Recognition of No-Stress and Stress with Gaussian Mixture Models (GMM) and K-Means Algorithms

Youssef Ahmed Rawy Baghdady
Ebraam Adel Latief Khalifa
Akshesh Chavda
Akanchha Choudhary
(University of Kassel)
e-mail:uk057346@student.uni-kassel.de

Abstract*— Recognition of No-Stress and Stress in smart environment has been developing certainly in recent years and it is significant for many applications due to the accurate predicted results. This paper presents a review of different algorithms techniques. These techniques used to recognize human feelings from Empatica E4 wristband sensor data. The sensors data were used in this study and these data took from the device. Two main steps describe the stress or no stress recognition process: data pre-processing and data classification. Two algorithms techniques namely, Gaussian Mixture Model (GMM) and K-Means are compared in terms error percentage and specificity. This comparison highlights which approach gives better performance in both algorithms.

***Keywords*— Wristband-E4-GMM-K-Means**

I. INTRODUCTION

These days are in everyday terms in almost everyone's life who are overburdened with the expectation and demands of their life. These might be related to anything from a high earning job to a peaceful relationships and comes a feeling of anxiety and unbalanced mental health that creates stress. So, it is important to recognize this problem because it comes along with underlying diseases like heart disease, headaches, and diabetes which further ruins normal life for worst. Now, the question is how can we make this recognition easier and reliable? There are many techniques like speech analysis and many more used before to perform this task but in our project we are going to explore a new method which we believe no have used it before to detect if a person is No-Stress and Stress in a certain situation. So, we are using GMM and K-means Clustering algorithm. In Fig. 1, GMM can limit the covariance of the different classes for example: diagonal, tied, spherical or whole. [1]

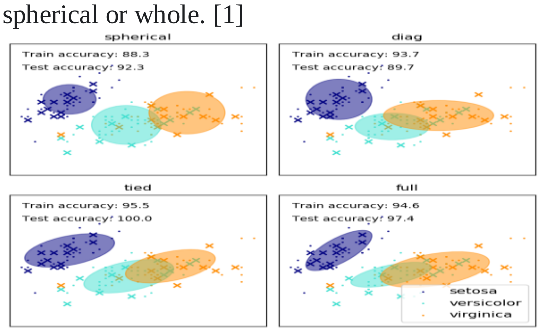


Fig. 1 GMM with different classes [1]

For case 1 D-dimensional vector of variables, The GMM is defined as weighted sum of K – Gaussian density function as: [2]

$$P(x) = \sum_{k=1}^{K} w_k N(x/\mu_k, \textstyle\sum_k)(1.1)$$

Where, $N(x/\mu_k, \sum_k)$ is component of distribution and $w_k$is mixing coefficient given for all w_k: $0 \le w_k \le 1$ and $\sum_{k=1}^{K} w_k = 1$.

To define Likelihood Function, the term 'Bayesian' comes from the usage of Bayes' theorem which was named after Thomas Bayes who first discovered the equation, so this basically define likelihood. The probability of event w conditioned on knowing event D is defined as: [2][3]

$$P(w|D) = \frac{P(D|w)P(w)}{P(D)} \, (1.2)$$

Where, $P(D) = \sum_w P(w|D)P(D)$

and the sum is over all possible values of D .

$P(w|D)$: The posterior probability that describes how certain hypothesis w is true, given that observed data D. It expresses all the necessary information in order to perform predictions.

$P(w)$: The prior probability which describes how certain hypothesis w is true, given that observed data D.

$P(D|w)$: Likelihood is like if assume that w is true, this is the probability that you would have observed data D.

$P(D)$: The probability that we would have observed data D, whether w is true or not.

II. STATE OF ART

The purpose of this section is to show the application of Gaussian mixture model and K-means clustering in estimation of human stress using physiological sensors. It will discuss and compare the data of different physiological sensors like Photo plethysmograph (PPG), Electrodermal activity (EDA) and Optical temperature sensors etc.

