# Cursor Movement Detection in Brain-Computer-Interface Systems Using the Hybrid K-Means Clustering Method and LSVM

Leila Mohammadi
   tehran north branch, islamic azad university
zahra einalou ( ✉ z.einalou@srbiau.ac.ir )
   Islamic Azad University Tehran North Branch
Hamidreza Hosseinzadeh
   Islamic Azad University Tehran North Branch
Mehrdad Dadgostar
   Islamic Azad University Central Tehran Branch

Research

# Cursor movement detection in Brain-Computer-Interface systems using the hybrid K-means clustering method and LSVM

Leila Mohammadi[1], Zahra Einalou[1*], Hamidreza Hosseinzadeh[1], Mehrdad Dadgostar[2]

[*1] Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran.
[2] Department of Biomedical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran.

z.einalou@srbiau.ac.ir

**Corresponding author**

Zahra Einalou

Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran,

Iran

E-mail: z.einalou@srbiau.ac.ir

Phone: +98 21 77318896

Fax: +98 21 77318896

**Abstract:**

In this study, we present the detection of the up- and downward as well as the right- and leftward motion of cursor based on feature extraction. Feature Extraction and selection for finding the proper classifier among the data mining methods are of great importance. In the proposed method, the hybrid K-means clustering algorithm and the linear support vector machine (LSVM) classifier have been used for extracting the important features and detecting the cursor motion. In this algorithm, the K-means clustering method is used to recognize the available hidden patterns in each of the four modes (up, down, left, and right). The identification of these patterns can raise the accuracy of classification. The membership degree of each feature vector in the proposed new patterns is considered as a new feature vector corresponding to the previous feature vector and then, the cursor motion is detected using the linear SVM classifier. The database of the Karadeniz Technical University of Turkey has been used in the present article. Applying the proposed method for data based on the hold-up cross validation causes the accuracy of the classifier in the up- and downward and left- and rightward movements in each person to increase by 2-10%.

**Keywords**:Brain-computer-interface system, EEG signal, hybrid K-means clustering, linear support vector machine (LSVM) classifier.

## 1  Introduction

The relationship between humans and computer is one of the research fields in which many investments have been made in recent years. Over the past two decades, various human-computer-interfaces have been designed that utilized audition, vision, somatosensation or a combination of them. Serious efforts have been made since the past decade to relate the computer to the environment based on analysis of brain electrical signals. The main purpose of these systems

is to help people with corticospinal damages. Despite of having healthy brains, these people are usually not able to move or build a normal and natural relationship with their surrounding environment. Recent advances in neuroscience, physiology, signal analysis, machine learning, and hardware have made it possible to design direct brain-computer communication systems, known as brain-computer-interfaces (BCI), which enable the physically disabled patients to do their affairs without the help of others. In general, a brain-computer interface is a system that allows people with disabilities to operate electrical devices like a computer pointer, a robotic arm, or even a cellular phone with their mind. Regarding its implementation, a BCI system evaluates the specific features of the brain activity, and accordingly translates them into the signals related to device control. The input signals may be electromagnetic [1], infrared spectroscopy [2-7], magnetic resonance imaging [8, 9], electro-echography [10] and EEG signal [11, 12]. Among these techniques, the EEG signal is used more widely due to its simplicity, low cost, non-invasiveness and ability to register and analyze the data.

A BCI program specifies appropriate physical or mental tasks and selects electrodes related to these tasks. It extracts features of signal recording by means of these electrodes, and ultimately produces an algorithm with the highest classifiers. Then, it the information is transferred to communication and control units through implementing the algorithms to translation. Developing a quick and precise classification method is a key subject for communication and control over the EEG signal. Computer cursor control is one of the most useful applications of BCI. The first studies on the cursor control based on EEG signal were focused on one-dimensional control of cursor [13, 14]. After obtaining satisfactory results, the attention of researchers was drawn to the multi-dimensional cursor control with the aim of increasing the relationship between the user and machine. Based on the studies, the multi-dimensional cursor can

be controlled by implementing different brain signals, for example the P300 potential [15, 16], synchronous and asynchronous signals [17] and evoked potentials [18, 19]. They showed well that the input signals of an EEG-based brain-computer-interface system have commonly weak, non-constant and mind tasks-related noises with various artifacts such as external electromagnetic waves and electromyogram and electro-echogram waves. All of these defects were considered by researchers to improve the key features of BCI system function including the classifier velocity and precision. In addition to improving the functional features, the operational and realistic construction of the EEG-based BCI systems was another major challenge from the viewpoint of researchers. It highly relies on using a training process, which is based on a small number of training signals, classifying channel algorithms and features which can represent the studies tasks better. The proper input design, features extraction and appropriate classification of a BCI system has been undertaken over the past 20 years. Various features of the EEG signal serve as inputs for a BCI, such as the Mu band (8-12 HZ) and beta band (18-25 HZ) [29], event-related potential such as P300 [30, 31], steady-state visual-evoked response or the surface slow cortical potential (SPCs) [32, 33], and several feature extraction methods for motor imagery data as input signals, including AR parameter [34] estimation in a specific frequency band [29, 36] ,domain of slow cortical potential [38], common space pattern, event-related synchronization (ERD / ERS), wavelet correlation coefficient [37], and spatial pattern spectrum [35]. Many of the classifier algorithms, such as K-nearest, Linear Discriminate Analysis (LDA) [39], hidden Markov model [40], Neural Networks [24, 41], Multilayer Perceptron, and Mahalanobis Distance (MDA) [42, 43] were tested for BCI programs. The feature extraction and classifier algorithm has been investigated in BCI applications [44]. Since 2001, several BCI competitions have been organized for improved target assessment and comparison of methods for EEG data derived from mental tasks and also

