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

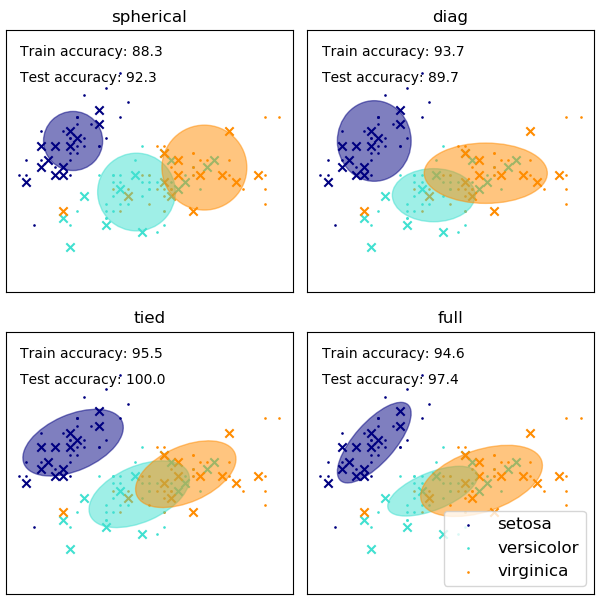
Youssef Ahmed Rawy Baghdady  
Ebraam Adel Latief Khalifa  
Akshesh Chavda  
Akcanchha Choudhary  
(University of Kassel)

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

Keywords— Wristband-E4-GMM-K-Mean

# Introduction

A Gaussian mixture model (GMM) can be defined as a probabilistic model that undertake or hypothesize that all the data points are produced from a mixture or combination of a limited number of Gaussian distributions with unfamiliar parameters. We can think of mixture models as rationalizing k-means clustering to assimilate information about the covariance structure of the information collected (data). In GMM our goal is to fit the mixture of Gaussian models. To achieve this, we use an algorithm called Expectation minimization. This algorithm basically gives an idea about which data points are originating from which hidden components in the observation by a reparative process. In Fig. 1, GMM can have different ways to limit the covariance of the different classes for example: diagonal, tied, spherical or whole. [1]



**Fig.** 1 GMM with different classes [1]

# motıvatıon

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

Where, is component of distribution and is mixing coefficient given for all : and

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 conditioned on knowing event is defined as: [2][3]

(1.2)

Where,

.

: The posterior probability that describes how certain hypothesis is true, given that observed data . It expresses all the necessary information in order to perform predictions.

: The prior probability which describes how certain hypothesis is true, given that observed data .

: Likelihood is like if assume that is true, this is the probability that you would have observed data .

: The probability that we would have observed data , whether is true or not.

# 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. First, the participants are moved to one room and after few minutes they are asked to submerge their right hand in cold water (0 ◦ C - 4 ◦ C), until the wrist. It should continue maximum for 3 minutes, but the participant can remove the hand earlier. They can however maintain the hand if possible. Heart rate, blood pressure and skin temperature are recorded during the immersion. Immediately after the submersion, the participant is asked to judge the discomfort, stress and pain level, on a scale from 0 to 10. 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], the task uses a mental arithmetic stress that explained to the participant for a quick mental test is performed in order to compare the result with other participants. This explained that other participants experienced the test as an easy task while the test consists of counting backwards from 2,193 in steps of seven. To rephrase, the mental arithmetic tests is similar in the Trier Social Stress Test. If a mistake is made, the participant should do repeat the producer. If the participant hesitates or pauses, the producer should be continued. This continues for five minutes, then the participant can be relaxing.

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.

# METHODOLOGY OF RECOGNITION WITH GMM AND K-MEAN ALGORITHMS

## 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, the test is performed during morning in order to avoid influences from external factors. First, the participants are moved to one room and after few minutes they are asked to submerge their right hand in cold water (0 ◦ C - 4 ◦ C), until the wrist. It should continue maximum for 3 minutes, but the participant can remove the hand earlier. They can however maintain the hand if possible. Heart rate, blood pressure and skin temperature are recorded during the immersion. Immediately after the submersion, the participant is asked to judge the discomfort, stress and pain level, on a scale from 0 to 10. 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. [5]

In our measurement results, we have measured four participants in two cases Stressed and No-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.

## K-Means algorithm

The K-Means algorithm segments the input data points

into K-clusters. In this algorithm, the distance can be measure by using the Euclidean distance, given by:

(1.3)

where and 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:

(1.4)

The cluster centroid could be updated, given by:

(1.5)

The K-Means algorithm has the following producer steps:

Step 1: Label the number of clusters.

Step 2: Evaluate the centroid.

Step 3: Calculate the distance of each object to the centroid.

Step 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.

## 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 The parameters are estimated using the EM algorithm. the steps for EM algorithm are given below. [2]

Step 1: Initialize all mean values , covariances , and mixing coefficients

Step 2: Expectation step: Evaluate the posterior probabilities that belongs to component , given by:

(1.6)

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

(1.7)

(1.8)

(1.9)

Step 4: Evaluate the log likelihood function, given by:

(2.0)

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

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.

