Clustering Gait Data using different machine learning techniques and finding the best technique

Presented by
ANUBHA PARASHAR

Computer Science and Engineering
MANIPAL UNIVERSITY, JAIPUR.

Outline

- GAIT/ Human GAIT
- Problem Statement
- Proposed Solution
- AIM
- Proposed work and Methodology
- Data Collection & Demos
- Clustering of gait data using SOM, K-Mean
- Result analysis
- Conclusion

What is GAIT?

Gait

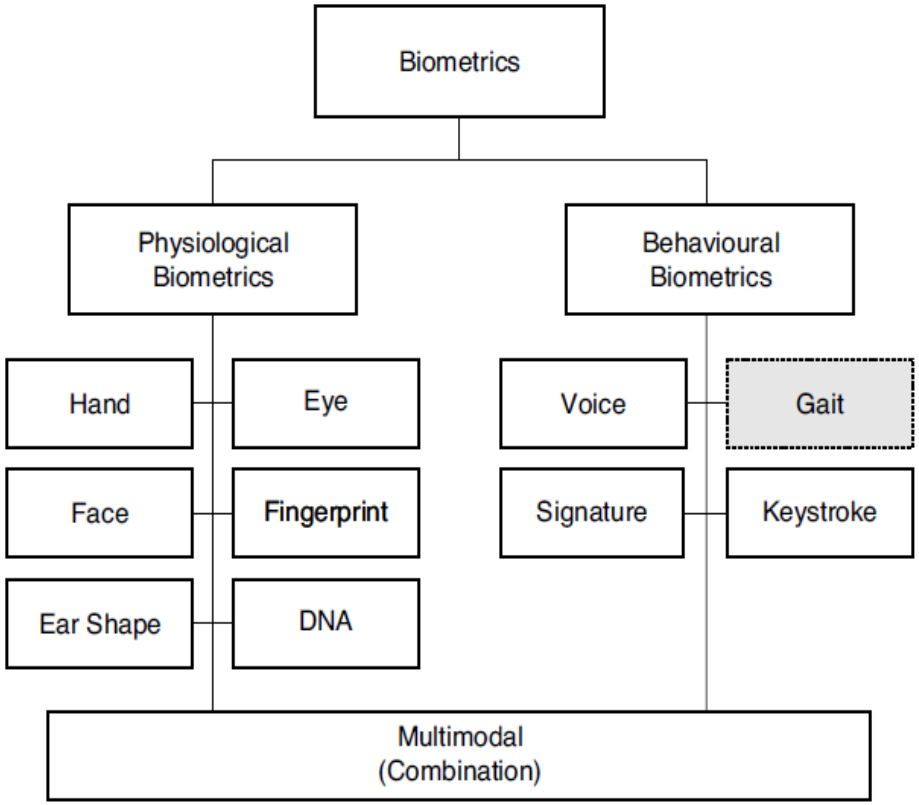
- Gait refers to the manner in which a person walks, and is one of the few biometric traits that can be used to recognize people at a distance.
- Therefore, this trait is very appropriate in surveillance scenarios and gait recognition system.

Human gait

- Human gait analysis refers to the study of human locomotion.
- Applications are in a large variety of domains, including entertainment, healthcare and pedestrian navigation.
- The act of walking involves the coordination of different human body parts, such as the skeleton, the muscle and the neural systems.
- Various factors can affect the complex interaction between the body parts, for example any pathological nature will require a distinction between — **normal gait** and **pathological gait**.
- Normal gait refers to the general human walking parameters without differentiating the age, sex or physical parameters.
- pathological gait refers to an abnormal gait, for example affected by pathologies such as muscle weakness or skeleton deformities.

Problem Statement

- To verify/validate the identity of people is called as authentication.
- And now a day's authentication is becoming very popular to check whether a person is authentic or not.
- Previously authentication was done with secret passwords or PIN codes. The disadvantage of such authentication systems was anyone could crack the password or PIN codes.
- So, biometric (in fig.) technology came to existence where authentication is done using a particular body part like fingers, iris etc. But now a days these biometrics are not user friendly as people dislike accessing finger print machine or avoid iris scanner as it may weaken or damage their vision.