In [4], author represents Trier Social Stress Test (TSST) where psycho-social stress in participants is measured using highly standardized methods. The test is performed by taking the participant to a room where three judges, a video camera and an audio recorder are present. The first part consists of 5 minutes of anticipatory stress, where the participant must prepare a presentation lasting 5 minutes, commonly explained to be a job interview. The participant can use a paper and pen to prepare the presentation, but not when it is time to present. The judges should stay neutral during the test and the participant is asked to continue the presentation if it is finished before the 5 minutes. Directly after the presentation, a 5 minutes mental arithmetic test is performed in which participant has to count backwards from 1,022, subtracting 13 in each step. They must start again from 1,022 if a mistake is made in the calculations. After this, a recovery period is performed, followed by a debriefing. The participant is informed that the only purpose of the test was to create stress. Samples are collected for a while even after the stress tasks have ended.

In cold pressor test [5], physiological stress of participant is measured. The test is performed during morning in order to avoid influences from external factors. One task is performed by participants where they had to submerge their right hand in cold water (0 ◦ C - 4 ◦ C), until the wrist. It should continue maximum for 3 minutes, but they can however maintain the hand if possible or can remove the hand earlier. Heart rate, blood pressure and skin temperature are recorded during the immersion. The experiment leader watches the participants during the whole test and instruct them to remove their hand if the full 3 minutes are reached. Participants should be in good health and non-smoking.

In [6], another stress measuring task called mental arithmetic stress task is discussed. A quick mental test is performed among the participants and their results are compared. The task consists of reverse counting from 2,193 in steps of seven as rapidly as possible. It is similar to mental arithmetic test in Trier Social Stress Test. If any mistake occurs during the task, then participant has to start it again. The task continues for five minutes and then participant is asked to relax.

From all different methods mentioned above, we decided to go with cold pressor test for estimating human stress using GMM and K-means clustering. We had E4 empatica wristband as a tool to take measurement of stress using its inbuilt PPG, EDA, HR and Temperature sensors. The readings obtained from that sensors are preprocessed with data sets and clustering it with GMM and K-means algorithm.

III. METHODOLOGY OF RECOGNITION WITH GMM AND K-MEANS ALGORITHMS

A. Experimental Setup

The process of analysis is to set up the smart device which is a sensor device measures a physiological sensor. As the state of art mentioned in case of cold pressor test, the participants should stay in a room at room temperature then align the device in right position on hand while wearing it. After this, the participants should do in case of No-Stress and Stress at maximum for 3 minutes. In case of stress, the participants have to submerge their right hand in cold water (0 ◦ C - 4 ◦ C) till the maximum for 3 minutes, but the participant can remove the hand then the physiological sensor has recorded during the immersion. After the submersion, the participant is asked to judge the discomfort, stress and pain level, on a scale from 0 to 10. [5]

In our measurement results, we have measured four participants in two cases No-Stress and Stress. In case of stressed, we have used cold water at 1.9 Celsius degree and it was measured by thermometer then the participants put their hands during taking the measurement results. In case of not stressed no water used and it was in normal room temperature.

B. GMM algorithm

In GMM, the likelihood function should be maximized with respect to the parameters means and variances of GMM with given D-dimensional data points $\overrightarrow{x_n}$. The parameters are estimated using the EM algorithm. the steps for EM algorithm are given below. [2]

1. Initialize all mean values $\overrightarrow{\mu_k}$, covariances $\sum_k$, and mixing coefficients w_k.
2. Expectation step: Evaluate the posterior probabilities $P(k|\overrightarrow{x_n})$ that $\vec{x}$ belongs to component k, given by:

$$P(k|\overrightarrow{x_n}) = \frac{w_k N(\overrightarrow{x_n}/\overrightarrow{\mu_k}, \Sigma_k)}{\sum_{j=1}^{K} w_j N(\overrightarrow{x_n}/\overrightarrow{\mu_j}, \Sigma_j)} \quad (1.6)$$

