4) &= 0.6 \quad \mu^X_{P2}(x_4) = 0 \\
\mu^X_{P1}(x_5) &= 0 \quad \mu^X_{P2}(x_5) = 0 \\
\mu^X_{P1}(x_6) &= 0.6 \quad \mu^X_{P2}(x_6) = 0.6 \\
\mu^X_{P1}(x_7) &= 0.5 \quad \mu^X_{P2}(x_7) = 0 \\
\mu^X_{P1}(x_8) &= 0.6 \quad \mu^X_{P2}(x_8) = 0.6 \\
\mu^X_{P1}(x_9) &= 0.6 \quad \mu^X_{P2}(x_9) = 0.6 \\
\mu^X_{P1}(x_{10}) &= 0 \quad \mu^X_{P2}(x_{10}) = 0
\end{aligned}$$

wi^{α} varies according to the $\mu^X_P(x_i)$ [10]. If $\mu^X_P(x_i)$ is greater than α then wi^{α} is 1/l, where l is the number of partitions. If $\mu^X_P(x_i)$ is less than α then wi^{α} is 0 [10]. When the product of wi^{α} and $\mu^X_P(x_i)$ is greater than or equal to β then it is added to lower approximation [10].

$$\begin{aligned}
\underline{VP}(x)^{0.3}_{0.3} &= \{x_1, x_2, x_3, x_4, x_6, x_8, x_9\} \\
\underline{VP}(x)^{0.6}_{0.6} &= \{x_2, x_3, x_6, x_8, x_9\} \\
\underline{VP}(x)^{0.7}_{0.3} &= \{x_1\}
\end{aligned}$$

5 Experimental Results and Discussion

In Variable precision multigranulation rough set Classification accuracy is evaluated using accuracy measures. Confusion matrix visualizes performance of the algorithm. Each column of the matrix represents the instances in a predicted class, while each row of the matrix represents the instances in an actual class. The accuracy (AC) gives total number of correct predictions. The recall or true positive rate (TP) is the proportion of correctly identified positive cases. The false positive rate (FP) is the proportion of incorrectly classified negatives cases. The true negative rate (TN) is defined as the proportion of correctly classified negatives cases that were classified correctly The false negative rate (FN) is the proportion of incorrectly classified positives cases. A confusion matrix is a visualization tool used

in supervised machine learning for classification. Instances of the predicted class and actual class are represented in each column and each row. The confusion matrix calculates the number per class of well classified and mislabeled instances and evaluate the performance of a classifier.

Cohen's kappa measures the similarity between two raters who each classify N items into C mutually exclusive categories [46, 47]. The motivation of this measure is to extract from the correctly classified percentage the actual percentage expected by chance [46, 47]. The equation is as follows

$$\kappa = \frac{P(D) - P(E)}{1 - P(E)} \tag{34}$$

where *P(D)* is the percentage of classified instances that are correct and *P(E)* is the Expected Proportion by chance [46, 47]. A κ Coefficient equals to one means perfect agreement and zero means poor agreement. In BCI Competition III Dataset IIIa contains three data with 60 channels and 4 class labels where 1 indicates Left Hand, 2 indicates Right Hand, 3 indicates Foot and 4 indicates Tongue. In BCI Competition IV Dataset I Seven Calibration data with 59 channels has been used. It contains two classes where 1 indicates Right Hand (or) Left Hand and 2 indicates Foot. The proposed methodology follows k-fold cross validation. In *k*-fold cross-validation, the original sample is randomly partitioned into k equal sized subsamples. Of the k subsamples, a single subsample is retained as the validation data for testing the model, and the remaining $k - 1$ subsamples are used as training data. The cross-validation process is then repeated k times (the *folds*), with each of the k subsamples used exactly once as the validation data. The k results from the folds can then be averaged (or otherwise combined) to produce a single estimation. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and validation, and each observation is used for validation exactly once.

Thresholds α and β are determined with game theoretic rough set and assigned the constant $\alpha = 0.5$, $\beta = 0.25$. The proposed methodology has been compared with existing techniques Naïve Bayes, Multi layer Perceptron, IBk, Decision Table, Random Tree, J48. IBk is K-Nearest Neighbour Classifier where k is chosen as 2 for BCI Competition IV and k is chosen as 4 for BCI Competition III dataset. Distance weighting method Linear NN Search (Euclidean distance) method is used in IBk.