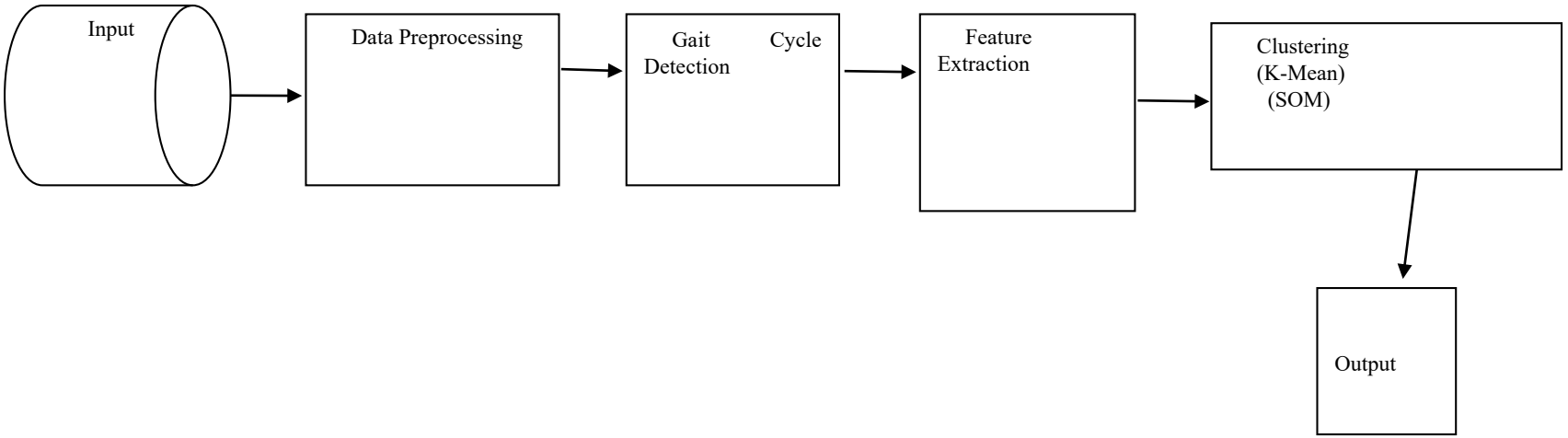


Proposed Solution

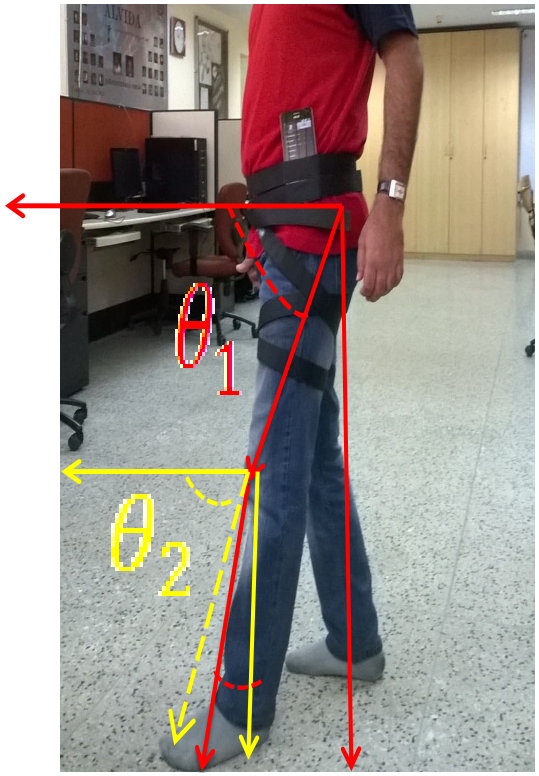
- Identifying normal or pathological human gait. Therefore a robust method for detecting and capturing the body, face and/or gait of each individual person has to be found.
- Human gait is supposed as unique biometric identification like thumb print.
- So in GAIT analysis to examine the human walking moment we collect and analyze the data.
- Human gait can be utilized to identify people for diverse security reason and exercised before to detect many abnormalities.

Proposed Solution

We have developed the computational model for prediction, formal verification and analyses of bipedal locomotion using machine learning technique for modelling.



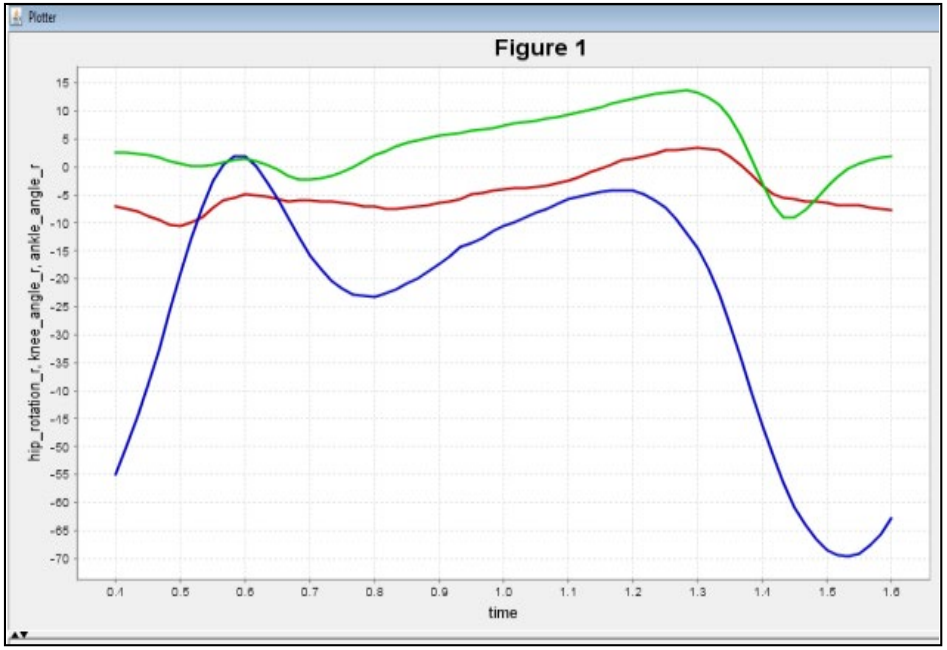
Empirical Study of Gait by HMCD



METHODOLOGY

TO GATHER DATA AND EXTRACT FEATURE

- Data is captured using body sensor based Human Motion Capture Device (HMCD) and potentiometer.