## Methodology of the data

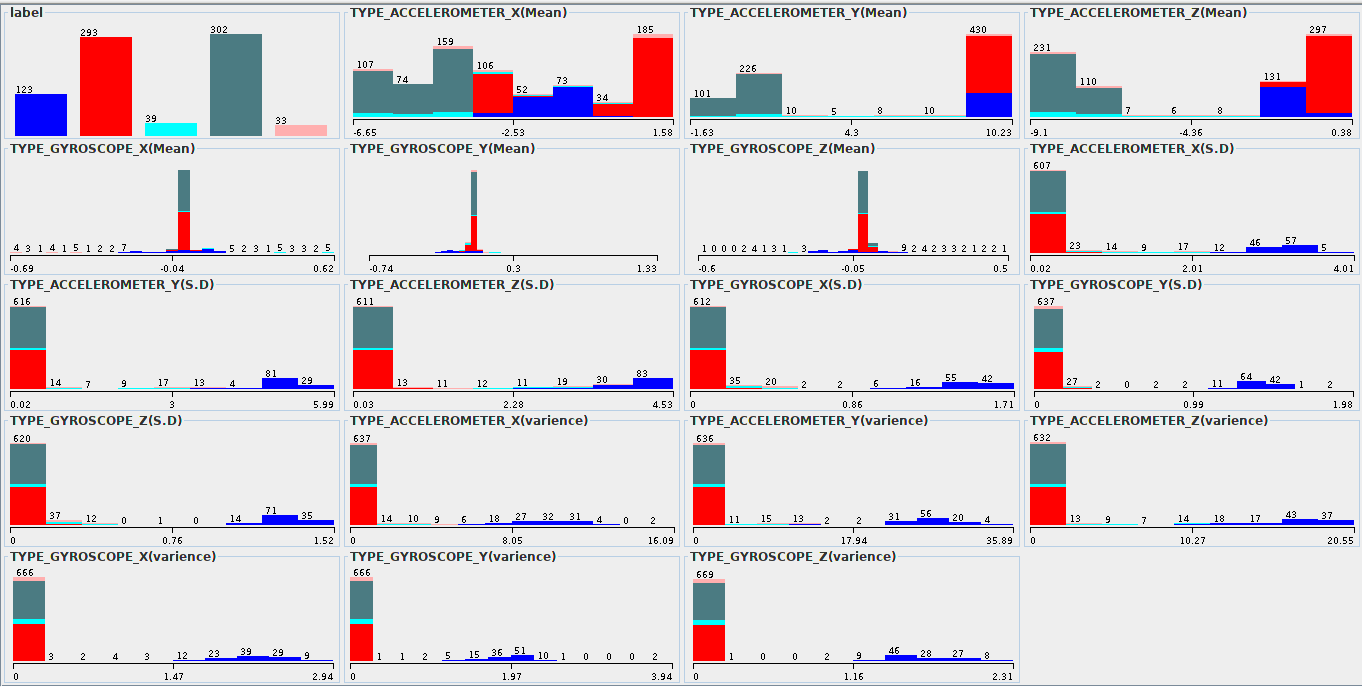
To evaluate the data, the smart device has created the data sets for all sensors was measured for participant in case of Stress and No-Stress. For pre-processing, the data sets for Stress and No-Stress will labelled as “Stressed” and “Notstressed” in a Comma-Separated Values (CSV) file for each sensor 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.