3. Maximization step: For each component k re-estimate the parameters using the current probabilities, which are mixing coefficient means, and variances, given by:

$$w_k^{\ new} = \frac{1}{N} \sum_{n=1}^{N} P(k|\overrightarrow{x_n}) \quad (1.7)$$

$$\overrightarrow{\mu_k}^{new} = \frac{1}{N\, w_k^{\ new}} \sum_{n=1}^{N} P(k|\overrightarrow{x_n})\, \overrightarrow{x_n} \quad (1.8)$$

$$\sum_k^{\ new} = \frac{\sum_{n=1}^{N} P\left(k|\overrightarrow{x_n}\right)(\overrightarrow{x_n} - \overrightarrow{\mu_k}^{new})(\overrightarrow{x_n} - \overrightarrow{\mu_k}^{new})^T}{N\, w_k^{\ new}} \quad (1.9)$$

4.Evaluate the log likelihood function, given by:

$$\ln P(X|\boldsymbol{\mu}, \Sigma, \vec{w}) = \sum_{n=1}^{N} \ln\{\sum_{k=1}^{K} w_k N(\overrightarrow{x_n}/\overrightarrow{\mu_k}, \Sigma_k)\} \quad (2.0)$$

If the convergence criterion is not satisfied, return to step 2.

C. K-Means algorithm

The K-Means algorithm is deterministic clustering algorithms that dividing observations into k groups. Each of observations only belong to one cluster which has the nearest cluster centroid to this observation [10]. This algorithm segments the input data points into K-clusters. In this algorithm, the distance can be measure by using the Euclidean distance, given by: [7]

$$\sqrt{\sum_{i=1}^{d} (x_i - y_i)^2} \quad (1.3)$$

where x_i and y_i are two points in a D-dimensional Euclidean space. Then, the objective is that function has to be minimized is so-called Sum of Squared Error (SSE), given by:

$$SSE = \sum_{i=1}^{d} \sum_{x_i \in c_k} (x_i - c_k)^2 \quad (1.4)$$

The cluster centroid c_k could be updated, given by:

$$c_k = \frac{\sum_{x_i \in c_k} x_i}{c_k} \quad (1.5)$$

The K-Means algorithm has the following producer steps: [8]

1. Label the number of clusters.
2. Evaluate the centroid coordinate.
3. Calculate the distance of each object to the centroid.
4. Group the objects based on minimum distance.

Moreover, The K-Means clustering is a method of grouping the data observations into K-clusters by minimum distance between center of cluster and the data.

As difference of two algorithms, the K-Means algorithm defines as hard assign a data point to one cluster, but GMM algorithm is soft assigns a point to clusters which give a kind of probability of any point to any centroid. In practical way, the K-Means can only detect the spherical cluster and GMM can make an elliptic shape cluster. In other words, the K-Means can be a special case of GMM.

D. Methodology of the data

To evaluate the data, the smart device has created the data sets for all sensors was measured for participant in case of No-Stress and Stress. For pre-processing, the data sets for No-Stress and Stress will labelled as "Notstressed" and "stressed" in a Comma-Separated Values (CSV) file for each sensor expect accelerometer sensor. Because, the accelerometer is measuring the movement of participant like, for instance, walking or running, but the participant is not moving any kind of direction. So, the participant is just stay stable on certain position.

Therefore, the evaluation process of the data is to see the data sets result of the error in percentage for each sensor, expect accelerometer, by using GMM algorithm for fitting the data in a shape of distribution then calculate the centroid by K-Means algorithm for result the error in percentage.

IV. EVALUATION RESULTS OF REGONITION WITH GMM AND K-MEANS ALGORITHMS

In Evaluation part, we will evaluate the test measurement for all sensors and comparison of four participant for three sensors which are Blood Pressure Volume (BVP) in nanoWatt(nW), Heart Rate (HR) in beats per minute (bpm) and Skin Temperature (TEMP) in Celsius (℃) excluding Electrodermal Activity (EDA)in microSimens(μS) and Inter-Beat Interval (IBI) in milliseconds(ms) because those are most effect in our measurement that can be compared with participant.