Naive Bayes Classifier is based on estimator classes. Numeric estimator precision values are chosen based on analysis of the training data. Multilayer perceptron classifier that uses backpropagation to classify instances. The network can also be monitored and modified during training time. The nodes in this network are all sigmoid (except for when the class is numeric in which case the output nodes become unthresholded linear units), Rate is assigned with value 0.3, Momentum (weight updation) is assigned with value 0.2, number of hidden layers is assigned based on formula number of attributes including class attribute divided by 2 and Validation threshold is assigned with value 20. J48 generates a pruned C4.5 decision tree where confidence factor (For Pruning) is assigned with value 0.25, MinNumObj

(Minimum number of Instances per leaf) is assigned with value 2. numFolds is assigned with value 3 and it determines the amount of data used for reduced-error pruning. One fold is used for pruning, the rest for growing the tree. seed is assigned with value 1 for randomizing the data when reduced-error pruning is used.

Table 2 Classifiication of BCI competition IV dataset 1

Method used	Calib ds1a	Calib ds1b	Calib ds1c	Calib ds1d	Calib ds1e	Calib ds1f	Calib ds1 g
	Kappa value						
Naïve bayes	0.56	0.50	0.55	0.55	0.59	0.51	0.52
Multi layer perceptron	0.57	0.54	0.62	0.61	0.65	0.67	0.68
IBk	0.58	0.58	0.69	0.69	0.79	0.62	0.63
Random tree	0.58	0.53	0.59	0.59	0.62	0.56	0.56
J48	0.58	0.52	0.58	0.59	0.58	0.57	0.57
VPMG RS	1	0.68	0.63	0.63	1	0.84	0.77

Table 3 Classification of BCI competition III data set 3a

Technique used	K3b	K6b	L1b
Naïve bayes	0.43	0.55	0.53
Multilayer perceptron	0.44	0.53	0.50
IBK	0.49	0.57	0.57
Random tree	0.46	0.46	0.50
J48	0.45	0.47	0.52
VPMGRS	1	0.74	0.74

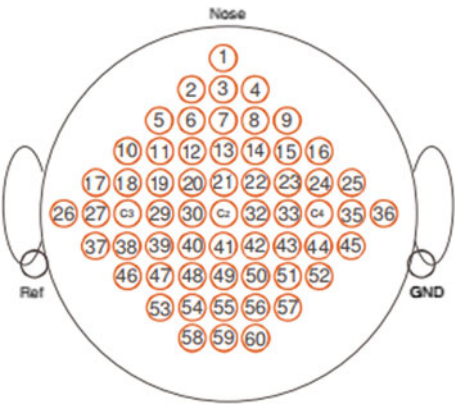

Fig. 3 EEG electrode placement for BCI competition III dataset 3a

In Random Tree Method a tree is constructed by considering K randomly chosen attributes at each node. It performs no pruning. Also has an option to allow estimation of class probabilities based on a hold-out set (backfitting). Max Depth is assigned the value 0. It determines the maximum depth of the tree, 0 for unlimited. minNum is assigned with value 1.0 It determines the minimum total weight of the instances in a leaf. numFolds is assigned with value 0. It determines the amount of data used for backfitting. One fold is used for backfitting, the rest for growing the