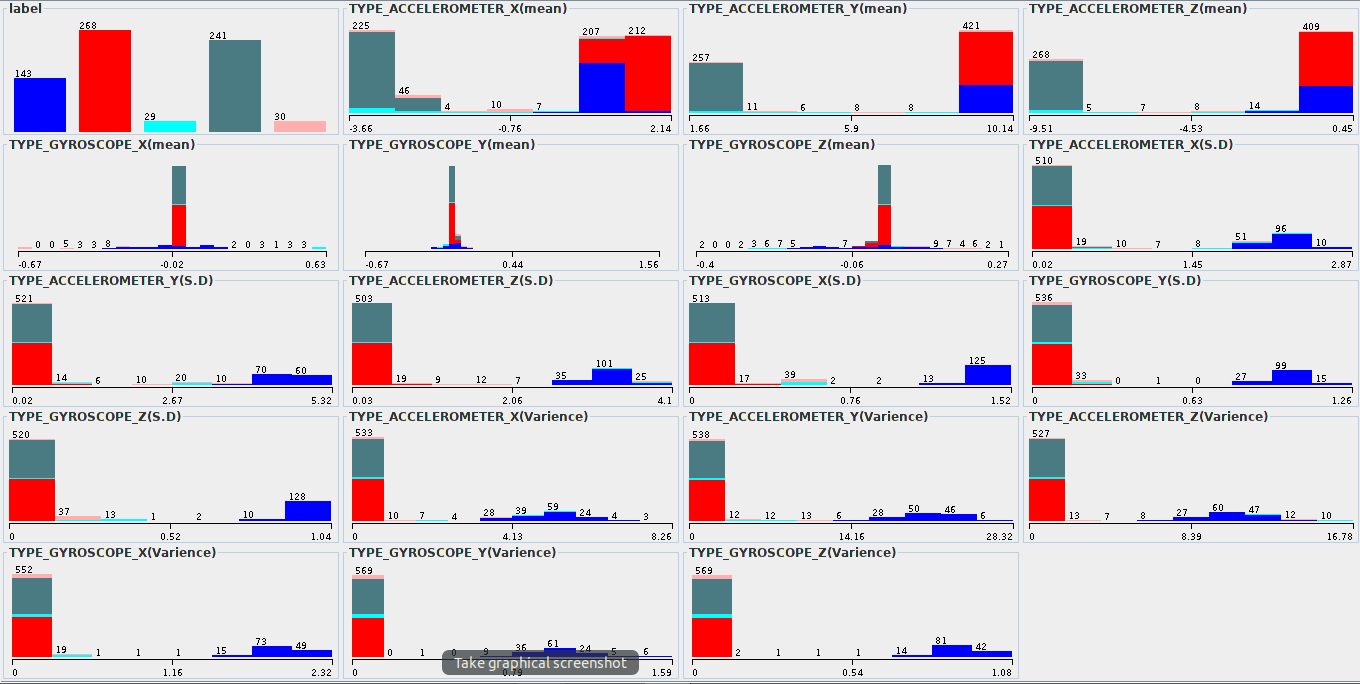
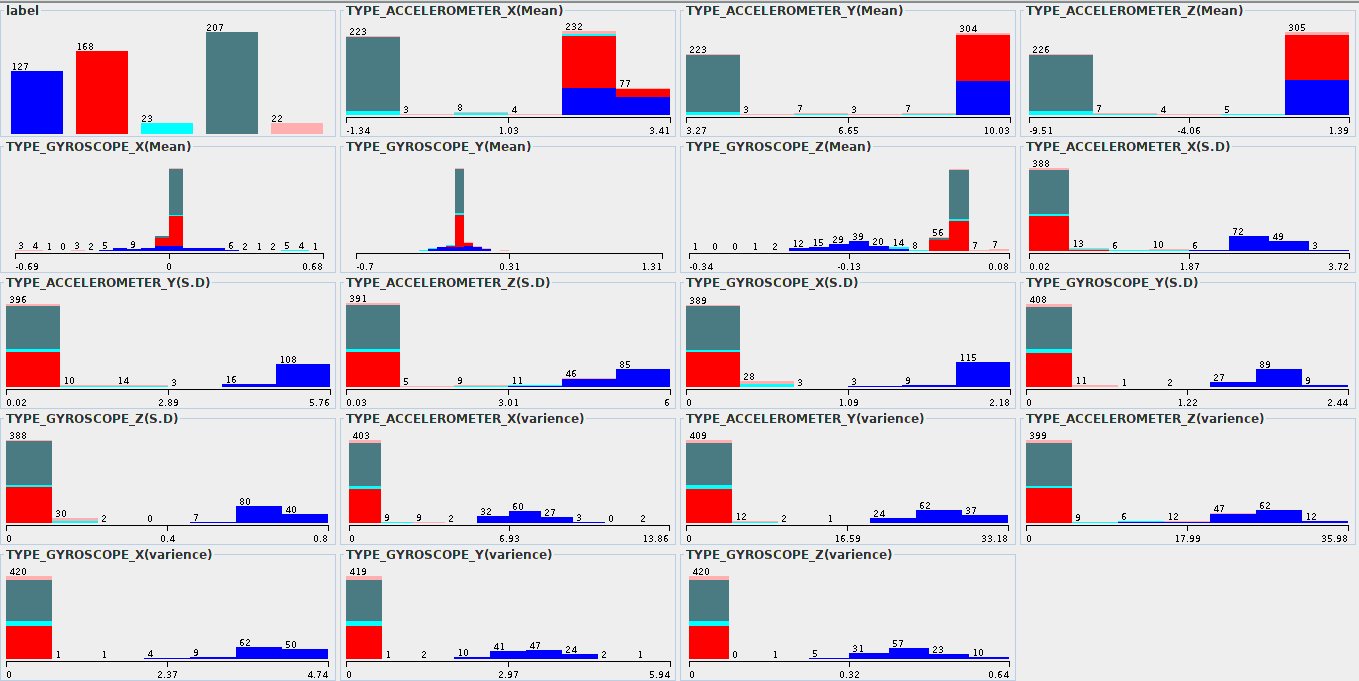
Therefore, the evaluation process of the data is to see the data sets result of the error in percentage for each sensor, expect accelerometer, by using K-Mean algorithm then the fitting the data in a shape of distribution by GMM algorithm.

