www.ijecs.in
International Journal Of Engineering And Computer Science ISSN: 2319-7242
Volume 4 Issue 9 Sep 2015, Page No. 14429-14435

Comparative Analysis of performance of K-means algorithm for skin detection using wearable Sensors

Anshul Sharma[1], Varun Sharma[2], Anil Saroliya[3]

[1,2,3]Amity School of Engineering & Technology

Amity University Rajasthan, India

[1]ansh.zeuss@gmail.com, [2]vsharma@jpr.amity.edu, [3]asaroliya@jpr.amity.edu

Abstract: The main idea of this project to provide an alternative method of analysis of K-means algorithm using wearable sensors. The presented work is a give the step of classification using K-means algorithm of data. It is well known classification approach for clustering in data representation. The analysis verifies and motivates the presented method which increases the advantage for intensity independent activity recognition. It is also present the design, implementation and testing of the K-means algorithm through complete experimentation by using wearable sensors. Finally performance analysis of the K-means algorithm on skin detection as well as features extracted discuss in detail.

Introduction: The K-means algorithms used as Classifier for skin detection in health monitoring to reduce the loss of the lives specifically for patients need to consider the related issues of its acceptability by the patients. In this paper the comparative analysis of the performance of K-means classifier based on the performance of the basic K-means with modified k-means algorithm skin detection for health monitoring using skin conductance signal wearable computing parameter[4]. It also provides the details of data partitioning, methodology adopted for analysis, design & implementation of algorithms and tests results for individual version of three K-Means algorithms and the comparative analysis of all the algorithms for specific case of skin detection data. The Skin detection is a very important pre-processing step in many different applications, such as face detection, people detection, image content interpretation, de-identification for privacy protection in multimedia content. There is no a perfect solution for skin detection, since this process is a compromise on speed, simplicity and precision (detection quality). It is not possible to achieve all three aspects in the same time. If it is fast and simple it is imprecise (large number of false positive detections). The skin conductance is used to determine the electrical conductance of the skin. This is control and dependent on the value of sweat induced skin moisture. The skin detection are represented by different name that are Galvanic Skin Response (GSR), Electro dermal Response (EDR), Psycho Galvanic Reflex (PGR), Skin Conductance Response (SCR) and Skin Conductance Level (SCL). In most applications certain level of precision and speed is sufficient and that determines complexity of a model. Health monitoring means monitoring of a person to identify changes in the person's health status because of exposure to certain substances. Health monitoring must be carried out by or be done under the supervision of a registered medical practitioner with experience in health monitoring. Health monitoring does not include air monitoring or other measures used to assess or control exposure to hazardous chemicals in the workplace. Health monitoring must never be used as an alternative to putting in place effective control measures. The exact difference between the galvanic skin resistance and galvanic skin potential is represented the term of Galvanic Skin Response (GSR). The GSR identify as a recorded electrical resistance between two electrodes when a fragile current passed between them. The electrodes are sited about the inch separately, and recorded resistance is modified with respect to emotional state of the user. The Galvanic skin potential represented to the computed voltage between two electrodes without any externally applied current. The compound modification between galvanic skin resistance and galvanic skin potential make up the galvanic skin response.

Data used: In this section the details of data used, features extracted and performance measures used for performance analysis of K-mean algorithms as classifier has been discussed. In order to detect the skin of skin cancer patient, real time physiological sensory data related to post state(cancer state) and normal state, collected and used for further processing. The data recorded was 2-3 minutes recording of Normal state patient and Post state of patient and stored in text file format. We selected some of the signals of the skin conductance wearable parameter for further processing in this paper. The two state signals of skin conductance were processed for extraction of predefined features using MATLAB. Various sets of features were extracted over full length of the signal collected. Two different kinds of sets were extracted from Skin Conductance signal such as Feature Set I and Feature Set II (wavelet features: File Name SC_DB3_L6). The details of feature vectors and their size is mentioned. The feature vectors of SC signals for two class data **Pre state** and **Post state** were used as input to the algorithms. Each column of the feature vector was treated as separate input to the algorithms and consequently the decision of state was also derived independently. Fifty percent of the selected data has been used as training and remaining fifty percent as test dataset.