AIM

- The aim of the paper is to evaluate and implement robust and accurate method to identify whether it's a **normal gait** or **pathological gait.**

Methodology

METHODOLOGY

- Firstly we collect the GAIT data then features are extracted and we cluster the GAIT dataset.

- Training and testing of system is done.

- After the Clustering is done result is shown as normal or pathological GAIT.

- Then 10 fold cross-validation test is also conducted to validate the statistical significance of the results.

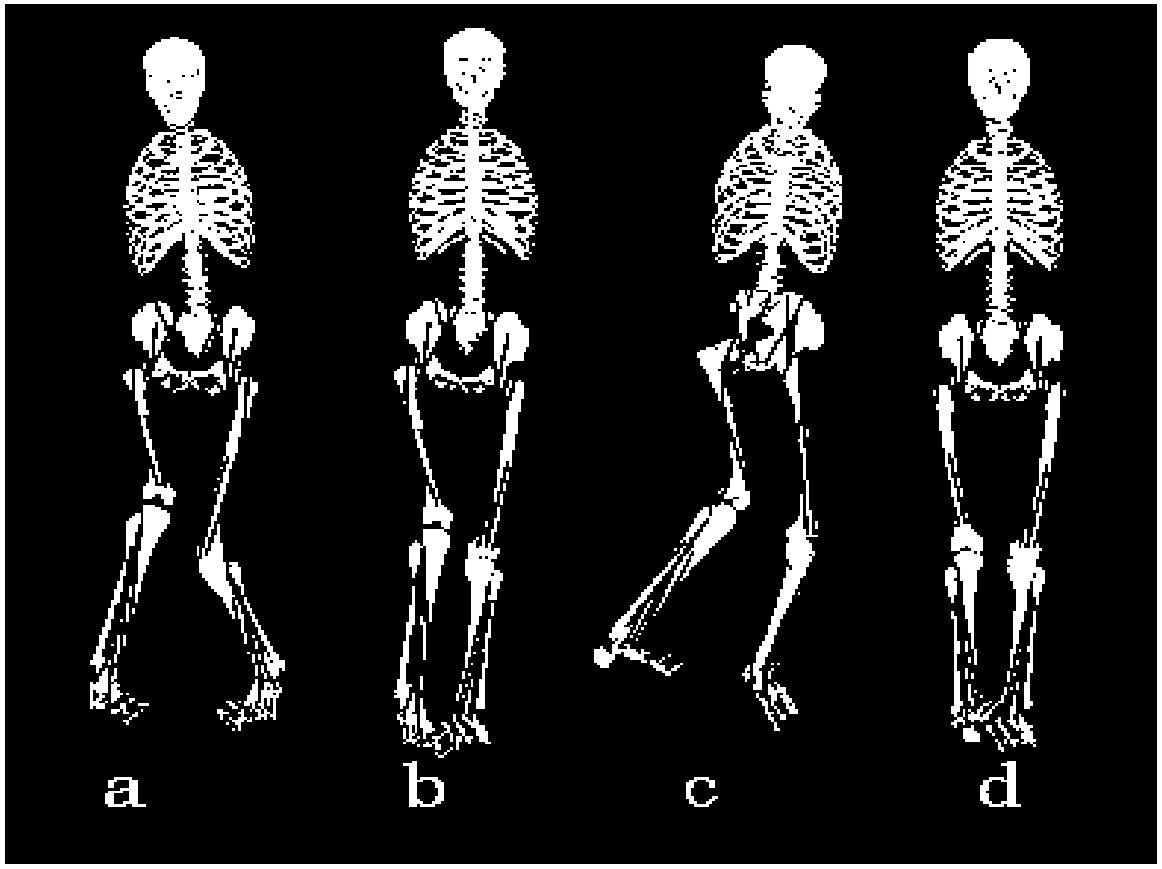
Data Collection

- **HMCD**
- To capture the human data we use HMCD. From this device different joint angles are captured to get the human GAIT data.