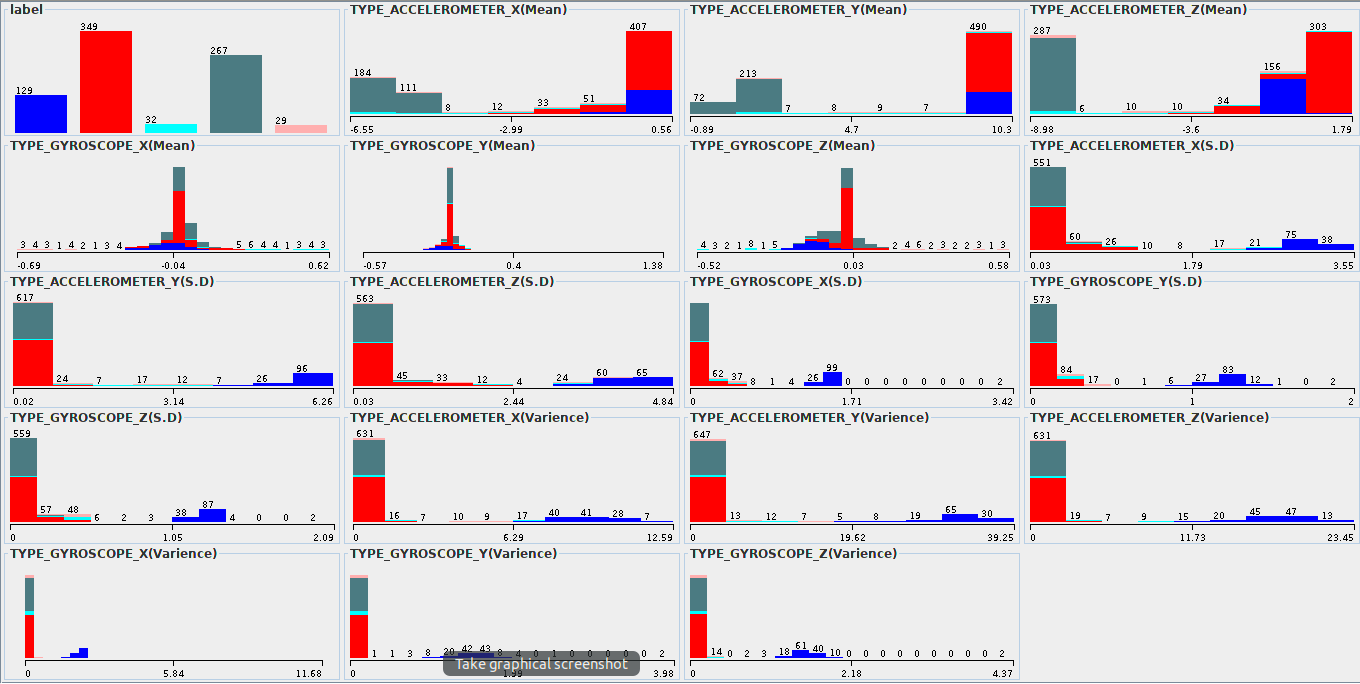
# EVALUATION RESULTS OF REGONITION WITH GMM AND K-MEAN ALGORITHMS

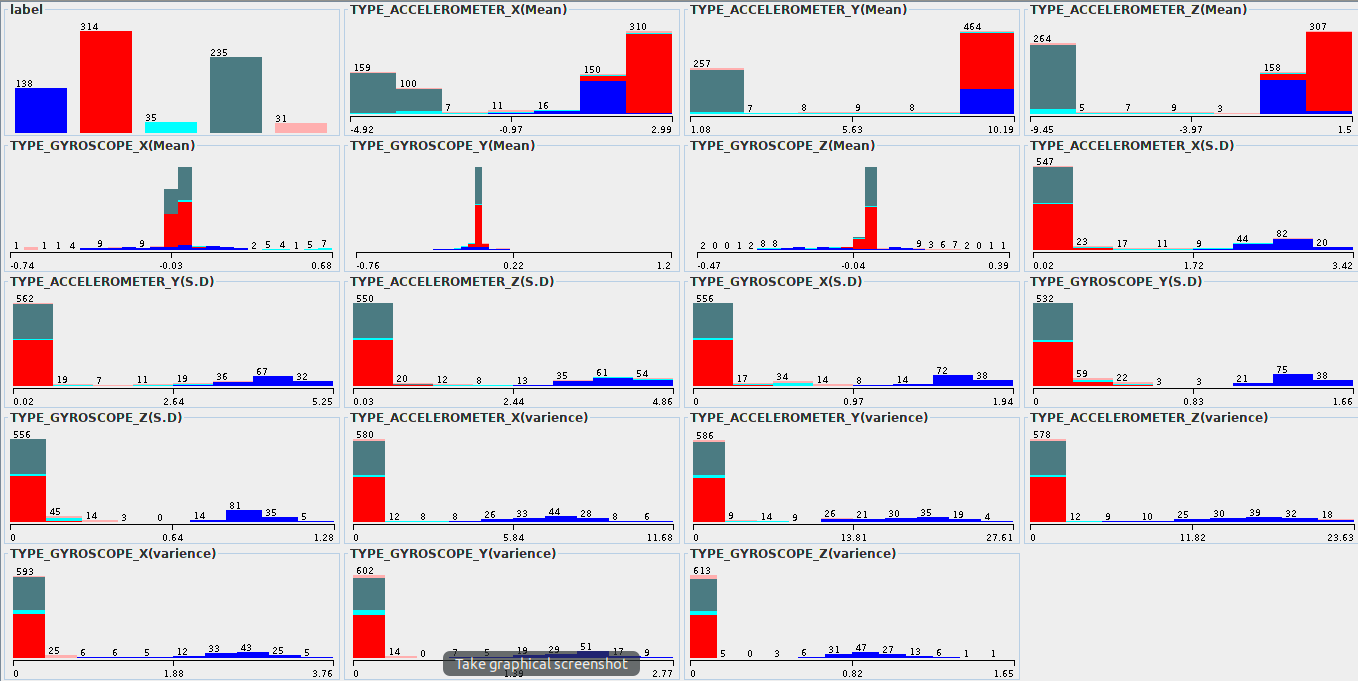
In Evaluation part, we will evaluate the test measurement for all sensors and comparison of four participant for three sensors which are Blood Pressure Volume (BVP), Heart Rate (HR) and Skin Temperature (TEMP) excluding Electrodermal Activity (EDA) and Inter-Beat Interval (IBI) because those are most 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 and “Stressed” is -20.2456 with mostly near 50 % of error as mentioned in Table 1. For EDA, the range from 0.22 to 0.32 points gives the result of K-Means for “Notstressed” is 0.2988 and for “Stressed” is 0.2629, so the centroid is nearly to each other might cause much errors. Within range from 50 to 90 points, the HR shows K-Mean’s 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 33.141and “Stressed” is 32.9815 with a very less error.

**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.

**Table.** 2 GMM’s Mean and Standard deviation for test

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

measurement.

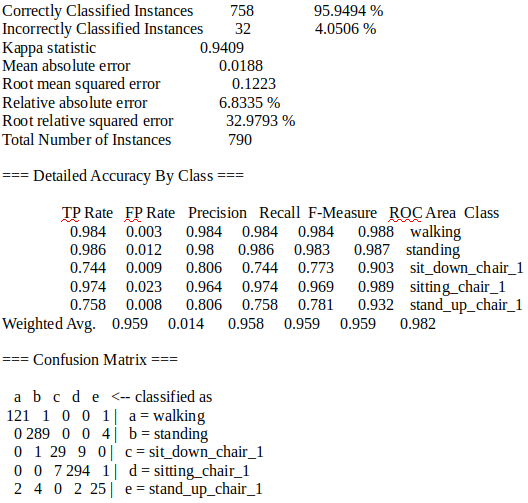
In the next tables we can see the comparison of the four participants for the three sensors, which are BVP, HR and TEMP, in both cases the GMM and K-Mean and the percentage of error in each case because those are most important of our comparison.