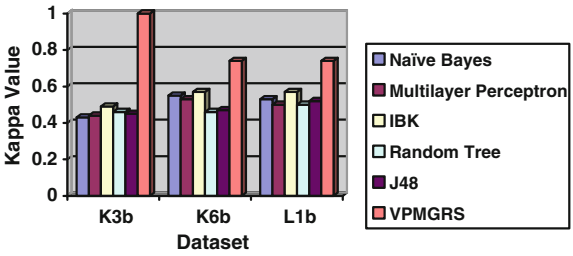

Fig. 4 Classification of BCI competition III data set 3a

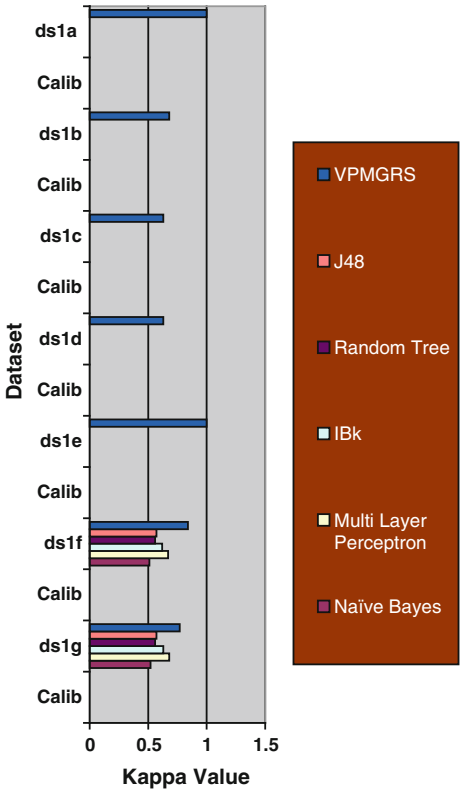

Fig. 5 Classification of BCI competition IV data set 1

tree. (Default Value: 0 denotes no backfitting). seed is assigned with value 1. The random number seed used for selecting attributes.

It is clear from Tables 2, 3 and Figs. 3, 4, 5 that the proposed methodology (VPMGRS) gives higher kappa value compared to existing methods. In Calib_ds1a, Calib_ds1e highest accuracy obtained using existing technique is about 0.58, 0.79 but proposed technique gives about 100 % accuracy. In Calib_ds1b, highest accuracy obtained using existing technique is about 0.58 but proposed technique gives 0.68. In Calib_ds1f, highest accuracy obtained using existing technique is about 0.67 but proposed technique gives 0.84. In Calib_ds1g, highest accuracy obtained using existing technique is about 0.68 but proposed technique gives 0.77. In Calib_ds1c and Calib_ds1d highest accuracy obtained using existing technique is about 0.69 but proposed technique gives 0.63 which is almost nearer to highest accuracy. In k3b dataset existing techniques gives accuracy of less than 50 % accuracy but proposed methodology gives 100 % accuracy. In k6b and L1b dataset highest Kappa obtained from existing techniques is about 0.57 but proposed methodology gives 0.74 kappa value.

6 Conclusion

Thus the BCI Competition IV dataset 1, BCI Competition III data set 3a shows higher accuracy when compared with existing techniques. BCI Competition IV dataset 1 is a two class problem (Right Hand, Foot). BCI Competition III data set 3a is a four class problem left hand, right hand, both feet, and tongue. Variable precision multigranulation rough set involves partition by multigranulation rough set and then based on partition μ values are computed and supporting characteristic function are computed based on partition. Both Supporting Characteristic Function w_i^{α}, μ, α, β values determine classification accuracy. The variation of α, β values in both BCI Competition Data set depicts higher accuracy. The flexibility to handle uncertain information in EEG data is the main advantage over state of art methods.

References

1. Frolov, A.A., Husek, D., Snasel, V., Bobrov, P., Mokienko, O., Tintera, J., Rydlo, J.: Brain-computer interface based on motor imagery: the most relevant sources of electrical brain activity. In: Proceedings of Soft Computing in Industrial Applications, pp. 1–10 (2012)
2. Lécuyer, A., Lotte, F., Reilly, R.B., Leeb, R., Hirose, M., Slater, M.: Brain-computer interfaces, virtual reality, and videogames. Computer **10**(41), 66–72 (2008)
3. Rebsamen, Brice, Burdet, Etienne, Guan, Cuntai, Zhang, Haihong, Teo, Chee Leong, Zeng, Qiang, Laugier, Christian, Jr Ang, Marcelo H.: Controlling a wheelchair indoors using thought. Intell. Syst. IEEE **22**(2), 18–24 (2007)