For comparison of the Fig. 2 and Fig. 3, for instance, EDA for testing using Weka and Python in case of GMM, The Weka shows all data sets for two different clusters, which are "Notstressed" on bottom with blue color and "Stressed"on top with red color of Fig. 1 , are kind of overlapping and this will case an error, but in Fig. No. for Python just make a boundary between two cluster with red for "Notstressed" and black for "Stressed" to avoid overlapping, so Python would give a good result of. Also, the Weka gives the mean and standard deviation in order to define the cluster for "Notstressed" and "Stressed".

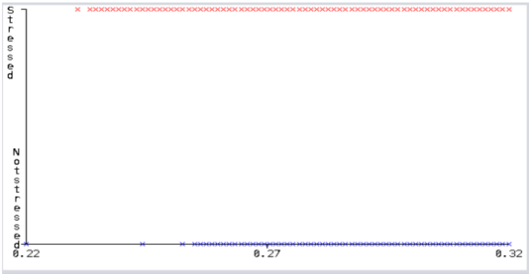


Fig. 2 GMM EDA result on Weka.

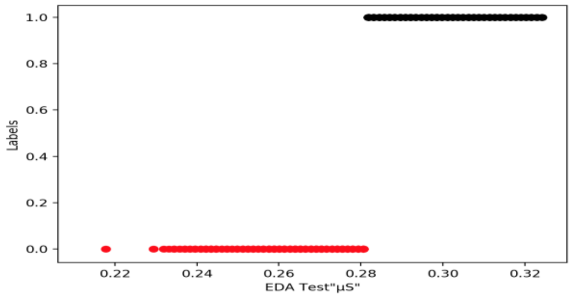


Fig. 3 GMM EDA result on Python.

For test measurement, the data sets are assigned as one-dimension space, so this will align on x-axis. For each sensor of data sets, K-Means calculates the centroid for "Notstressed" and "Stressed" as shown in Table. 1 are resulted by Weka and Python. BVP is resulted in K-Means for "Notstressed" is 8.9493 nW and "Stressed" is -20.2456 nW with mostly near 50 % of error as mentioned in Table 1. For EDA, the range from 0.22 to 0.32 μS points give the result of K-Means for "Notstressed" is 0.2988 μS and for "Stressed" is 0.2629 μS, so the centroid is nearly to each other might cause much errors. Within range from 50 to 90 bmp points, the HR shows K-Means results for "Notstressed" and "Stressed" are quite close together, so this could have somehow less error between them. IBI is result of HR divided by BVP gives beats/min with range from 0.5 to 1, so K-Means for "Notstressed" and "Stressed" will have similarity error as BVP. For TEMP, it measures the skin temperature of participant, so K-Means for "Notstressed" is 33.141℃ and "Stressed" is 32.9815 ℃ with a very less error.

Sensor type	K-Means with Notstressed	K-Means with Stressed	Error %
BVP test with range from -550 to 230 points	8.9493 nW	-20.2456 nW	49.3884%
EDA test with range from 0.22 to 0.32 points	0.2988 μS	0.2629 μS	43.7026%
HR test with range from 50 to 90 points	74.0901bmp	80.3454 bmp	33.119%
IBI test with range from 0.5 to 1 points	0.7403 ms	0.8256 ms	45.4167%
TEMP test with range from 32 to 34 points	33.141 ℃	32.9815 ℃	10%

Table. 1 K-Means results for test measurements.