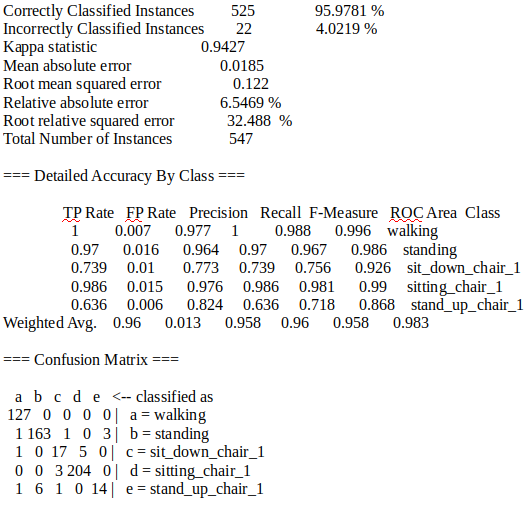
For K-Mean, The BVP sensor has the highest error than other sensors in Table. 3 and most give the good result with less error is TEMP sensor as shown in Table. 5 while HR sensor is most almost with sufficient results as in Table. 4. Participant P2 is one of best results for three sensors with less error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

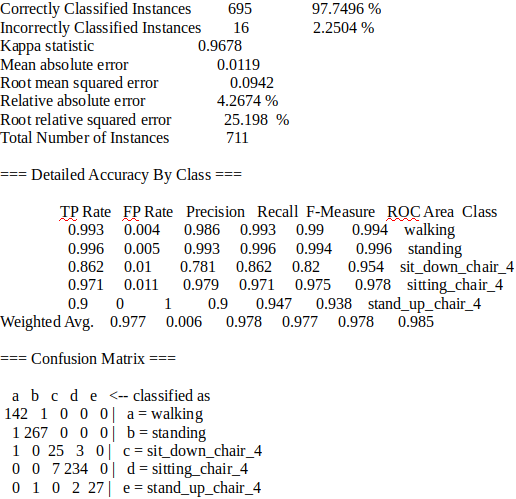
**Table**. 3 K-mean for BVP sensor

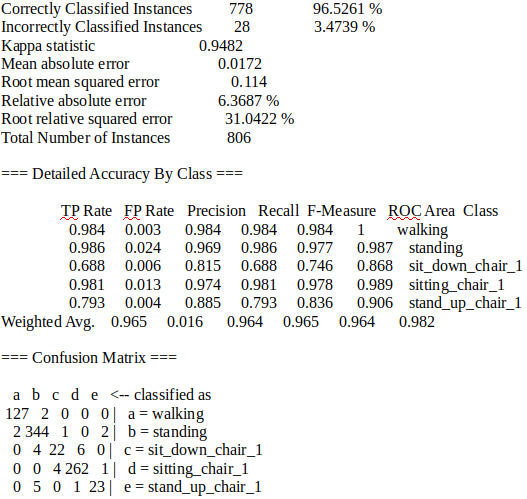
**Table**. 4 K-mean for HR sensor

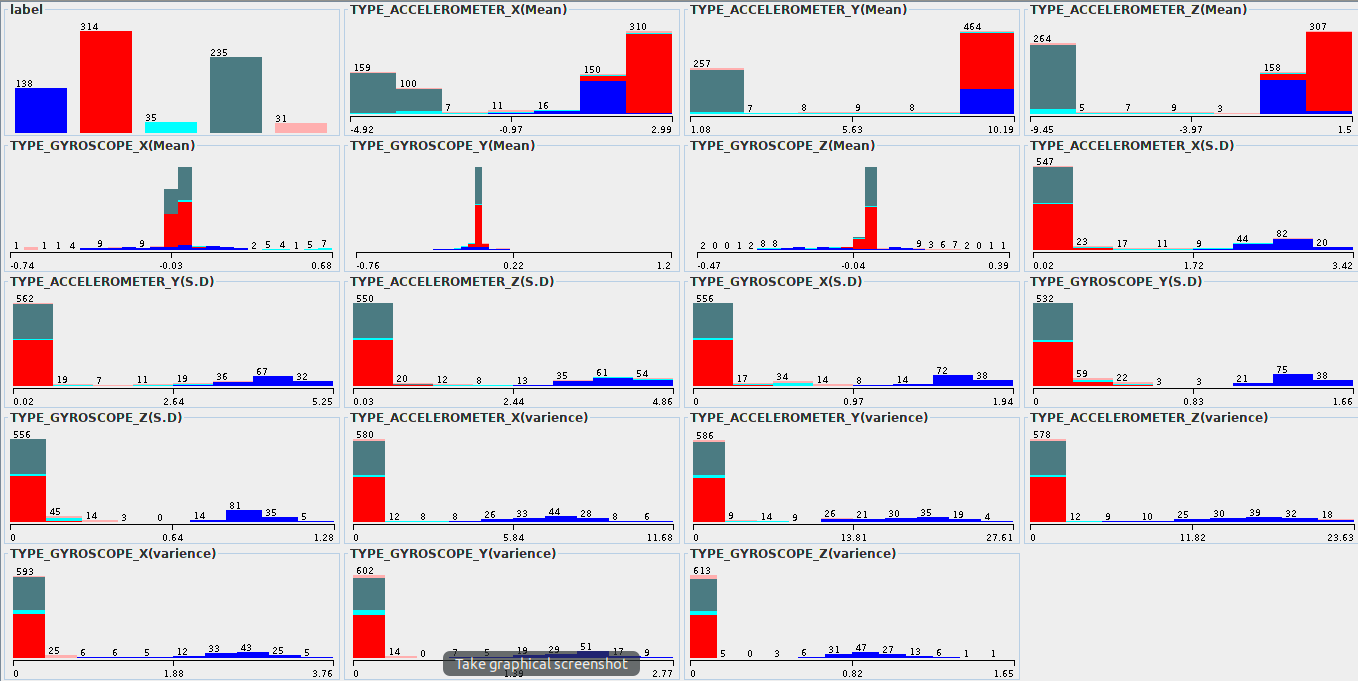
**Table**. 5 K-mean for temp sensor



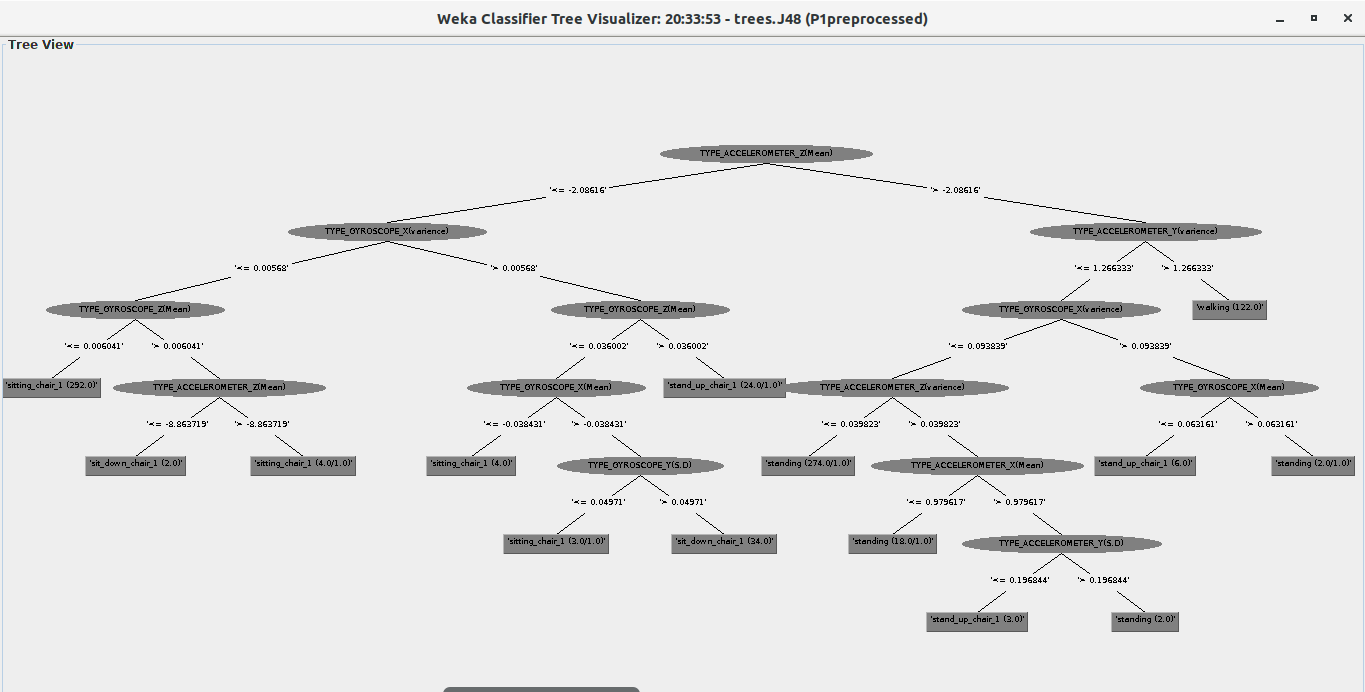




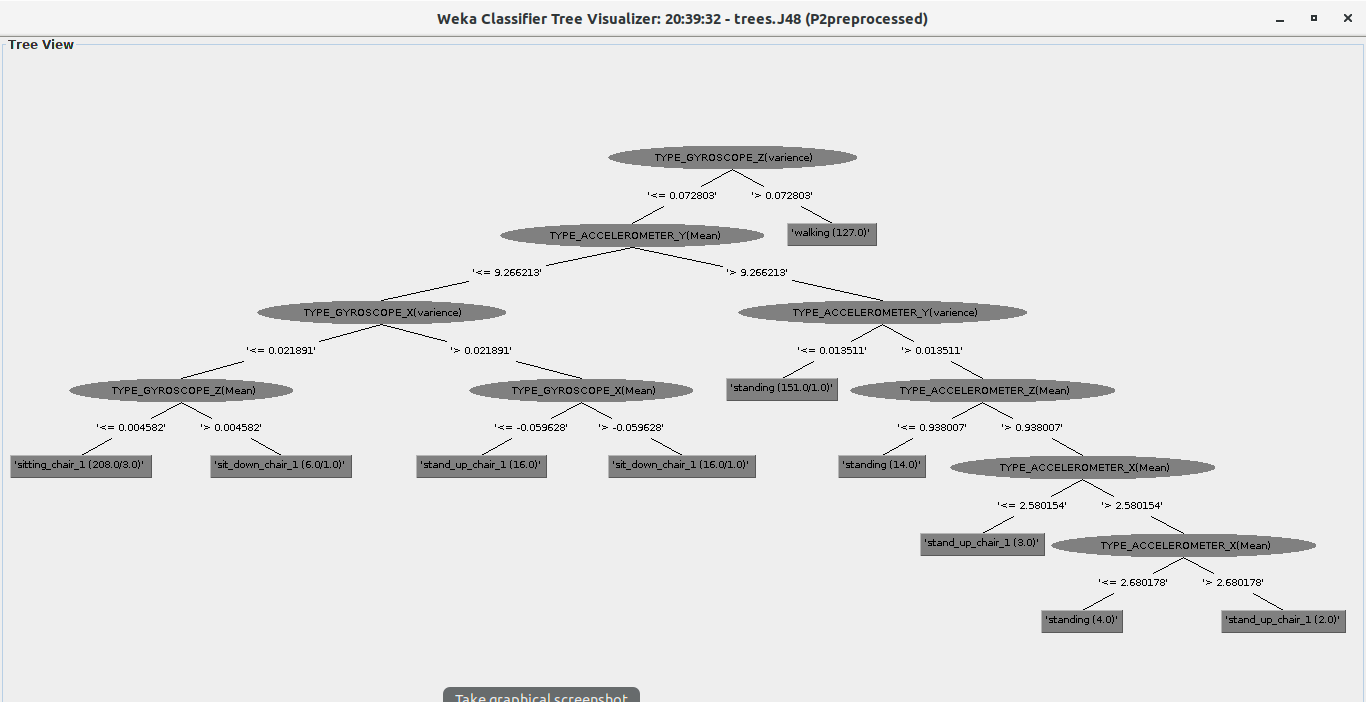


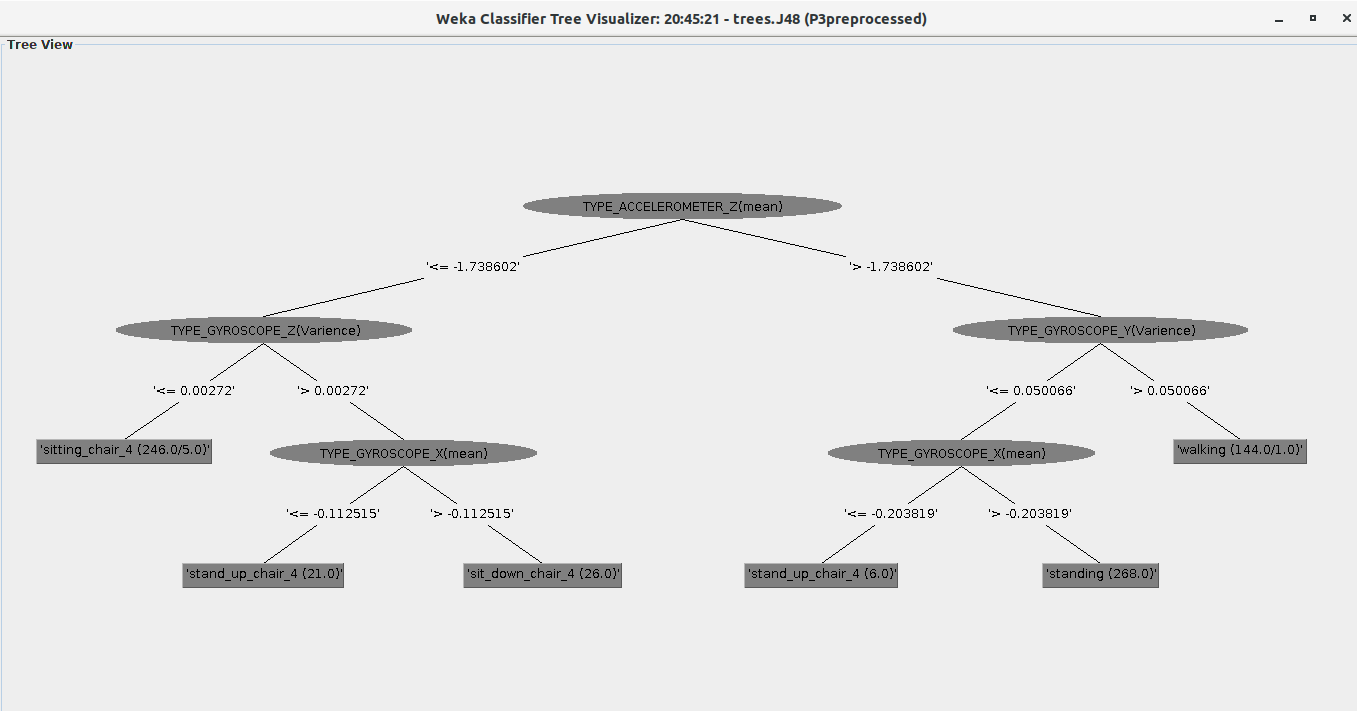


For GMM parameters as shown in Table. 6, 7 and 8, the one of the best result is P2 as mention in K-Means because it has less error so the two clusters of distributions for three sensors will have less overlapping between two clusters. Moreover, the P4 is worst result, so it will have more overlapping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