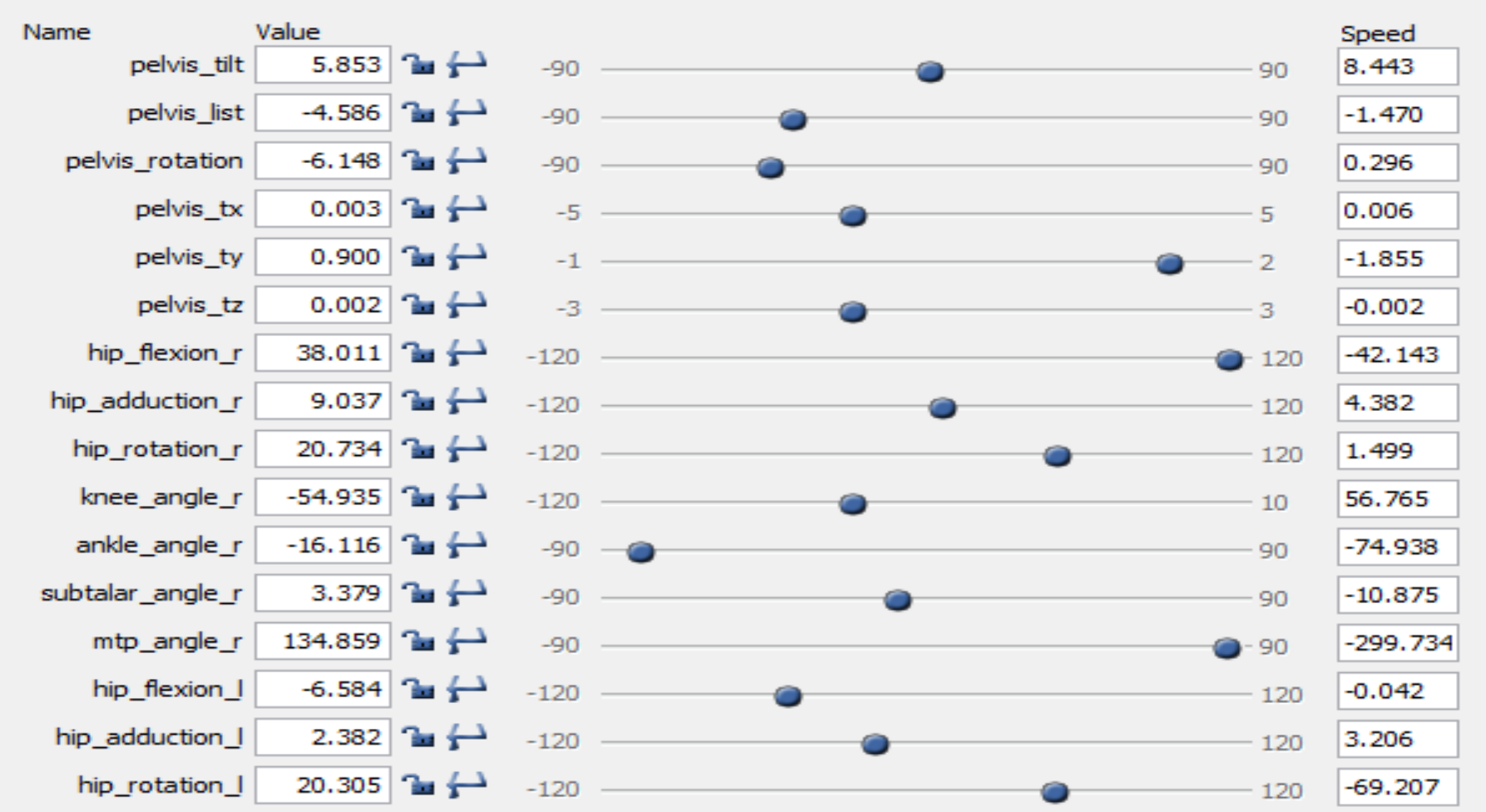
Analysis of dataset

When the dataset is simulated, then the coordinates of the human GAIT data set is captured and shown in figure for crouch2, crouch3, crouch4, normal datasets respectively.