Table 1: Details of Input Parameters

Data Signal Used	File Name	No. of Feature Parameters	Size of Training dataset	Size of Testing dataset
Skin Conductance	SC Features Set_1	6	400 rows of 6 columns	400 rows of 6 columns
	SC Feature Set_DB3_L6	77	400 rows of 77 columns	400 rows of 77 columns

Proposed Approach:

5.3.1 Algorithm to compute the Basic k-means (Version 1) work as Classifier:

Step 1: Read the Two class data (de Chile (Pre, Post)) from Excel data file.

Step 2: Find the number of rows and number of columns.

Step3: Initially store the dataset i=0 to rows/3 and j=0 to number of columns into 2-D array along with row tag.

Step 4: For i=0 to rows/3 & For j= 0 to (Number of Columns-1).

Step5: Assign first two values of each column to corresponding two class cluster centers.

Step6: Read and Calculate the distance between next data point in each column to the corresponding cluster centers.

Step7: Assign data point to the cluster from which the distance is minimum for each column.

Step 8: Update value of each cluster centers of corresponding columns.

Step9: Recalculate the distance between previously assigned data point and the new cluster centers of each column.

Step10: If no data point of each column was reassigned then stop otherwise go to step6 for corresponding column.

Step 11: Then store the dataset i=rows/3+ 1 to 2*rows/3 and j=0 to columns into 2-D array along with row tag.

Step 12: Assign each of the data point in each column to corresponding cluster on the basis of minimum distance with the current cluster centers.

Step 13: Prepare confusion matrix of test data for dataset i=rows/3+ 1 to 2*rows/3 and j=0 to column & write to output file.

Step 14: Then store the dataset i=2*rows/3+ 1 to 3*rows/3 and j=0 to columns into 2-D array along with row tag.

Step 15: Assign each of the data point in each column to corresponding cluster on the basis of minimum distance with the current cluster centers.

Step 16: Prepare confusion matrix of test data for dataset i=2*rows/3+ 1 to 3*rows/3 and j=0 to column & write to output file.

Modified K-means algorithm: The basic difference in the concept from Version 1 is that here that the data rows are treated as object arrays rather than individual columns.

Algorithm for Modified K-means Algorithm (Version 2) as Classifier:

Step 1: Read the Two class data (de Chile (Pre, Post)) from Excel data file.
Step 2: Find the number of rows.
Step3: Initially make a Classification class which contain numeric attribute along with string tag as a member variables.
Step 4: Store the dataset i=0 to rows/3 into 1-D object array of Classification class.
Step 5: For i=0 to rows/3.
Step6: Assign first two values of 1-D object array of classification class to corresponding class 1-D cluster centers array of classification class to store the 2-cluster center values.
Step7: Read and Calculate the distance between next data element in 1-D object array to the corresponding cluster centers values in 1-D cluster center array.
Step8: Assign data element to the cluster from which the distance is minimum.
Step 9: Update value of each cluster centers.
Step10: Recalculate the distance between previously assigned data point of 1-D object array and the new 1-D cluster centers of array.
Step11: If no data point of 1-D object array was reassigned then stop otherwise go to step 6.
Step 12: Then store the dataset i=rows/3+ 1 to 2*rows/3 into 1-D object array along with row tag.
Step 13: Assign each of the data element of 1-D object array to corresponding cluster on the basis of minimum distance with the 1-D current cluster centers object array.
Step 14: Prepare confusion matrix of test data for dataset i=rows/3+ 1 to 2*rows/3 & write to output file.
Step 15: Then store the dataset i=2*rows/3+ 1 to 3*rows/3 into 1-D array along with row tag.
Step 16: Assign each of the data element of 1-D object array to corresponding cluster on the basis of minimum distance with the element of 1-D current cluster centers object array.
Step 17: Prepare confusion matrix of test data for dataset i=2*rows/3+ 1 to 3*rows/3 & write to output file.

Comparative Results of K-means algorithm for skin conductance features:

DOI: 10.18535/ijecs/v4i9.63

A. *Feature Set-I of Skin Conductance:* Total six parameters were used as input feature set (800x6) including Training data (400x6) and Test data (400x6) for the classification. The major performance measure of the classifier is Classification Accuracy. The performance of the classifier for MOS has been 100 % and 100 % for both Training and test data of Pre driving class and Post driving class. For FE, MOFS, STDFS, and SDOS performed well in Post condition case while poor performance in Pre-condition case. All other remaining did not perform well in either case, hence not included in table. It can further be observed that the average classification accuracy has been above 80 % for FE and MOS for both Training and Test datasets.