addressing of key issues in BCI research, and significant results have been reported for BCI data. Currently, BCI technology lacks desirable speed and precision, and these are two important variables affecting current attempts to address issues related to the creation of real BCI systems in the future.To this aim, the classifier can discriminate between the EEG signals recorded in different sessions and days with the mental tasks belong to the BCI system. It should also be borne in mind that BCI systems have the potential to help people of all ages with severe mobility impairments. Although some studies on the BCI performance in older people didn't yield good results due to poor rate of data transfer, .In the BCI competitions, [34] obtained a classifier precision of 88.7% with a linear classifier using gamma band power (Channel 4) and SCPs (Channel 1) [44]. Then, by selecting a neural network as classifier, the classifier accuracy was enhanced to 91.47%. In addition, [30, 39] by extracting features of 6 channels and using wavelet transform algorithm, they obtained a classifier accuracy of 90.8% by the neural network. [39, 45] used principal component analysis (PCA) in 6 channels and their spectral properties, reporting an accuracy of 90.44%. Trejo et al. [46] showed that people of various ages can benefit from this technology. It has also been observed in recent studies that the combination of EEG and EMG signals in BCI systems can be used for several activities in cases of mobility impairments. They believed that users would be able to understand how to interact with such interactive systems [20]. Although these methods have desirable classifier accuracy, according to complete training data, they were capable of using more than one electrode and complex vectors. Therefore, they are complex, low portable and in need of extensive training and computational time. Their results not only revealed the accuracy of diagnosis, but also significantly reduced the calculation time for training due to the reduced feature space. However, the drawback of these methods is that they used training accuracy as a criterion for the evaluation of different combinations of features. In

fact, different subsets of features were trained for creating the optimal subset with the best accuracy for diagnosis, although it was ineffective. The objective of this study is to classify the EEG signal in the BCI systems using data mining technique. As we know, data mining plays a significant role in extraction of hidden information or patterns and relationships in a large amount of data. Feature extraction and selection are very important in data mining. In addition, an increase occurs in the computational time when the number of descriptive features increases. To deal with the large number of features in the present study, a method for recognizing the pattern and extracting the essential information for detection of the cursor movement has been studied. The present study aimed to find a method for the input motion efficently and efficently by implementing the data mining technique. This is done using the hybrid K-means clustering method and the linear support vector machine (LSVM) classifier [21] for detecting the upward, downward, leftward and rightward motion.

The present paper's agenda is as follows: Section 2 describes the collection of data used for classification. Besides, it presents the K-means algorithm was proposed as an unsupervised algorithm to extract motion features and prevent from repetiting the training in different subsets. All of the data related to feature can be stored in a dense method for one-time training instead of performing multiple training pilots in different feature subsets since the K-means algorithm clusters the principal feature space with unsupervised learning. Therefore, a membership function was used to obtain optimal results from the K-means algorithm for training. Further, Section 3 provides a summary of a combination of K-means and LSVM classifier for motion detection with high accuracy has been proposed in this study. Lastly, Sections 4 and 5 give critical insights into the performed interpretation and accuracy, and address some limitations and possibilities of the presented method.

## 2  Materials and Methods

### 2.1 Data Acquisitions

The research was performed by online collection of EEG signals from three normal subjects at two different sessions with a delay of one week in 2014 at the Karadeniz Technical University of Turkey [22]. The sampling frequency was 256 Hz and a notch filter was used to eliminate noise. The device consisted of 18 electrodes based on the international system 10-20 shown in Fig. 1, in which the two electrodes FP1 and FP2 were related to EOG artifacts, the two electrodes O1 and O2 were associated with EMG artifacts and the remaining 14 electrodes were used in the proposed method for extracting features. Each person sat on a chair facing a $19 \times 19$ screen and was asked to remain motionless during the experiment. Each experiment lasted at least 10 seconds and was performed with a delay of 2 seconds. During the experiment, the target appeared in one of the four possible modes on the display and the person was asked to imagine the appeared mode for 8 seconds; each experiment ended with a beep sound. The data recording was repeated for each person one week later. Each data recording contained 4 runs and each run contained 40 experiments (10 experiments for each class or movement). Therefore, each data recording contained 160 experiments ($40 \times 4 = 160$). The length of data for each experiment was 2048 samples ($8 \times 256 = 2048$). The data were generally divided into two categories of train and test data as shown in Table 1. The number of experiments varied from person to person.

Table 1. The total number of experiments [22].

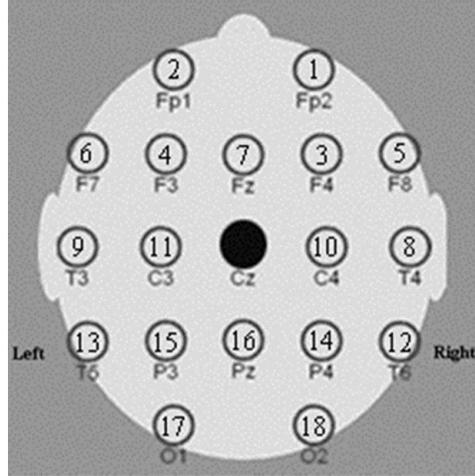| Subject | Train samples | Test samples |
|---------|---------------|--------------|
| A | 140 | 152 |
| B | 148 | 152 |
| C | 148 | 152 |

Fig. 1. The manner of electrodes placement based on the international system 10-20 [22].

## 2.2 Feature Selection

The four following feature extraction methods were implemented to represent the experimental signals as as shown in Table 2.

Table 2. Selected features

| Feature No. | Feature name |
|---|---|
| 1 | Mean & SD of continuous wavelet coefficients in alpha, beta & theta bands |
| 2 | Autoregressive model |
| 3 | Skewness |
| 4 | Mean of derivatives |

### 2.2.1 Continuous wavelet transform

The continuous wavelet transform (CWT) is a strong signal processing tool used as a feature in many BCI data analyses [24, 25]. CWT is a convolution of the signal $X(t)$ with the wavelet function $\psi_{\tau.s}(t)$ so that:

$$CWT_x^{\psi}(\tau,s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t)\psi^*(\frac{t-\tau}{s})dt \tag{1}$$

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{2}$$

Where, $\psi(t)$ is a continuous function in time and frequency, known as the mother wavelet. S and

$\tau$ are the scaling and shifting parameters, respectively.

$$\int_{-\infty}^{\infty} \psi_{\tau,s}(t)\,dt = 0 \tag{3}$$

The mother wavelet applied in this method is the Morlet wavelet transform. The analysis of the statistical features in the train dataset by applying the Morlet wavelet transfer showed that the mean and standard deviation of the continuous wavelet coefficients can be used as feature. The mean and standard deviation as the first feature are calculated in accordance with the following relationships.

$$CWTC^{Avr} = \frac{\sum |CWTC|}{L_{CWTC}} \tag{4}$$

$$CWTC^{std} = \sqrt{\frac{\sum (|CWTC| - CWTC^{Avr})^2}{L_{CWTC} - 1}} \tag{5}$$

Where $L_{CWTC}$ is the length of CWT. The mean and standard deviation of the continuous wavelet coefficients for all channels are calculated as the first feature based on the following three parameters.