By using Python for test measurement, the data set of each sensor is cluster by boundary region that on y-axis shows cluster of "Notstressed" defined as "0" with red color and "Stressed" as "1" with black color, which are labels, shown in Fig. 4 for each type of sensor. Not only just Python, but also Weka can give us the mean and standard deviation of GMM, so Gaussian distribution can be drawn for "Notstressed" and "Stressed" as mentioned in Table. No for each type of sensor. But, the distribution is not drawn in Fig. No because the plotting has defined the cluster on y-axis, so they are already mentioned as "0" for "Notstressed" with red color and "1" for "Stressed" with black color. In case for IBI sensor, both cluster have same mean and standard devivation, so this will case an overlapping and more much error.

Sensor type	Mean for Notstress	Standard deviation for Notstressed	Mean for Stressed	Standard deviation for Stressed
BVP test	-12.12 nW	11.9001 nW	8.645 nW	7.0983 nW
EDA test	0.3012 μS	0.0073 μS	0.2446 μS	0.0064 μS
HR test	78.96 bmp	1.6418 bmp	81.7 bmp	0.5033 bmp
IBI test	0.7662 ms	0.0512 ms	0.7662 ms	0.0512 ms
TEMP test	32.9512℃	0.0148℃	33.1553℃	0.005℃

Table. 2 GMM's Mean and Standard deviation for test

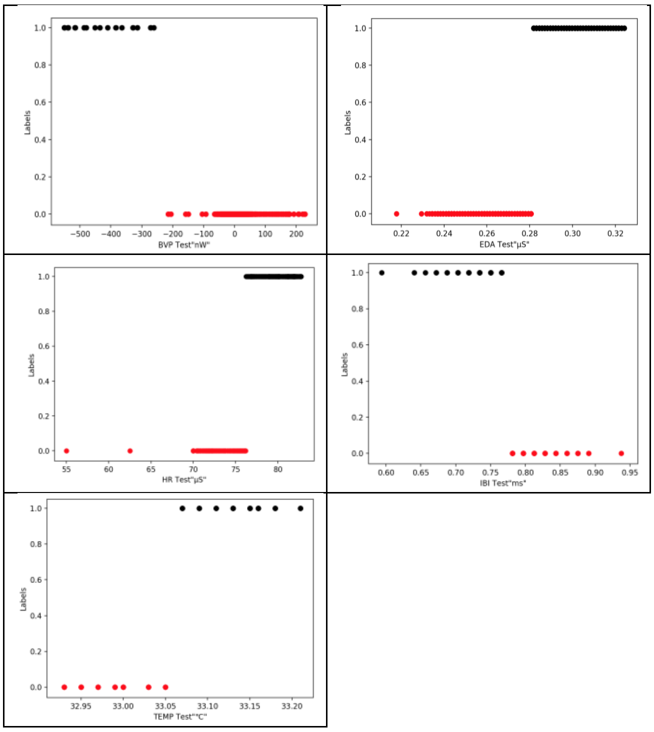


Fig. 4 GMM clustering model for test measurement.
measurement.

In the next tables we can see the comparison of the four participants for the three sensors, which are BVP, HR and TEMP, in both cases the GMM and K-Means and the percentage of error in each case because those are most important of our comparison.

For K-Means, The BVP sensor has the highest error 49.84 % than other sensors in Table. 3 and most give the good result with less error 22.88 % is TEMP sensor as shown in Table. 5 while HR sensor is most almost with sufficient results error 19.1 % as in Table. 4. Participant P2 is one of best results for three sensors with less error while P4 has the worst result with high errors, but only in BVP sensor P1 is the most worst result error is 49.84 %.