4. Mugler, E.M., Ruf, C.A., Halder, S., Bensch, M., Kubler, A.: Design and implementation of a P300-based brain-computer interface for controlling an internet browser. In: Neural Syst. Rehabil. Eng., IEEE Trans. **18**(6), 599–609 (2010)
5. Lotzea, Martin, Halsband, Ulrike: Motor imagery. J. Physiol. **99**, 386–395 (2006)
6. Grosse-Wentrup, Moritz: Understanding brain connectivity patterns during motor imagery for brain-computer interfacing. Adv. Neural Inf. Process. Syst. (NIPS) **21**, 561–568 (2008)
7. Nicolas-Alonso, L.F., Corralejo, R., Álvarez, D., Hornero, R.: Adaptive classification framework for multiclass motor imagery-based BCI, Chapter XIII. In: Mediterranean Conference on Medical and Biological Engineering and Computing, vol. 41, pp. 762–765 (2013)
8. Miranda, Eduardo Reck, Brouse, Andrew: Interfacing the brain directly with musical systems: on developing systems for making music with brain signals. Leonardo **38**(4), 331–336 (2005)
9. Cabredo, Rafael, Legaspi, Roberto, Inventado, Paul Salvador, Numao, Masayuki: Discovering emotion-inducing music features using EEG signals. J. Adv. Comput. Intell. Intell. Inform. **17** (3), 362–370 (2013)
10. Renuga Devi, K., Hannah Inbarani, H.: Motor imagery classification based on variable precision multigranulation rough set. Adv. Intell. Syst. Comput. **412**, 145–154 (2016)
11. Sitaram, Ranganatha, Zhang, Haihong, Guan, Cuntai, Thulasidas, Manoj, Hoshi, Yoko, Ishikawa, Akihiro, Shimizu, Koji, Birbaumer, Niels: Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface. NeuroImage **34**, 1416–1427 (2007)
12. Corralejo, Rebeca, Nicolás, Luis F., Alonso, Álvarez, Daniel, Hornero, Roberto: A P300 based brain–computer interface aimed at operating electronic devices at home for severely disabled people. Med. Biol. Eng. Comput. **52**, 861–872 (2014)
13. Kaiser, Vera, Bauernfeind, Günther, Kaufmann, Tobias, Kreilinger, Alex, Kübler, Andrea, Neuper, Christa: Cortical effects of user learning in a motor-imagery BCI training. Int. J. Bioelectromag. **13**(2), 60–61 (2011)
14. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Braincomputer interfaces for communication and control. Clin. Neurophysiol. **113**(6), 767–791 (2002)
15. Velásquez-Martínez, L.F., Álvarez-Meza, A.M., Castellanos-Domínguez, C.G.: Motor imagery classification for BCI using common spatial patterns and feature relevance analysis. Nat. Artif. Comput. Eng. Med. Appl. **7931**, 365–374 (2013)
16. Duan, L., Zhong, H., Miao, J., Yang, Z., Ma, W., Zhang, X.: A voting optimized strategy based on ELM for improving classification of motor imagery BCI data. Cogn. Comput. **6**, 477–483 (2014)
17. Rodrıguez-Bermudez, G., García-Laencina, P.J.: Analysis of EEG signals using nonlinear dynamics and chaos: a review. Appl. Math. Inf. Sci. **9**(5), 2309–2321 (2015)
18. Kubler, A., Kotchoubey, B., Kaiser, J., Wolpaw, R.J., Birbaumer, N.: Brain-computer communication: unlocking the locked in. Psychol. Bull. **127**(3), 358–375 (2001)
19. Calvo-Dmgz, D., Galvez, J.F., Glez-Pena, D., Gomez-Meire, S., Fdez-Riverola, F.: Using variable precision rough set for selection and classification of biological knowledge integrated in DNA gene expression. J. Integr. Bioinf. **9**(3), 1–17 (2012)
20. Wei, Jin-Mao, Wang, Ming-Yang, You, Jun-Ping: VPRSM based decision tree classifier. Comput. Inform. **26**, 663–677 (2007)
21. Beynon, Malcolm J., Peel, Michael J.: Variable precision rough set theory and data discretisation: an application to corporate failure prediction. Omega **29**, 561–576 (2001)
22. ziarko, Wojciech: Variable precision rough set model. J. Comput. Syst. Sci. **46**, 39–59 (1993)
23. Ang, K.K., Chin, Z.Y., Wang, C., Guan, C., Zhang, H.: Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Front. Neurosci. 6–39 (2012)
24. Zhou, H., Wang, Y., Xu, Q., Huang, J., Wu, J.: An improved support vector machine classifier for EEG-based motor imagery classification. Adv. Neural Netw. **5552**, 267–275 (2009)
25. Mohd Tumari, S.Z., Sudirman, R., Ahmad, A.H.: Selection of a suitable wavelet for cognitive memory using electroencephalograph. Signal Eng. **5**, 15–19 (2013)