1. Based on channel selection; the proposed method includes 14 channels.

2. Based on recording duration, which has been set to 8 seconds in the proposed method.

3. Based on EEG signal bands consisting of three bands: theta (4-8 Hz), alpha (8-13) Hz and beta (13-20 Hz).

### 2.2.2 Autoregressive model

A raw signal $x(t)$, as an autoregressive model, is expressed as follows:

$$x(n) = A_1 x(n-1) + A_2 x(n-2) + ... + A_p x(n-p) + e(n) \tag{6}$$

Where, $A_1. A_2. .... A_p$ are the parameters of the AR model, p is the order of the model and n is an integer representing the discrete time points of the signal. $e(n)$ is the white noise with zero variance and mean.

The suggested AR model is calculated as a second feature based on the following three parameters:

1. Based on channel selection; the proposed method includes 14 channels.

2. Based on recording duration, which has been set to 8 seconds in the proposed method.

3. The order of the model; the model order in the proposed method is $p = 8$.

### 2.2.3 Skewness and the mean of derivatives

Skewness and the mean of derivatives as third and fourth features are calculated based on the following relationships and parameters.

$$S = \frac{\frac{1}{L}\sum_{i=1}^{L}(x_i - \bar{x})^3}{\left(\sqrt{1/L\sum_{i=1}^{L}(x_i - \bar{x})^2}\right)^3} \tag{7}$$

$$AD = \frac{\sum_{n=1}^{L-1}x(n-1) - x(n)}{L-1} \tag{8}$$

1. Based on channel selection; the proposed method includes 14 channels.

2. Based on recording duration, which has been set to 8 seconds in this method and the order of the model, which has been considered $p = 8$.

## 2.3 Description of the proposed method

Figure 2 shows the block diagram of the proposed method, which aims to detect cursor

movement in up- and downward and left- and rightward directions. The proposed method is a hybrid k-means clustering algorithm and LSVM classifier.
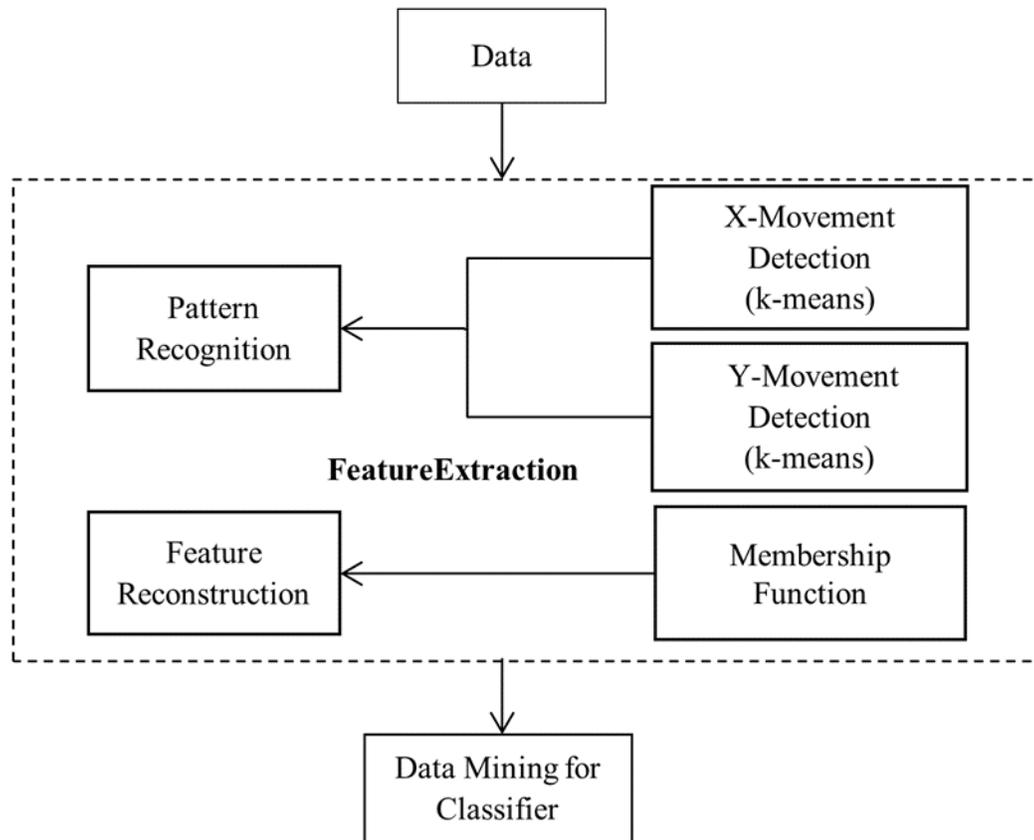


Fig. 2. Block diagram of the proposed method.

The feature extraction is used for patterns identification and the K-means clustering algorithm is used for clustering the movements based on the similarity of the features of the up-, down-, left- and rightward movements considering the membership degree of data. To reduce the big dimensions of the feature space, the algorithm extracts some movement patterns before training of data for proper classification. This method considers a new pattern of k-means algorithm as a new feature of movements. The new features are different from the previous features that only included the pattern of movement features. Therefore, the dimension of feature

space decreases and the new features show the similarity between the movements and patterns. Further, the boundary between up- and downward and left- and rightward movements are estimated by the classifier based on these new features. As the dimensions of the feature space decrease, the dataset is reconstructed with the new features. The optimal channel for each person has been calculated. In this method, once the up- and downward mode and then the left- and rightward mode are classified.