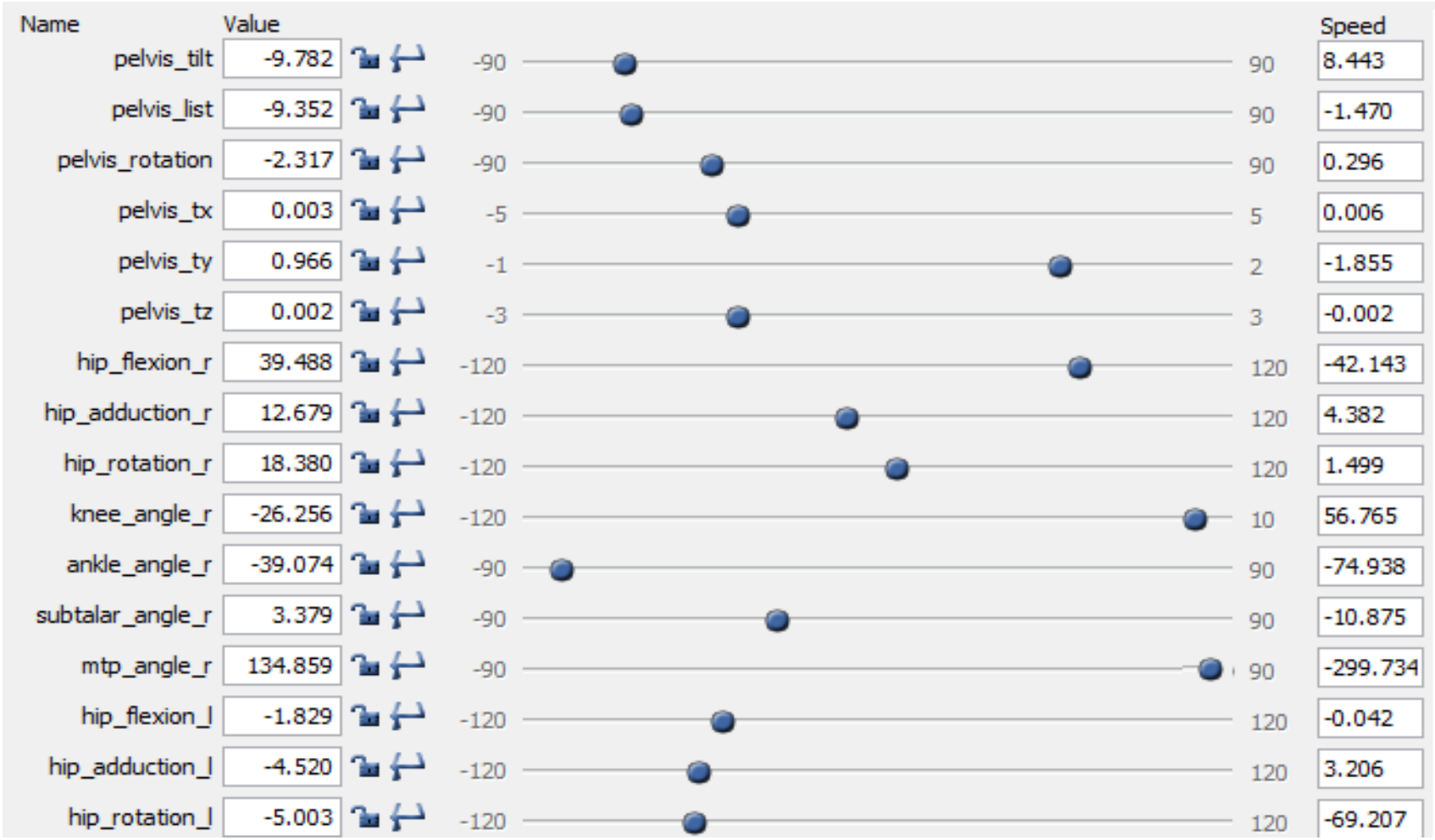
Simulation of dataset in openSim



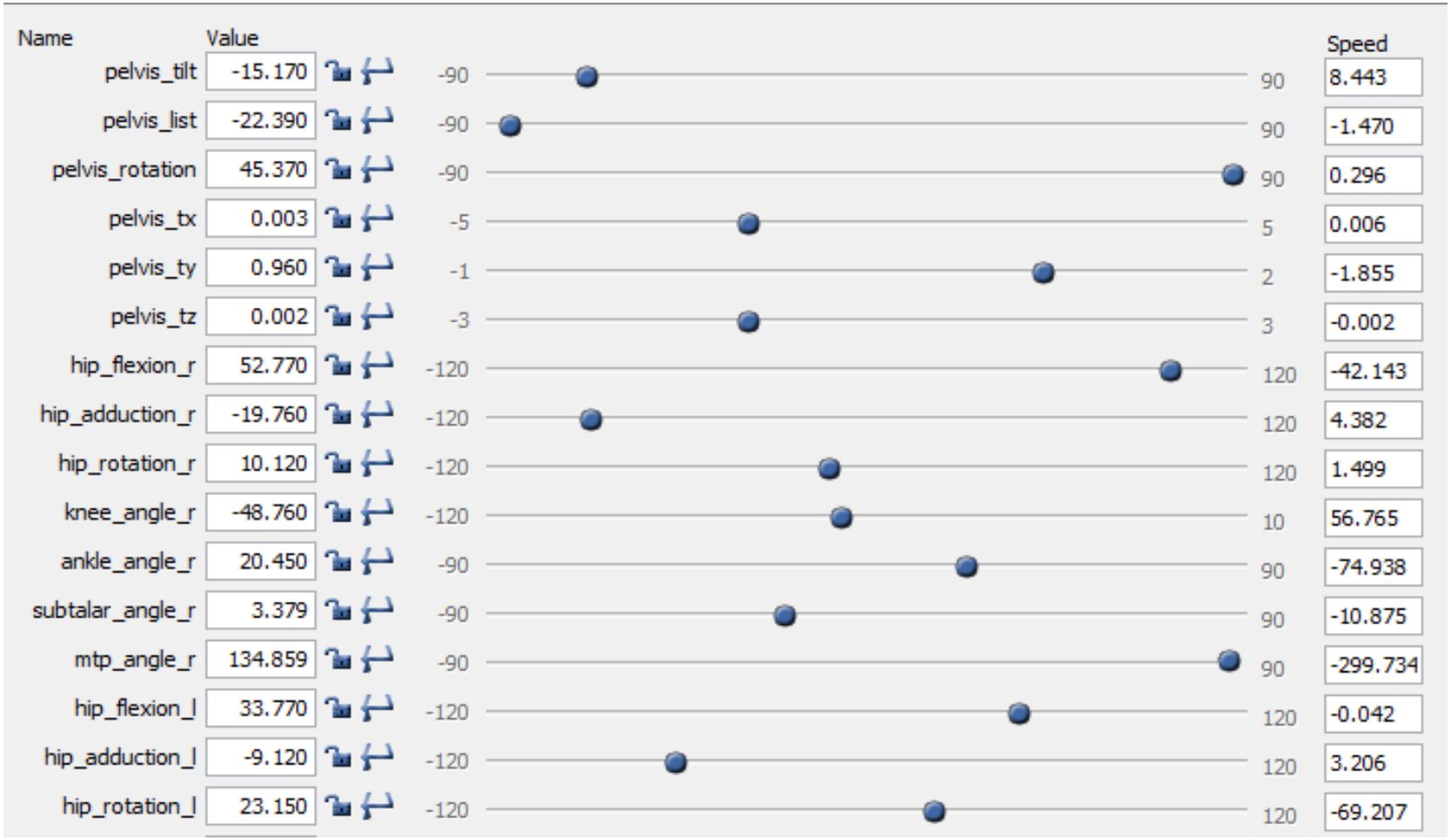
Values of crouch2 dataset



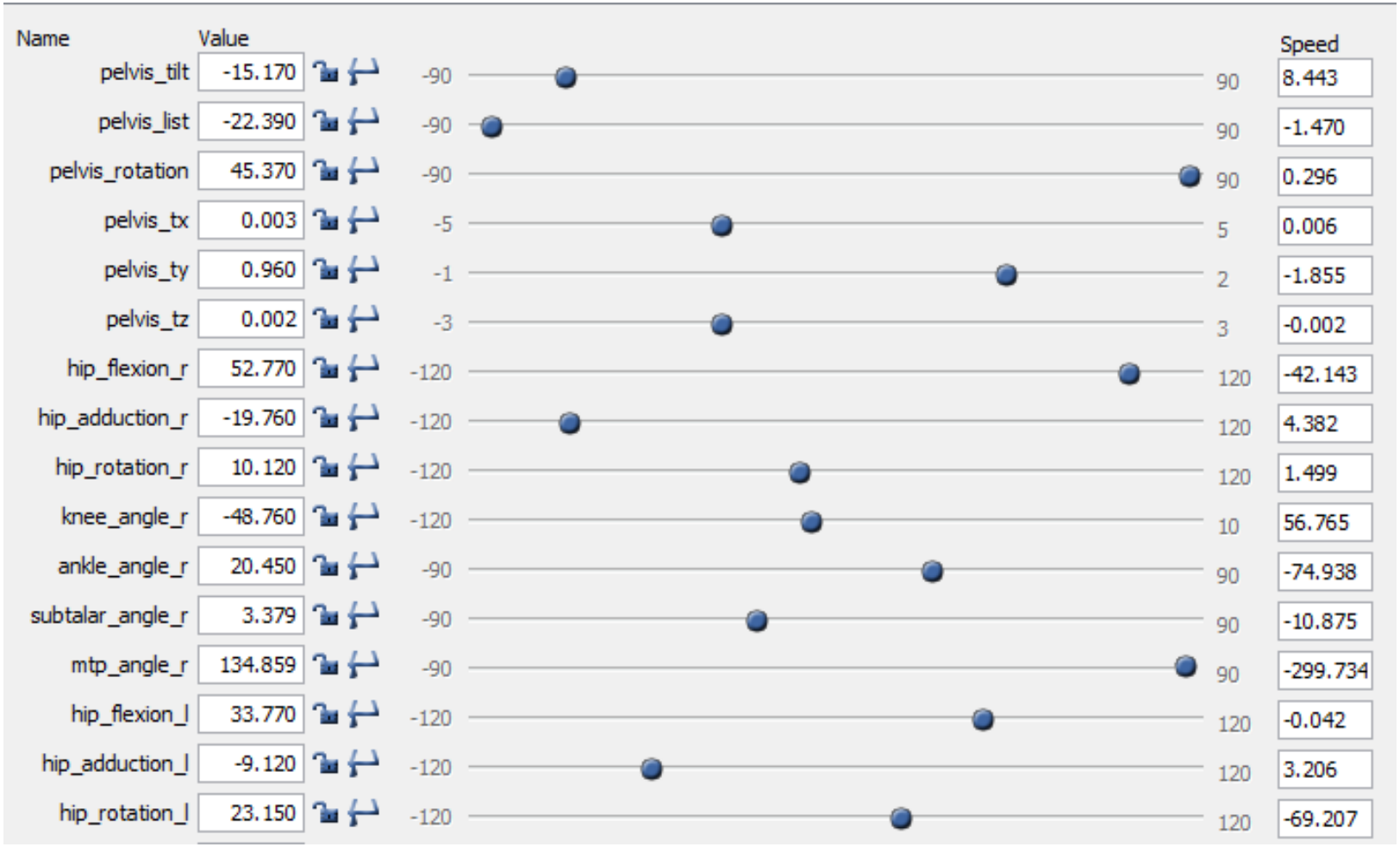
Values of crouch3 dataset



Values of crouch4 dataset



Values of normal dataset



Feature Selection

Feature Category	Feature Name
F1	Pelvis_tilt
F2	Pelvis_list
F3	Pelvis_rotation
F4	Pelvis_tx
F5	Pelvis_ty
F6	Pelvis_tz
F7	hip_flexion_r
F8	hip_adduction_r
F9	hip_rotation_r
F10	knee_angle_r
F11	ankle_angle_r
F12	ankle_angle_l
F13	knee_angle_l
F14	hip_flexion_l
F15	hip_adduction_l
F16	hip_rotation_l

CLUSTERING

K-Mean, SOM

CLUSTERING

- CLUSTERING is a collection of different methods for categorizing data.
- To correctly cluster the data one wishes to find features that separate each group of data into clusters.