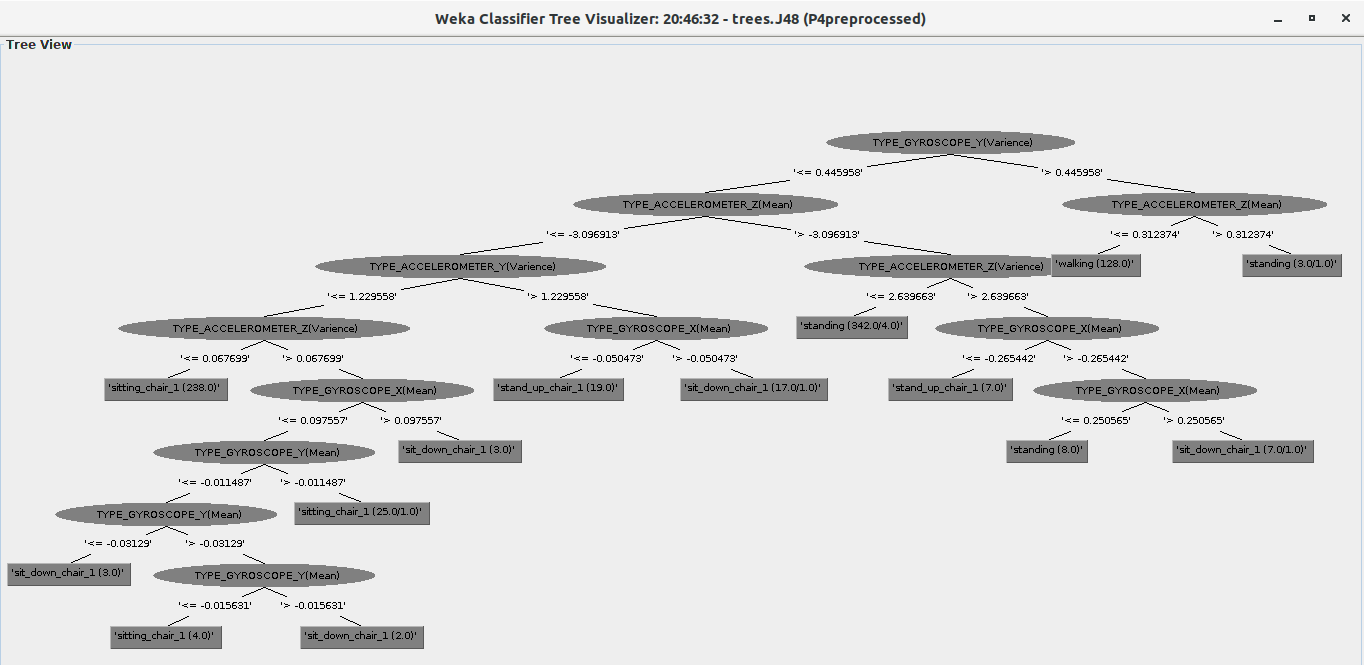


**Table**. 6 GMM for BVP sensor

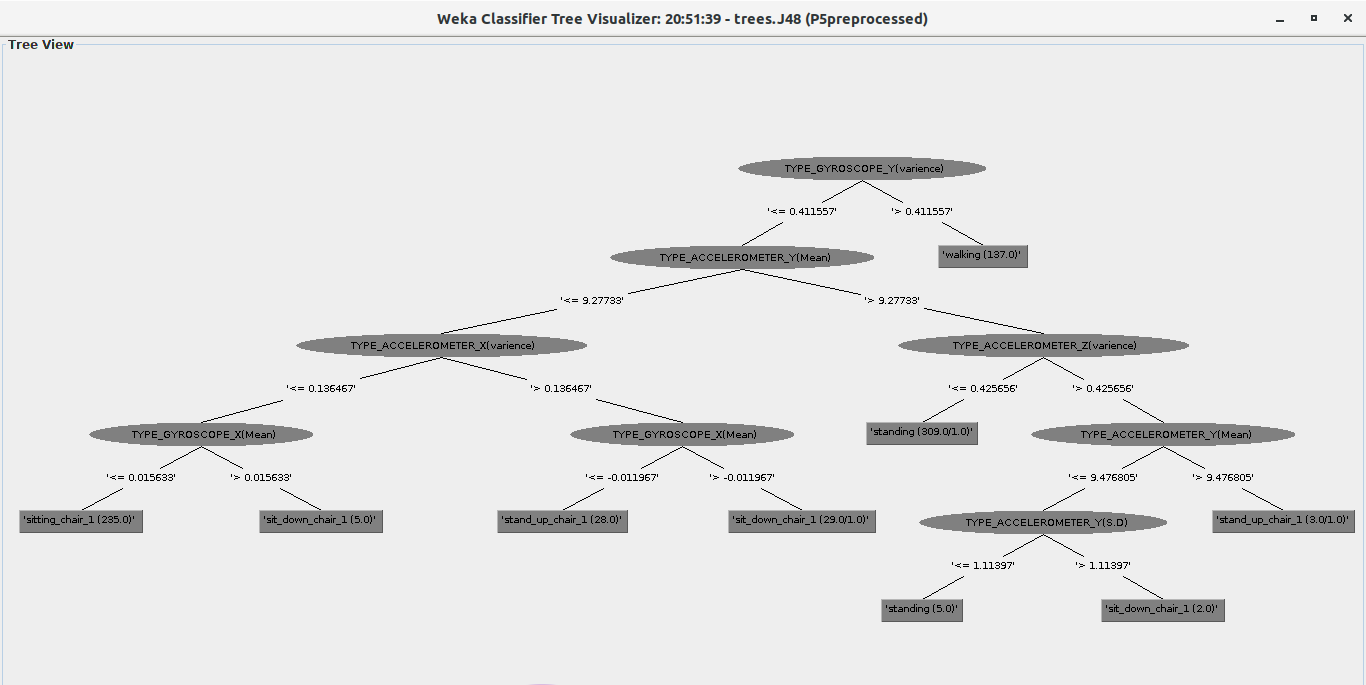




**Table**. 7 GMM for HR sensor



**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.

# 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”, unpublished.
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.