26. Procházka, A., Mudrová, M., Vyšata, O., Gráfová, L., Araujo, S.P.S.: Computational intelligence in multi-channel EEG signal analysis. Recent Adv. Intell. Eng. Syst. **378**, 361–381 (2012)
27. Deepa, V.B., Thangaraj, P.: A study on classification of EEG data using Filters, (IJACSA) Int. J. Adv. Comput. Sci. Appl. **2**(4) (2011)
28. Kang, D., Zhizeng, L.: A method of denoising multi-channel EEG signals fast based on PCA and DEBSS algorithm. In: International Conference on Computer Science and Electronics Engineering (ICCSEE), vol. 3, pp. 322–326 (2012)
29. Omerhodzic, I., Avdakovic, S., Nuhanovic, A., Dizdarevic, K.: Energy distribution of EEG signals: EEG signal wavelet-neural network classifier. World Acad. Sci. Eng. Technol. **4**, 1–24 (2010)
30. Sviderskaya, N.E., Bykov, P.V.: EEG spatial organization during intense hyperventilation (Cyclic breathing) EEG correlates of psychovisceral phenomena. Hum. Physiol. **32**(3), 270–277 (2006)
31. Chen, Chih-Wei, Ming-Shaung, Ju, Sun, Yun-Nien, Lin, Chou-Ching K.: Model analyses of visual biofeedback training for EEG-based brain-computer interface. J. Comput. Neurosci. **27**, 357–368 (2009)
32. Li, Xiaowei, Zhao, Qinglin, Liu, Li, Peng, Hong, Qi, Yanbing, Mao, Chengsheng, Fang, Zheng, Liu, Quanying: Improve affective learning with EEG approach. Comput. Inform. **29**, 557–570 (2010)
33. Yu, L.: EEG de-noising based on wavelet transformation. In: International Conference Bioinformatics and Biomedical Engineering, pp. 1–4 (2009)
34. Inuiguchi, Masahiro: Attribute reduction in variable precision rough set model. Int. J. Uncertainty Fuzziness Knowl. Based Syst. **14**(4), 61–479 (2006)
35. Inuiguchi, Masahiro: Structure-based attribute reduction in variable precision rough set models. J. Adv. Comput. Intell. Intell. Inform. **10**(5), 657–665 (2006)
36. Ningler, Michael, Stockmanns, Gudrun, Schneider, Gerhard, Kochs, Hans-Dieter, Kochs, Eberhard: Adapted variable precision rough set approach for EEG analysis. Artif. Intell. Med. **47**, 239–261 (2009)
37. Kusunoki, Yoshifumi, Inuiguchi, Masahiro: Variable precision rough set model in information tables with missing values. J. Adv. Comput. Intell. Intell. Inform. **15**(1), 110–116 (2011)
38. Gong, Zengtai, Shi, Zhanhong, Yao, Hongxia: Variable precision rough set model for incomplete information systems and its -reducts. Comput. Inform. **31**, 1385–1399 (2012)
39. Ziarko, W.: A variable precision rough set model. J. Comput. Syst. Sci. **46**, 39–59 (1993)
40. Yao, J.T., Herbert, J.P.: A game-theoretic perspective on rough set analysis. J. Posts Telecommun. **20**(3), 291–298 (2008)
41. Xu, W., Zhang, X., Wang, Q.: A generalized multigranulation rough set approach. ICIC 681–689 (2012)
42. Yao, Y., Qian, Y., Liang, J., Dang, C.: MGRS: a multigranulation rough set. Inf. Sc. 1–22 (2009)
43. Wei, W., Liang, J., Qian, Y., Wang, F.: Variable precision multi-granulation rough set. In: IEEE International Conference on Granular Computing, pp. 639–643 (2012)
44. Azam, Yao, J.T.: Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets. Int. J. Approx. Reason. **55**(1), 142–155 (2014)
45. Azam, Nouman, Yao, JingTao: Game-theoretic rough sets for recommender systems. Knowl. Based Syst. **72**, 96–107 (2014)
46. de Vries, Sjoerd, Mulder, Theo: Motor imagery and stroke rehabilitation. A Crit. Discuss. J. Rehabil. Med. **39**, 5–13 (2007)
47. Vaughan, T.M., Heetderks, W.J., Trejo, Lj, Rymer, W.Z., Weinrich, M., Moore, M.M., Kubler, A.: Brain-computer interface technology: a review of the second international meeting. IEEE Trans. Neural Syst. Rehabil. Eng. **11**(2), 94–109 (2003)