A K-means algorithm can play a role in applying an optimization solution in order to reduce the distance between cluster centers and cluster members, which is calculated according to the relation (9).

$$\min_{\mu_1,\mu_1,\ldots,\mu_k} \sum_{k=1}^{K} \sum_{i \in S_k} \left\| X^i - \mu_k \right\|^2 \tag{9}$$

In the relation (9), k is the cluster index, $S_k$ represents the k $^{th}$ cluster of the dataset, $\mu_k$ indicated the center point of the cluster and k is considered as the total number of clusters, which can contribute to the normalization of the points of data for removing the impact of the scale used for different features. The k-means algorithm adapts the location of the center repeatedly for decreasing the Euclidean distance in order to choose the cluster centers. In this regard, two methods are available for estimating the number of clusters:

1. Based on the experiences of physicians

2. Based on the similarity criterion.

The second method has been used for the similarity of movement patterns. The similarity criteria are calculated according to relations (10) and (11).

$$d_{avg} = \frac{\sum_{k=1}^{K} \sum_{i \in S_k} \sqrt{\sum_{j=1}^{F} (X_j^i - X_j^{\mu_k})^2}}{N} \tag{10}$$

$$d_{\min} = \min\left[\sqrt{\sum_{j=1}^{F}(X_j^{\mu_{k_1}} - X_j^{\mu_{k_2}})^2}\right], \forall k_1 \neq k_2 \tag{11}$$

where $d_{avg}$ shows the average distance from each member $i$ to the center $\mu_k$ in the same cluster $S_k$, $d_{\min}$ indicates the smallest distance between the two clusters, $X_j^i$ is considered as the $j^{th}$ input element of the member $i$. $X_j^{\mu k}$ is regarded as the $j^{th}$ input element of the center $\mu_k$. Accordingly, $N$ and $F$ indicate the total number of data points and the dimension of the input vector, respectively. The number of optimal clusters ($K^*$) is calculated by the function $validity$ ($\theta$) as indicated in relation (12).

$$K^* = \ arg_k \quad \min\theta = arg_k\min\frac{d_{avg}}{d_{\min}} \tag{12}$$

where ($\theta$) represents the validity ratio for assessing the various number of clusters, upon which an acceptable number of clusters is found to recognize the hidden patterns of movement. Based on the results, the average distance from each member to the center of that cluster ($d_{avg}$) decreases by computing the minimum ratio of validity. However, an increase occurs in the minimum distance between two cluster centers ($d_{\min}$). In other words, the determind clusters are concentrated in their own cluster and separated from others. No multiple pattern is shown with a large number of members when the value of k is close to the number of data points. In fact, the cluster density is calculated by the k-means algorithm, which is small and quite similar to the original data. Therefore, the least range of k in the proposed method ($2 < K < 14$) is considered for finding the number of local minima of clusters to show the accumulated patterns of each type of motion. In this case, the magnitude of the scale of the main dataset is reduced by the available labels for each person.

After recognizing the movement patterns, some movements fall into other clusters with symbolic labels. The similarity between the untested movement and symbolic movement can pave the way

for recognition. Thus, a membership function is used to assess the degree of similarity of the main data points and provided motion labels, which indicates the points' fuzzy membership of the recognized patterns. The membership function is as follows:

$$f_c(X_j^i) = \begin{cases} 1 - \dfrac{\left|X^{\mu c} - X_j^i\right|}{\max\left|X_j^{\mu c} - X_j^n\right|} & \text{if } \min(X_j^n) \le X_j^i \le \max(X_j^n),\ \forall n \in S_c \\ 0, & \text{otherwise;} \end{cases} \tag{13}$$

$$p_c = \frac{1}{F}\sum_{j=1}^{F} f_c(X_j^i),\ \ 1 \le c \le K^m + K^b \tag{14}$$

In relation (14), $c$ is the index for the new pattern, $X_j^i$ represents the $j^{th}$ feature of the main input $i$ and $X_j^{\mu c}$ denotes the $j^{th}$ feature of the center $\mu_c$ in the cluster $S_c$, which is obtained from the previous $K$-means algorithm. $K^m$ and $K^b$ are the number of clusters related to the class data.

Using this membership function, the similarity between movements and patterns is assessed from the viewpoint of how well-suited a movement is for pattern recognition.

## 3  Results

The results have been obtained from three subjects by determining the optimal channel for each of them. The proposed method is the hybrid K-means clustering and LSVM classifier.

The $K$-means clustering has been used for recognizing the hidden patterns and the LSVM classifier for separating the up- and downward and the left- and rightward movements. The optimal channel for each person has been calculated according to Table 3. In this method, once the up- and downward mode and then the left- and rightward mode are classified. Based on the results, the LSVM classifier in the proposed method has effective results, as the applied hybrid method has led to an increase of 2-10% in the accuracy of the classifier.

Table 3. The general results

| Subject | Optimal channel | the LSVM classifier accuracy in up- & downward movement hybrid | the LSVM classifier accuracy in up- & downward movement feature | the LSVM classifier accuracy in left- & rightward movement hybrid | the LSVM classifier accuracy in left- & rightward movement feature |
|---|---|---|---|---|---|
| 1 | 12 | 69 ± 3.8 | 67.1 ± 4.9 | 63.1 ± 2.8 | 62.1 ± 3.6 |
| 2 | 14 | 73.3 ± 5.4 | 64.4 ± 4.3 | 68.6 ± 3.6 | 60.9 ± 3.2 |
| 3 | 12 | 66.3 ± 4.3 | 62.1 ± 3.3 | 74.4 ± 5.8 | 65.5 ± 2.8 |

Sensitivity and specificity are two statistical measures for evaluating the outcome of a binary classification test. When data can be divided into two positive and negative groups, the accuracy of the results of a test that divides the data into these two groups can be measured and described using the indices of sensitivity and specificity. Sensitivity (the true positive rate) is the proportion of positives that are correctly identified as positive by the test. Specificity (the true negative rate) is the proportion of negatives that are correctly identified as negative by the test. Mathematically, sensitivity is expressed as the ratio of true positives to combined true and false positives.

$$sensitivity = \frac{number \quad of \quad true \quad positives}{number \quad of \quad true \quad positives + number \quad of \quad false \quad negtives} \qquad (15)$$

$$specificity = \frac{number \quad of \quad true \quad negatives}{number \quad of \quad true \quad negatives + number \quad of \quad false \quad positives} \qquad (16)$$

Sensitivity and specificity for up- and downward and left- and rightward movements in the proposed method are calculated based on the relations (15) and (16), as seen in Table 4.