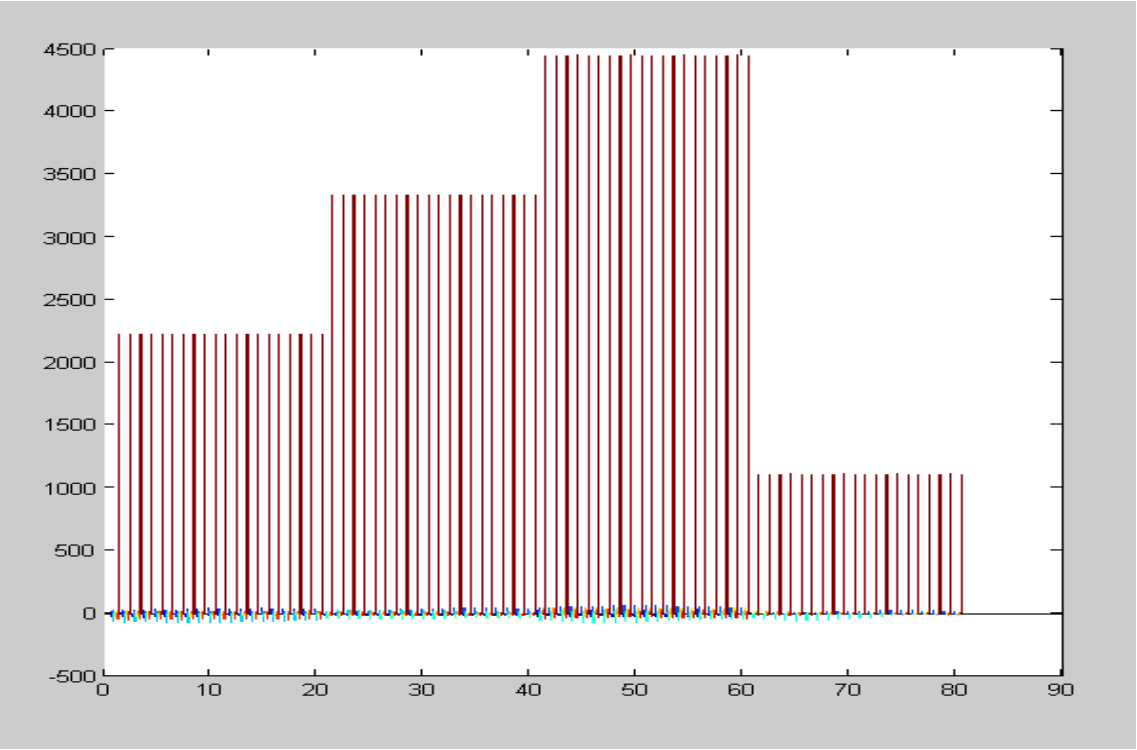
CLUSTERING

- CLUSTERING of GAIT data is done using two following techniques:
- K-Mean
- SOM

K-Mean

- The k-mean algorithm is generally used for compact clustering.
- It is very much sensitive for outlier and noise.
- It is applicable only for numerical data.