Participant number	K-Means not stressed [nW]	K-Means stressed [nW]	Error (%)
P1	53.1491	-89.9962	49.84 %
P2	26.2996	-47.5136	46.92 %
P3	-24.8529	14.6255	49.72 %
P4	-26.6405	13.6411	47.73 %

Table. 3 K-Means for BVP sensor

Participant number	K-means not stressed [bmp]	K-means stressed [bmp]	Error (%)
P1	90.3939	80.7918	28.72 %
P2	87.5324	71.85	19.1 %
P3	87.12	82.1493	30.3867 %
P4	80.5445	90.0447	44.04 %

Table. 4 K-Means for HR sensor

Participant number	K-Means not stressed [℃]	K-Means stressed [℃]	Error (%)
P1	33.7557	32.697	0 %
P2	34.0892	33.5005	0 %
P3	32.5863	32.7654	8.9 %
P4	31.9393	32.0257	22.88 %

Table. 5 K-Means for temp sensor

For GMM parameters as shown in Table. 6, 7 and 8, the one of the best result is P2 as mention in K-Means because it has less error so the two clusters of distributions for three sensors will have less overlapping between two clusters. Moreover, the P4 is worst result, so it will have more overlapping between two clusters of distributions.

Participant number	Mean for Not stressed [nW]	Standard Deviation for Not stressed [nW]	Mean for Stressed [nW]	Standard Deviation for Stressed [nW]
P1	32.80	28.6106	-47.81	39.6807
P2	45.92	8.2922	-24.83	9.0126
P3	15.01	7.2237	7.16	6.4694
P4	-13.73	10.513	9.70	7.1154

Table. 6 GMM for BVP sensor

Participant number	Mean for Not stressed [bmp]	Standard Deviation for Not stressed [bmp]	Mean for Stressed [bmp]	Standard Deviation for Stressed [bmp]

P1	85.70	2.675	92.78	1.0993
P2	77.75	9.1342	77.75	9.1342
P3	85.79	1.6125	81.66	1.0093
P4	78.73	0.4358	81.01	1.0093

Table. 7 GMM for HR sensor.

Participant number	Mean for Not stressed [°C]	Standard Deviation for Not stressed [°C]	Mean for Stressed [°C]	Standard Deviation for Stressed [°C]
P1	32.8315	0.0222	33.7501	0.0017
P2	33.4296	0.0206	34.2347	0.0165
P3	32.7565	0.0097	32.64	0.0231
P4	32.03	0.011	31.9472	0.0194

Table. 8 GMM for temp sensor.

To adjust GMM cluster for P2 and P4 with three sensors, this implemented using Python with boundary as “0” with red color for “Notstressed” and “1” with black for “Stressed”shown in Fig. 5 and Fig. 6 As those figures, The P2 has good clustering boundary for three sensors that dose not have any kind of overlapping between two clustering, but the P4 shown worst result with overlapping, except the temp has less overlapping.

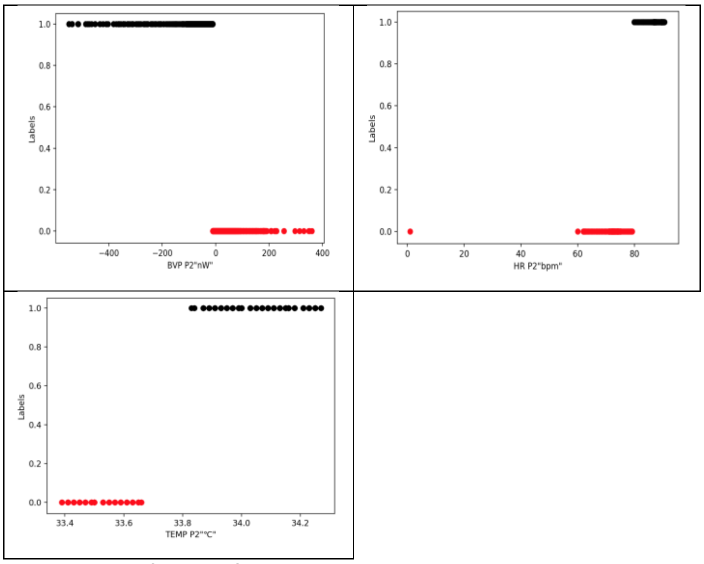
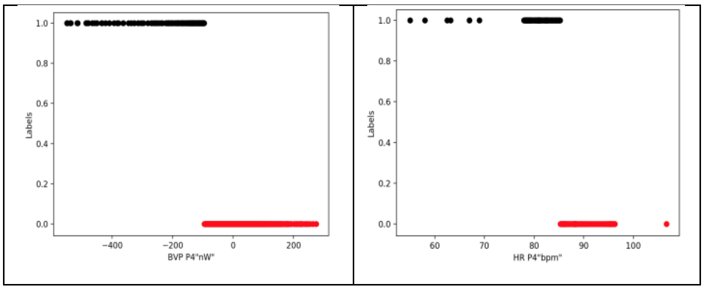


Fig. 5 GMM for P2 of three sensors.