Table 4. Sensitivity and specificity amounts

| Subject | Movement | Sensitivity | Specificity |
|---|---|---|---|
| 1 | up- and downward | 65% | 72% |
| | left- and rightward | 63% | 62% |
| 2 | up- and downward | 74% | 71% |
| | left- and rightward | 69% | 67% |
| 3 | up- and downward | 67% | 65% |
| | left- and rightward | 71% | 79% |

## 4  Discussion

In this study, the K-LSVM method was proposed based on known special patterns. This can be comparable to conventional data mining methods in motion detection. In the feature extraction step, conventional data mining methods are not used; instead, clustering is utilized to extract symbolic motions with the aim of representing the cluster of motions. The relationship between input motion and symbolic motion as the pattern membership is evaluated to predict new motions. The proposed method is comparable to Aydemir et al. [21] with same dataset and features. They classified the EEG signal based on the decision tree structure during the cursor movement. They could separate the up- and downward and left- and rightward movements by achieving the accuracy of 55.92%, 57.92% and 80.24% for the first, second and third person, respectively. As the feature applied in their approach is according to table 2 and the three classifiers LDA, KNN and SVM have been used for separating the movements, the applied decision tree structure is different in each person and this is one of the weak points of this approach, while the hybrid algorithm applied in the proposed method in the present study, considering the use of the same features, is similar for all subjects, and hence, the proposed method has caused an increase of 2-10% in the classifier accuracy. There are no researches on hybrid clustering and classifiers in BCI systems and recent studies using this approach have been concentrated on breast cancer diagnosis. Bennett K-P et al. [23] tested this algorithm on a set of breast cancer data for clustering the benign and malignant tumors and suggested the detection of data patterns as one of the applications of clustering algorithm. Since the method used by the researchers did not focus on the manner of determining the quantity of clusters in datasets, the test clusters were randomly divided into two benign and malignant categories. The number of patterns should be determined to evaluate the hidden patterns of breast cancer. There is also a need for

comprehensive exploration of the connection between the supervised learning and learning without supervision to improve the breast cancer diagnosis. In order to reduce the dimensions of the dataset, some researchers began to use the technique of clustering algorithms and classifier model in the field of machine learning. Dhillon I-S et al. [27] applied the hybrid clustering algorithm for a set of similar words to achieve a quick and accurate learning in text classification. They used C-means clustering for educational samples and SVM for Text identification. Radha et al. [28] used a breast cancer database for their researches in 2014. Their clustering was based on motor size. They applied the K-means algorithm to this dataset. Nine features were used in this method. In the first phase of the K-means clustering algorithm, the Manhattan distance criterion was used to measure the distance between nodes and the selection of the center of the primary cluster was done randomly. After implementing the algorithm at k = 2 on 286 data points, the average probability of the correct classification was 66.99%. In the second phase of K-means clustering, the Euclidean distance criterion was used to measure the distance between nodes and the selection of the center of the primary cluster was done randomly. After implementing the algorithm at k = 2 on 286 data points, the average probability of the correct classification was 74.98%. In the third phase of K-means clustering, the cluster centers were graded in order to raise the clustering accuracy and the Pearson's correlation criterion was used to measure the distance between nodes and the selection of the centers of the primary cluster was achieved through co integration. After implementing the algorithm at k = 2 on 286 data points, the same results were obtained from all clustering. The average probability of the correct classification was 100%. Based on the above-mentioned issues, a large number of methods have been utilized for diagnosing breast cancer based on implementing clustering and classifier. Having a large scale of data is considered as one of the main disadvantages of conventional and traditional methods. Although

the time calculations in different training are decreased by applying explorations in feature selection, no comprehensive search is necessary for selecting the feature for the clustering algorithm, especially the K-means algorithm, which results in better understanding of the number of training samples without any help for extracting and selecting the related feature. In the present study, a combination of K-means clustering algorithm and classifier were implemented to compress the feature extraction space in order to decrease the computational costs. The proposed K-LSVM method in this study emphasizes on discovering the knowledge framework, searching for more knowledge of motion detection through implementing data mining methods. In addition, descriptive features and data will be collected in the future. However, feature extraction and selection are considered as a challenging issue for researchers, although large data with lost values can cause another issue which should be taken into consideration with respect to computational time.

## Conclusions

In the present study, K-means was used as clustering algorithm in order to recognize and obtain motion patterns. These patterns are reconstructed as a new feature for the training phase. Pre-processing steps such as feature extraction and selection can play a role in creating a highly effective feature for the purpose of machine learning algorithms which is used for training the classifier. Regarding K-LSVM results, the calculation time can be significantly decreased without measuring the detection accuracy. However, using the k-means algorithm failed to decrease the size of training samples by filtering similar samples in the data, which can be regarded as a potential method for reducing the size of the training set. Based on the results, the LSVM classifier in the proposed method has effective results, as the applied hybrid method has led to an increase of

2-10% in the accuracy of the classifier. The dataset used in this study represents a complete dataset with lost values; however, applying this method to a broad set of low-size data opens a venue of research in the future. So far, no general rule has been adopted for evaluating the number of patterns for detecting motion. In the future, a method for evaluating the number of symbolic motions can be developed, which may decrease the calculation time for extracting feature and allows physicians to comprehend the motions better based on the patterns derived from the methods implemented for data mining methods.

**Declarations**

**Availability of data and materials**

Available.

**Competing interests**

The authors declare that they have no conflict of interest.

**Funding information**

This work did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Authors' contributions**

All Authors contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

**Compliance with ethical standards**

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the Karadeniz Technical university and its later amendments or comparable ethical standards.