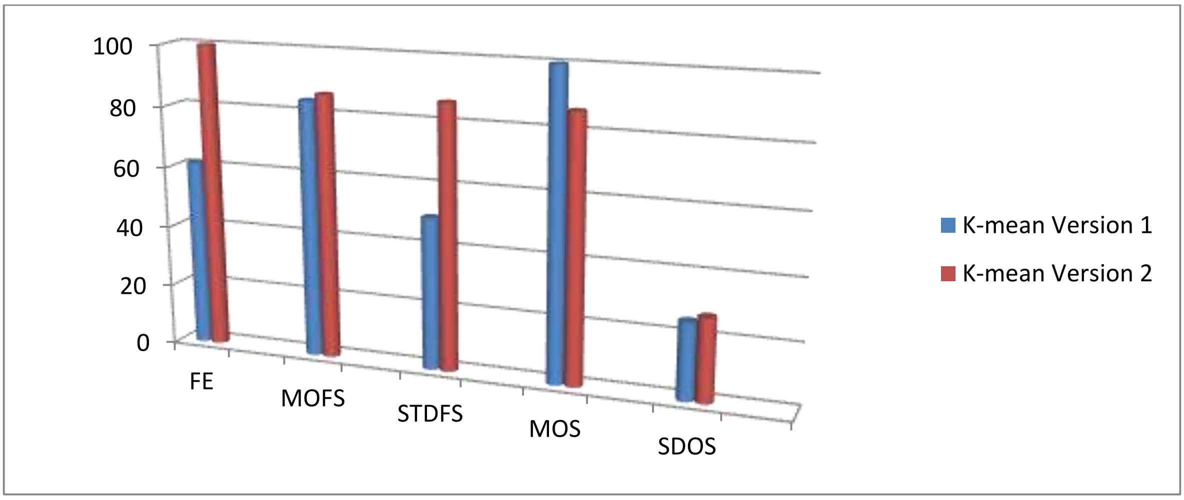

Fig 1: SC Feature Set I versus Performance classification accuracy for K-means 1 to 2

B. **Feature Sets-II of Skin Conductance (SC_DB3_L6) Dabechies Wavelet feature sets (SC_DB3_L6) of skin conductance:** Total 77 features were used using wavelet features extracted up level six of size 800x77 including 400x77 as training and 400x77 as testing data. The test results of the classifier for Training data and Test dataset are shown in Table 5.3.3 that includes results for those features only which fetched classification accuracy above 50 % in either class. It can be observed from the table that the features MAX, MIN, MEAN & MODE of CA6 and A6 i. e. decomposed approximate coefficient and reconstructed approximate coefficient along with Entropy parameters fetched 100 % classification accuracy in either classes for both training as well as test dataset. So these parameters could strongly correlate to the cognitive fatigue state with basic K-means algorithm. All other parameters in the table could perform well for one of the class only except MODE D1 which could give above 80 %.