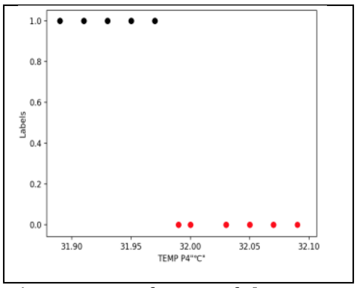


Fig. 6 GMM for P4 of three sensors.

V. CONCLUSION

The conclusions based on GMM and K-Means algorithm for case stress and no stress then goal of this is comparison of participant with best result and worst result. As analysis, K-Means performs K-observations from the dataset then uses the initial means for case stress and no stress. Moreover, GMM performs clustering for each class of label for stress and no stress with fitting estimation by maximum likelihood. Then, the preprocessing for analysis is that each sensor, which are the five sensors of test measurement, of data sets have assigned the labels for stress and no stress with choose of cold water in case for stress experiments.

In the end, we have evaluated the result of different participants with best result with three sensors, which are BVP, Hear Rate and Temp, that have chosen using the Weka program and Python. As participant number 2, the result is tested with very good accuracy after comparison in Section V while the participant number 4 does not give us a good accuracy.

REFERENCES

[1] Scikit Learn, “GMM Coveriances”, From: https://scikit-learn.org/stable/auto_examples/mixture/plot_gmm_covariances.html.

[2] C.J.Bishop,“Pattern recognition and machine learning,””pp.22,111,424,430”, New York: Springer-Verlag, 2009.

[3] A.Gelman, J.Carlin, H.Stern, D. Dunson, A.Vehtari and D.Rubin, Bayesian data analysis,“pp. 6,7”,3rd ed., 2014.

[4] Clemens Kirschbaum, Karl-Matrin, Pirke, and Dirk H.Helhammer, The “Trier Scoial Stress Test“ – A Tool for Investigating Psychobiological Stress Responses in a Laboroatory Setting”, “pp. 8,93”,Neuropsychobiology., 1993.

[5] Lars Schwabe, Leila Haddadm and Hartmut Schachinger, “HPA axis activation by a socailly evaluated cold-pressor test”, “PP. 8,94”, Psychoneuroendorcrinology, 2008.

[6] Morten Rostrup, Arne Westheim, Sverre E.Kjeldsen, and Ivar Eide, “Cardiovascular Ractivity, Coronary Risk Factors, and Sympathetic Activity in Young Men, “PP. 8,94”, Hypertension,1993.

[7] S.Banerjee, A.Choudhary,S.Pal, “Empirical Evaluation of K-Means, Bisecting K-Means, Fuzzy K-Means and Genetic K-Means Clustering Algorithms”, “PP. 169”, IEEE International WIE Conference,2015.

[8] S.Kapil, M.Chawla, M.Dilshad, “On K-Means Data Clustering Algorithm with Genetiv Algorithm ”, “PP. 202”,Fourth International Conference on Paralle,2016.

[9] K.Singh, S.Chatterjee, “ What is the difference between K-Means and the mixture model of Gaussian?”, from: https://www.quora.com/What-is-the-difference-between-K-means-and-the-mixture-model-of-Gaussian.

[10] F.Yi, I.Moon, “Extended K-Means Algorithm”, “PP. 263”,Fifth International Conference on Intelligent Human-Machine Systems and Cybernetics,2013.