**Informed consent** was obtained from all individual participants included in the study.

**Acknowledgements**

Not applicable

# References

1. Mellinger, J., Schalk, G., Braun, C., Preissl, H., Rosenstiel, W., Birbaumer, N., & Kübler, A. (2007). An MEG-based brain–computer interface (BCI). *Neuroimage*, *36*(3), 581-593..
2. Dadgostar, M., Setarehdan, S.K. and Akin, A., 2013, December. Detection of motion artifacts in fNIRS via the continuous wavelet transform. In 2013 20th Iranian Conference on Biomedical Engineering (ICBME) (pp. 243-246). IEEE.
3. Shirzadi, S., Einalou, Z., & Dadgostar, M. (2020). Investigation of Functional Connectivity During Working Memory Task and Hemispheric Lateralization in Left-and Right-Handers Measured by fNIRS. Optik, 221, 165347.
4. Dadgostar, M., Setarehdan, S. K., Shahzadi, S., & Akin, A. (2016). Functional connectivity of the PFC via partial correlation. *Optik-International Journal for Light and Electron Optics*, *127*(11), 4748-4754.
5. Einalou, Z., Maghooli, K., Setarehdan, S. K., & Akin, A. (2015). Functional near infrared spectroscopy for functional connectivity during Stroop test via mutual information. *Adv. Biores.*, *6*(1), 62-67.
6. Einalou, Z., Maghooli, K., Setarehdan, S. K., & Akin, A. (2017). Graph theoretical approach to functional connectivity in prefrontal cortex via fNIRS. *Neurophotonics*, *4*(4), 041407.
7. Shirzadi, S., Einalou, Z., & Dadgostar, M. (2020). Investigation of Functional Connectivity During Working Memory Task and Hemispheric Lateralization in Left-and Right-Handers Measured by fNIRS. *Optik*, *221*, 165347.
8. Weiskopf, N., Mathiak, K., Bock, S. W., Scharnowski, F., Veit, R., Grodd, W., ... & Birbaumer, N. (2004). Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE transactions on biomedical engineering*, *51*(6), 966-970.
9. Faskhodi, M. M., Einalou, Z., & Dadgostar, M. (2018). Diagnosis of Alzheimer's disease using resting-state fMRI and graph theory. Technology and Health Care, 26(6), 921-931.Deiber, M. P., Ibanez, V., Honda, M., Sadato, N., Raman, R., & Hallett, M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *Neuroimage*, *7*(2), 73-85.
10. Aydemir, O., & Kayikcioglu, T. (2011). Wavelet transform based classification of invasive brain computer interface data. *Radioengineering*, *20*(1), 31-38.
11. Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, *12*(2), 1211-1279.
12. Aydemir, O., & Kayikcioglu, T. (2013). Comparing common machine learning classifiers in low-dimensional feature vectors for brain computer interface applications. *International Journal of Innovative Computing, Information and Control*, *9*(3), 1145-1157.
13. Kayikcioglu, T., & Aydemir, O. (2010). A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data. *Pattern Recognition Letters*, *31*(11), 1207-1215.
14. Piccione, F., Giorgi, F., Tonin, P., Priftis, K., Giove, S., Silvoni, S., ... & Beverina, F. (2006). P300-based brain computer interface: reliability and performance in healthy and paralysed participants. *Clinical neurophysiology*, *117*(3), 531-537.

15. Gao, W., Guan, J. A., Gao, J., & Zhou, D. (2015). Multi-ganglion ANN based feature learning with application to P300-BCI signal classification. *Biomedical Signal Processing and Control*, *18*, 127-137.
16. Wang, M., Daly, I., Allison, B. Z., Jin, J., Zhang, Y., Chen, L., & Wang, X. (2015). A new hybrid BCI paradigm based on P300 and SSVEP. *Journal of Neuroscience Methods*, *244*, 16-25.
17. Royer, A. S., Doud, A. J., Rose, M. L., & He, B. (2010). EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies. *IEEE Transactions on neural systems and rehabilitation engineering*, *18*(6), 581-589.

18. Volosyak, I., Gembler, F., & Stawicki, P. (2017). Age-related differences in SSVEP-based BCI performance. *Neurocomputing*, *250*, 57-64.
19. Heidari, H., & Einalou, Z. (2017). SSVEP Extraction Applying Wavelet Transform and Decision Tree With Bays Classification. *International Clinical Neuroscience Journal*, *4*(3), 91-97.
20. Rouillard, J., Duprès, A., Cabestaing, F., Leclercq, S., Bekaert, M. H., Piau, C., ... & Lecocq, C. (2015). Hybrid BCI coupling EEG and EMG for severe motor disabilities. *Procedia Manufacturing*, *3*, 29-36.
21. Einalou, Z., Maghooli, K., Setarehdan, S. K., & Akin, A. (2016). Effective channels in classification and functional connectivity pattern of prefrontal cortex by functional near infrared spectroscopy signals. *Optik-International Journal for Light and Electron Optics*, *127*(6), 3271-3275.
22. Aydemir, O., & Kayikcioglu, T. (2014). Decision tree structure based classification of EEG signals recorded during two dimensional cursor movement imagery. *Journal of neuroscience methods*, *229*, 68-75.
23. Bennett, K. P., & Blue, J. A. (1998, May). A support vector machine approach to decision trees. In *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on* (Vol. 3, pp. 2396-2401). IEEE.
24. Einalou, Z., & Maghooli, K. (2010, November). Fuzzy neural network approach for noninvasive diagnosis of digestive diseases using wavelet comparing to classification followed by fuzzy C-mean algorithm. In *Biomedical Engineering (ICBME), 2010 17th Iranian Conference of* (pp. 1-4). IEEE.
25. Einalou Z, Mehran YZ, Maghuly K. Diagnosis and classification of digestive diseases by wavelet transform. InFirst Joint Congress on Fuzzy and Intelligent Systems (ISFS2007) 2007.
26. Dadgostar M, Setarehdan SK, Shahzadi S, Akin A. CLASSIFICATION OF SCHIZOPHRENIA USING SVM VIA fNIRS. Biomedical Engineering: Applications, Basis and Communications. 2018 Apr;30(02):1850008.
27. Dhillon, I. S., Mallela, S., & Kumar, R. (2003). A divisive information-theoretic feature clustering algorithm for text classification. *Journal of machine learning research*, *3*(Mar), 1265-1287.
28. Radha, R., & Rajendiran, P. (2014, February). Using K-Means Clustering Technique to Study of Breast Cancer. In *Computing and Communication Technologies (WCCCT), 2014 World Congress on* (pp. 211-214). IEEE.    27