DOI: 10.18535/ijecs/v4i9.63

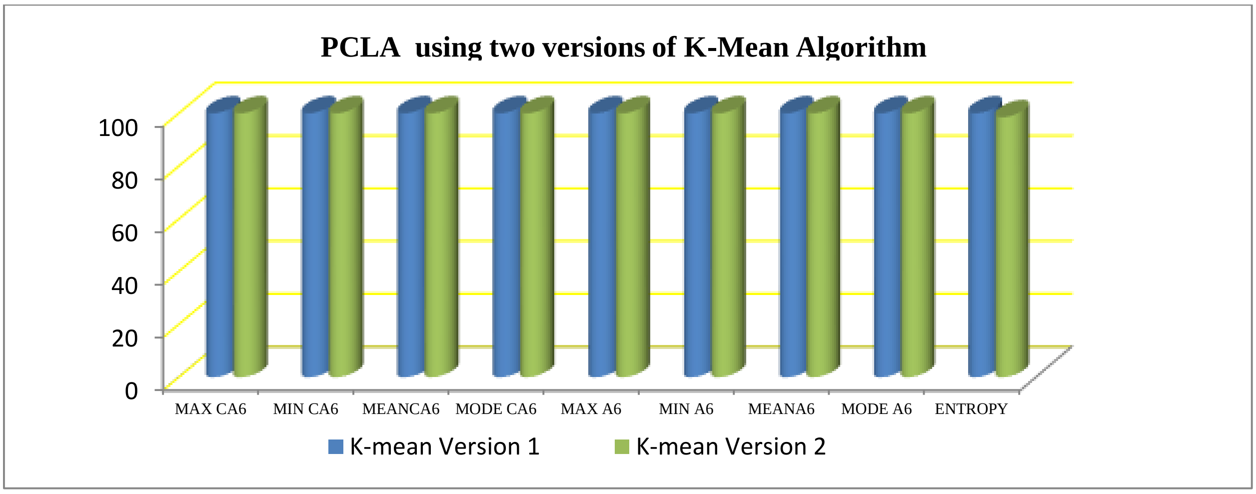

Fig2 (A): SC Feature Set II versus Performance classification accuracy for K-means 1 to 2

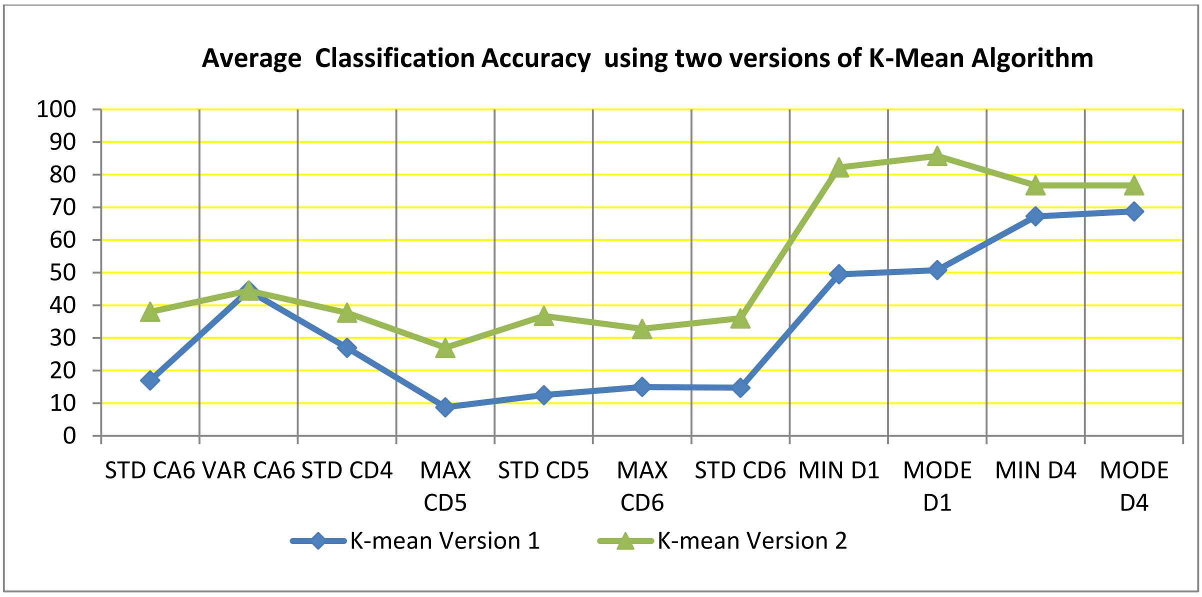

Fig2 (B): SC Feature Set II versus Performance classification accuracy for K-means 1 to 2

Optimal Results of Two Versions of K-means for Performing Parameters :

Here **Table 2** shows the summary of various features those performed the best in each class for training as well as test datasets in corresponding algorithms.