K-Mean Result

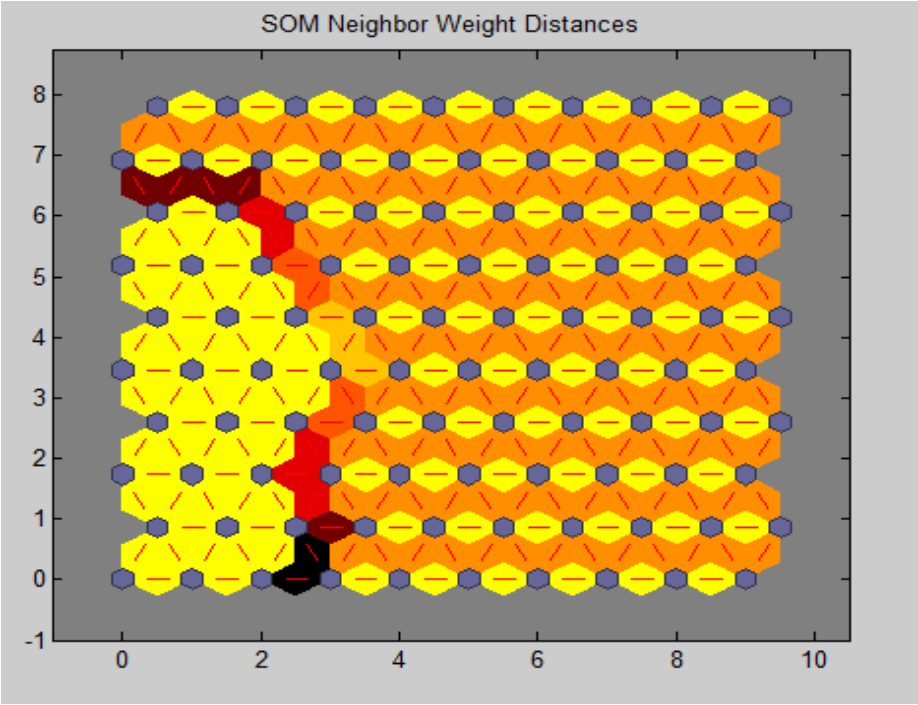


K-Mean Result

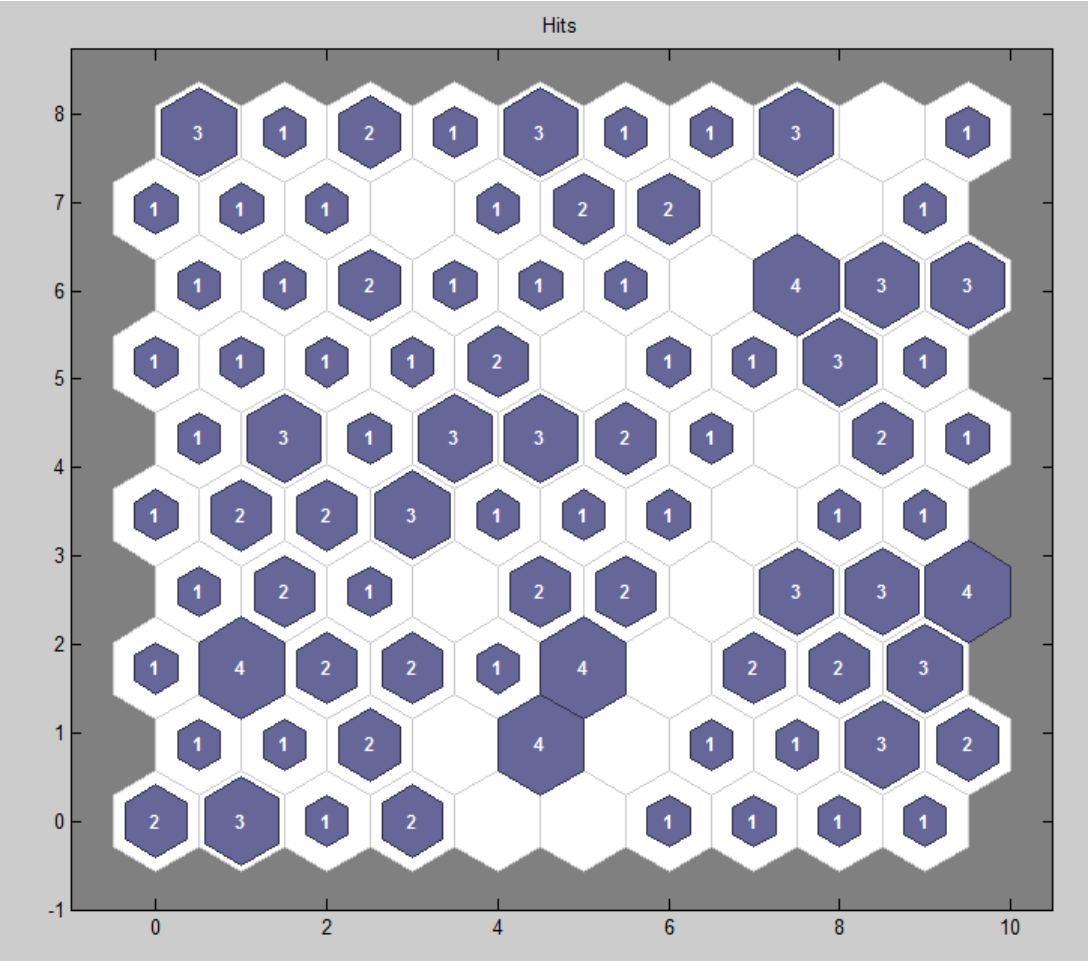
Exp. No	Training accuracy (%)	Test accuracy (%)
1	100	80

SOM

- SOM is unsupervised learning method.
- SOM is used to map the training samples and thus is known as self-organising maps



Cluster of Dataset



Result

Exp No.	Training accuracy (%)	Test accuracy (%)
1.	100	71

Comparison of success rate by different algorithms using 10–fold cross validation

Classifier	Rate of Success (%)
K-Mean	80
SOM	71

Correlation coefficient	0.9955
Mean absolute error	8.2394
Root mean squared error	117.3938
Relative absolute error	0.741 %
Root relative squared error	9.4381 %
Total Number of Instances	203

Conclusion

- The proposed system can be used for developing a gesture controlled hexadecimal keyboard making human-computer interaction easier, surveillance scenarios, in detecting disease of sports person and gait recognition system.
- After this, 10-fold cross validation was used to provide overall accuracy of the system.
- Overall accuracy was observed to be 90.23, 85.50% using the K-Mean and SOM.

Questions?

Thank you

Have a great day!

References

[1] Mark, Afolabi, and Olatoyosi Olude. Predicting NurnberofComponents Organizing Map (SOM). Proceedings ofthe 40th Hawaii International Conference on System Scienes, Figure 1. K-means clustering accuracy 2007.

[2] Faber, Vance. Clustering and the Continuous k machines proved to be the comparable methods for Means Algorithm. Los Alamos Science. Vol 22. this application for K-means clustering; which http://www.fas.org/sgp/othergov/doe/lanl/pubs/00412 are both fairly close values.