29. Guger, C., Ramoser, H., Pfurtscheller, G., 2000. Real-time EEG analysis with subject specific spatial patterns for a brain–computer interface (BCI). IEEE Trans. Neural Syst. Rehabil. Eng. 8, 447–456.
30. Donchin, E., Spencer, K.M., Wijesinghe, R., 2000. The mental prosthesis: assessing the speed of a P300-based brain–computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. 8, 174–179.
31. Hoffmann, U., Vesin, J.M., Ebrahimi, T., Diserens, K., 2008. An efficient P300-based brain–computer interface for disabled subjects. J. Neurosci. Methods 167, 115–125.
32. Middendorf, M., McMillan, G., Calhoun, G., Jones, S.K., 2000. Brain–computer interfaces based on the steady-state visual-evoked response. IEEE Trans. Rehab. Eng. 8 (2), 211–214.
33. Mensh, B.D., Werfel, J., Seung, H.S., 2004. BCI competition 2003—Data set Ia: Combining gamma-band power with slow cortical potentials to improve singletrial classification of 51, 1052–1056.
34. Burke, D.P., Kelly, S.P., Chazal, P., Reilly, R.B., Finucane, C., 2005. A parametric feature extraction and classification strategy for brain–computer interfacing. IEEE Trans. Neural Syst. Rehabil. Eng. 13 (1), 12–17.
35. Lal, T.N., Schröder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N Schölkopf, B., 2004. Support vector channel selection in BCI. IEEE Trans. Biomed Eng. 51 (6), 1003–1010.
36. McFarland, D.J., Sarnacki, W.A., Wolpaw, J.R., 2003. Brain–computer interface (BCI)operation: Optimizing information transfer rates. Biol. Psychol. 63, 237–251.
37. Qin, L., He, B., 2005. A wavelet-based time–frequency analysis approach for classification of motor imagery for brain–computer interface applications. J. Neural Eng. 2, 65–72.
38. Wu, W., Gao, X., Hong, B., Gao, S., 2008. Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). IEEE Trans. Neural Syst. Rehabil. Eng. 55 (6), 1733–1743.
39. Schlogl, A., Lee, F., Bischof, H., Pfurtscheller, G., 2005. Characterization of four-class motor imagery EEG data for the BCI-competition. J. Neural Eng. 2, L14–22.
40. Obermaier, B., Guger, C., Neuper, C., Pfurtscheller, G., 2001. Hidden Markov models for online classification of single trial EEG data. Pattern Recognit. Lett. 22, 1299–1309.
41. Pfurtscheller, G., da Silva, F.H.L., 1999. Event-related EEG/MEG synchronization and desynchronization: Basic principles. Clin. Neurophysiol. 110, 1842–1857.
42. Lotte, F., Congedo, M., Lecuyer, A., Lamarche, F., Arnaldi, B., 2007. A review of classification algorithms for EEG-based brain–computer interfaces. J. Neural Eng. 4, R1–R13.
43. Einalou, Z., Najafi, Z., Maghooli, K., Zandi, Y., & Sheibeigi, A. (2007). Long-term simulation of digestive sound signals by CEPSTRAL technique. Int J Appl Math Comput Sci, 4(2), 844-8.
44. Wang, B., Jun, L., Bai, J., Peng, L., Li, Y., Li, G., 2005. EEG recognition based on multiple types of information ] using wavelet packet transform and neural networks. In: Proc. 2005 IEEE Engineering in Medicine and Biology 27th Annual Conf. Shanghai, China, pp. 5377–5380.
45. Sun, S., Zhang, C., 2005. Assessing features for electroencephalographic signal categorization. Acoust. Speech Signal Process. 5, 417–420.