Table 2: Optimal Performing Skin conductance Parameters / Features & Algorithms

Physiological Parameters / Algorithms	**Feature File**	**Total Features**	**K-Means Version 1**	**K-Means Version 2**
	Feature Set I	6	**MOS**	**FE, MOFS, STDFS, MOS**
	Feature Set II-SC_DB3_L6	**77**	**MAX, MIN, MEAN & MODE of CA6 and A6, Entropy (9 Features)**	**MAX, MIN, MEAN & MODE of CA6 & A6 and Entropy , MIN D1, MODE D1, MIN D4, MODE D4 (13 Features)**

The K-means version 2 (Modified K-means) could perform very well compare to the basic version of K-means -1 for some of the features of all the physiological parameters such as Skin Conductance. Although K-1 & K2 performed well for some of the features, the test has been on individual features rather than in combination and hence these features only can be used for detection of skin conductance as the classification accuracies for both the classes has been up to 100 %. For Skin Conductance (SC) Feature set II, MOS feature performed well with K-means Version 1 while FE, MOS, MOFS; STDFS worked well in K-means version 2. In SC feature set II which was based on Daubechies order 3 and level 6 decomposition, approximate coefficient of decomposed as well as reconstructed i.e. CA6 & A6 and the entropy features performed the best with K-means version 1 and 2.

Conclusion: The selected physiological signals of two classes Pre condition and Post condition were considered for testing the performance of these algorithms. Various statistical and wavelet features including Daubechies wavelet function, were derived from the text files of Skin conductance of these two classes. Two kinds of feature sets were derived from SC files. These feature files were used as input and the performance of each feature individually has been checked through two versions of K-means algorithms in order to find out set of features those can work the best and as alternative solution and the best performing version of K-means. Percentage classification accuracy has been used as performance measures to test the performance of the classifiers. The test results proved that some of the features of all the parameters individually could be classified to the extent of 100 % by K-means version 1 and 2.

References:

[1] AlexandrosPantelopoulos and Nikolaos G. Bourbakis,"A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis", IEEE Transactions on Systems, Man, and Cybernetics, Vol. 40, No. 1, January 2010.

[2] A. Behrad, M. Salehpour, M. Saiedi, and M. N. Barati, "Obscene Video Recognition Using Fuzzy SVM and New Sets of Features", International Journal of Advanced Robotic Systems, 2013, Vol. 10, 99:2013.

[3] COST, "De-identification for privacy protection in multimedia content", [Online]. Available: http://www.cost.eu/domains_actions/ict/Actions /IC1206

[4] J. Ruiz-del-Solar and R. Verschae, "Skin Detection using Neighborhood Information", Proceedings of the 6th International Conference on Automatic Face and Gesture Recognition (FG2004), Seoul, Korea, 2004, pp. 463-468.

[5] Roopesh S O, Appaji M Abhishek, Dr H N Suma, "Prognostic Health Monitoring System", Volume 3, Issue 6, June 2013 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering

[6] M. Priyadharshini Padma, Mr.G.Komorasamy, "A Modified Algorithm For Clustering Based On Particle Swarm Optimization And K-Means", 2012 International Conference on Computer Communication and Informatics (*ICCCI* -2012), Jan. 10 – 12, 2012, Coimbatore, INDIA, pp. 978-1-4577-1583-9.

[7] Prajesh P Anchalia, Anjan K Koundinya, Srinath N K, “MapReduce Design of K-Means Clustering Algorithm” IEEE International Conference 2013 PP. 978-1-4799-0604-8.

[8] Quan Shi, Lu Xu, Zhen Quan Shi, yijun chen, yeqin shao,” Analysis and Research of campus network user behavior based on K-means Clustering”, 2013 fourth international conference on Digital Manufacturing & Automation PP. 978-0-7695-5016-9.

[9] Sardha Sangewar, Atish A. Pattishwar, Vilas Alagdeve, Rupali Balpande, “Liver Segmentation of CT scan images using K-means Algorithm”, 2013 International Conference on Advanced Electronics System (ICAES) PP. 978-1-4799-1441-8.

[10] Shanshu Wang, Shaosheng Fan, “Prediction of Fouling In Condenser Based On Kmeans Algorithms and Improved Chebyshev Neural Network”, Automatic Control and Artificial Intelligence (ACAI 2012), IET International Conference march 3rd to 5th 2012 pp 978-1-84919-537-9.

[11] SHEN-TU Han, XUE An Ke, YANG Jian Bo, “K-means Entropy Algorithm for Multiple Model Estimation”, Proceeding of the 32 Chinese control conference July 26th to 28th 2013 pp 61174024.

[12] Tingting Cui, Fangshi Li “Weight Computing in Competitive K-Means Algorithm”, “Vehicle Classification for single loop detector with Neural Genetic Controller: A Design Approach .” P.Sharma,Dr.P.R.Bajaj, ITSC-07 ,International Conf., Washington.