46. Trejo, L. J., Rosipal, R., & Matthews, B. (2006). Brain-computer interfaces for 1-D and 2-D cursor control: designs using volitional control of the EEG spectrum or steady-state visual evoked potentials. IEEE transactions on neural systems and rehabilitation engineering, 14(2), 225-229.
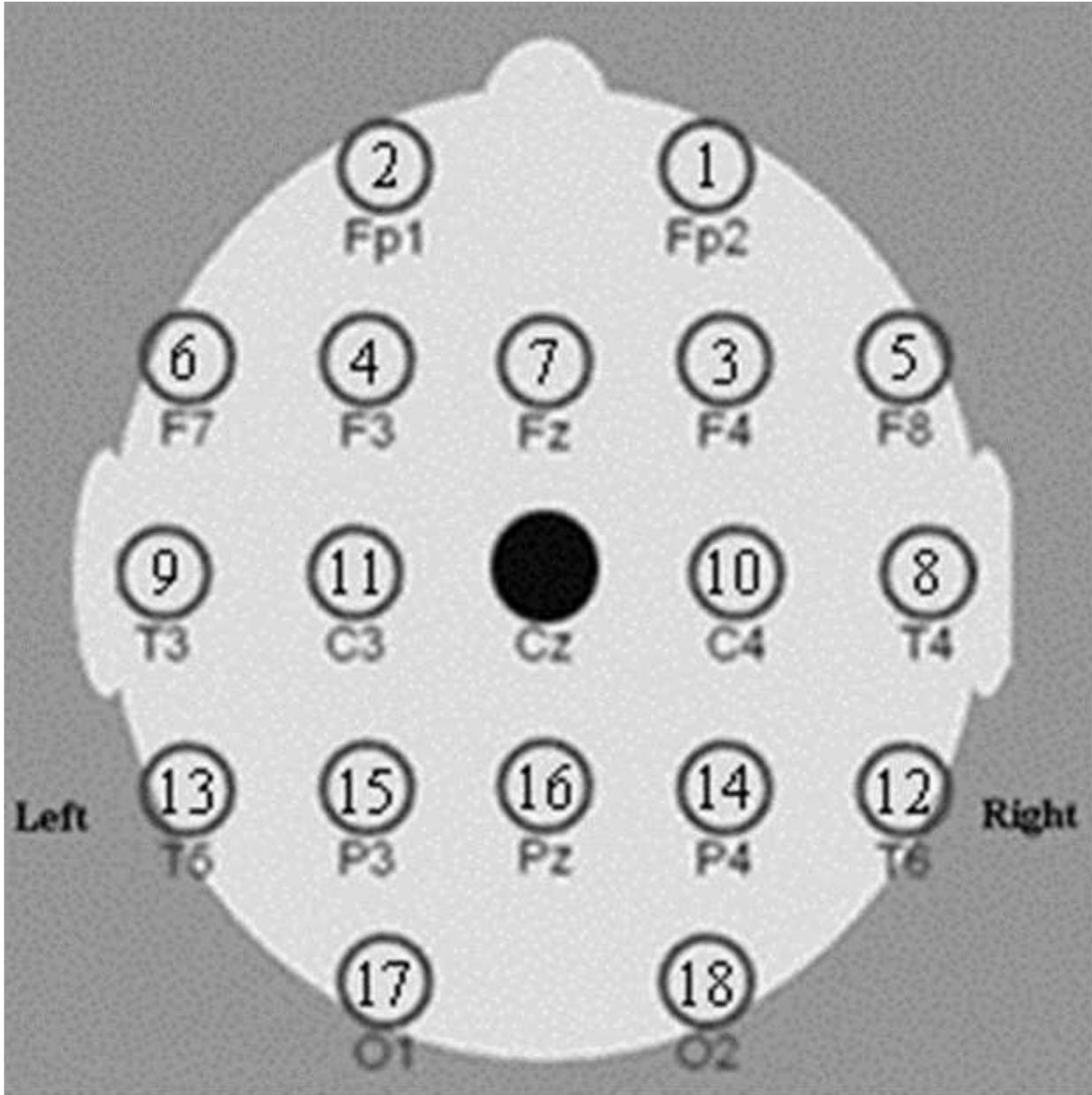
# Figures



**Figure 1**

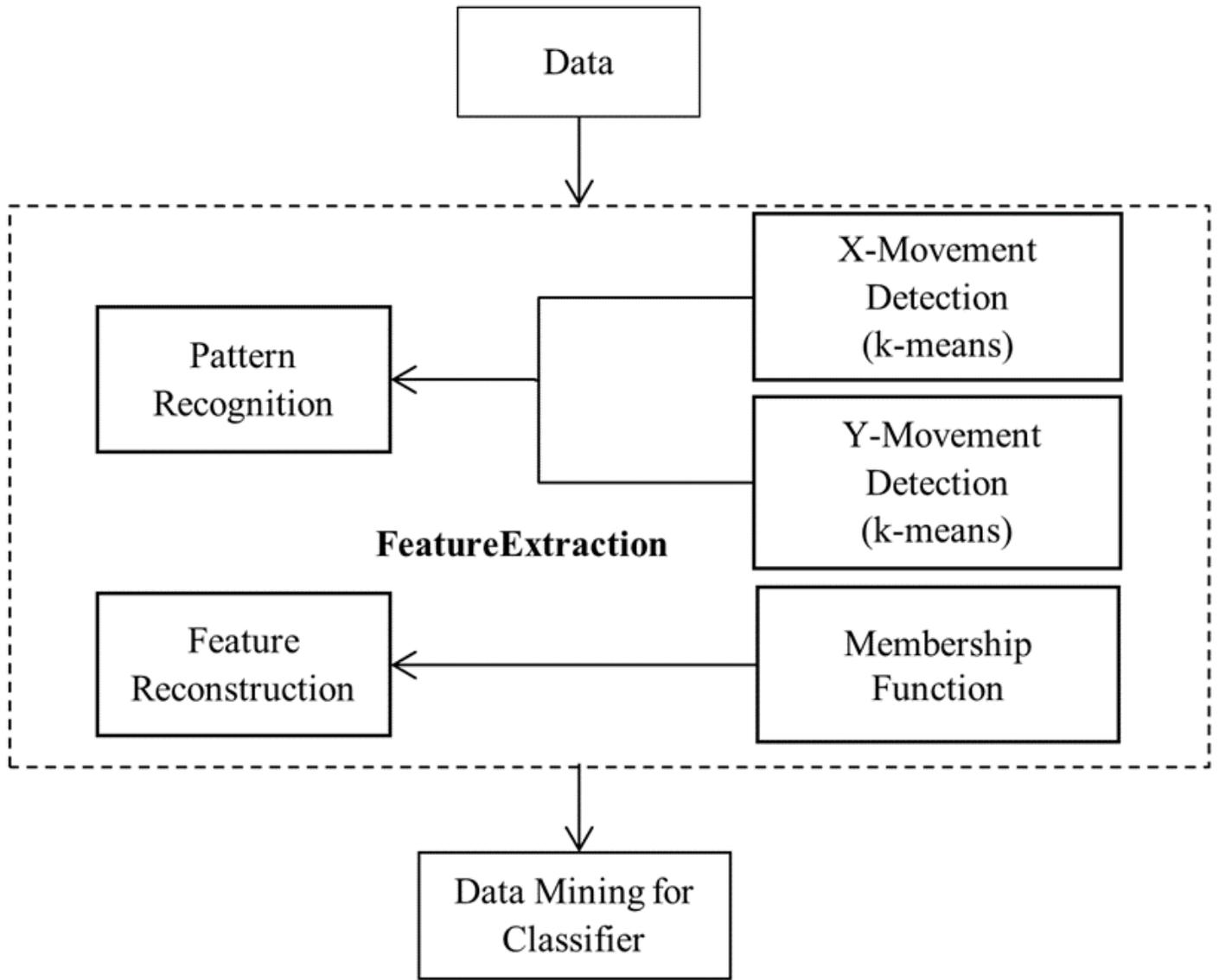The manner of electrodes placement based on the international system 10-20 [22].

Figure 2

Block diagram of the proposed method