**INTERIM REPORT**

MACHINE LEARNING MODEL FOR 3D HAND POSTURE RECOGNITION

*(Dataset:Mo-Cap Hand Postures Data Set UCI Repository)*

Problem Statement: Developing a Machine Learning model aimed at recognition of hand postures from the 3-Dimensional data points corresponding to reflective markers on a left handed glove

**Mentor Submitted by**

Mr. Naveen Divyam Tripathi

Shreya Goyal

Shweta Chopra Shahrukh Khan

Yash Bagaria

**Premable**

**Hand Posture Recognition:**

Today Hand posture recognition is gradually becoming a well-known means of Human Computer Interaction (HCI). It is efficiently making our live simpler and effective by achieving “hands free” interaction through eliminating the need to hold or press the device. The most realistic examples can be using hand postures to operate our smart phones, to play videogames like Xbox and to operate highly secure facilities. Industries like Automotive sector, Consumer electronics sector, Gaming sector are flourishing by making use of hand posture recognition daily.

To get the domain knowledge on how these various sectors are developing their hand recognition technique, we obtained data which involved users and their different hand postures which were captured using a Vicon motion capture camera system.

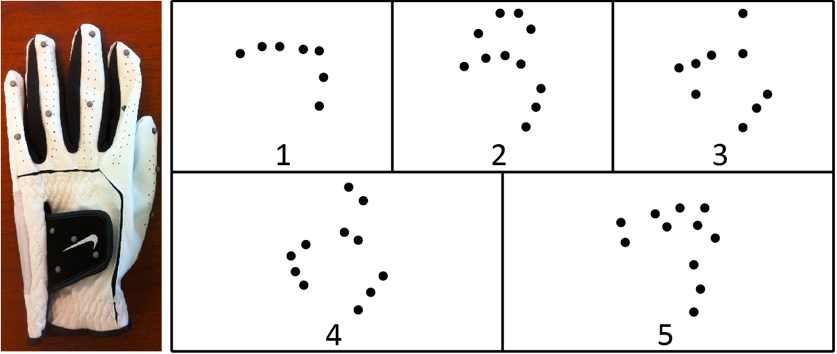
**

Figure 1: The glove used to capture data along with a sample from each class of posture projected onto the local XY plane. The classes are fist (1), stop (2), point with one finger (3), point with two fingers (4), and grab (5)

**Data Monography:**

As already mentioned A Vicon motion capture camera system was used to record 12 users performing 5 hand postures with markers attached to a left-handed glove.

A rigid pattern of markers on the back of the glove was used to establish a local coordinate system for the hand, and 11 other markers were attached to the thumb and fingers of the glove. 3 markers were attached to the thumb with one above the thumbnail and the other two on the knuckles. 2 markers were attached to each finger with one above the fingernail and the other on the joint between the proximal and middle phalanx. Their positions were not explicitly tracked.

Consequently, there is no a priori correspondence between the markers of two given records. In addition, due to the resolution of the capture volume and self-occlusion due to the orientation and configuration of the hand and fingers, many records have missing markers. Extraneous markers were also possible due to artifacts in the Vicon software’s marker reconstruction/recording process and other objects in the capture volume. As a result, the number of visible markers in a record varied considerably.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 78095 | **Area:** | Computer |
| **Attribute Characteristics:** | Real | **Number of Attributes:** | 38 | **Date Donated** | 2017-01-27 |
| **Associated Tasks:** | Classification, Clustering | **Missing Values?** | Yes | **Number of Web Hits:** | 26195 |

Dataset link:

<https://archive.ics.uci.edu/ml/datasets/Motion+Capture+Hand+Postures>

**Data Pre-Processing:**

The data available here is already pre-processed in the following ways:

* All markers were transformed to the local coordinate system of the record containing them.
* Each transformed marker with a norm greater than 200 mm was removed.
* Any record that contained fewer than 3 markers was removed and each record has at most 12 markers.

**Attributes’ Information:**

1. **Classes:**

5 types of hand postures from 12 users were recorded. Each class is recorded as an integer. The class ID of the given record ranges from 1 to 5 with :

* 1. Fist (with thumb out)
  2. Stop (hand flat)
  3. Point1 (point with pointer finger)
  4. Point2 (point with pointer and middle fingers)
  5. Grab (fingers curled as if to grab)



1. **Features:**

As there were 12 different people whose postures were recorded, the ID of the user that contributed the record is represented in the user column. Other features are X, Y, Z coordinates of all the markers used. No meaning other than as an identifier. Each record is a set.

• ‘Xi’ – Real Number. The x-coordinate of the i-th marker position. ‘i’ ranges from 0 to 11.

• ‘Yi’ - Real Number. The y-coordinate of the i-th marker position. ‘i’ ranges from 0 to 11.

• ‘Zi’ - Real Number. The z-coordinate of the i-th marker position. ‘i’ ranges from 0 to 11.

 The i-th marker of a given record does not necessarily correspond to the i-th marker of a different record. One may randomly permute the visible (i.e. not missing) markers of a given record without changing the set that the record represents. For the sake of convenience, all visible markers of a given record are given a lower index than any missing marker. A class is not guaranteed to have even a single record with all markers visible.

**METHODOLOGY:**

I Data Cleaning

II Exploratory Data Analysis

III Modelling Technique

I Data Cleaning:-

There is a large number of null values present in the dataset represented by ‘?’ symbol(markers not captured in a particular frame (row) due to occlusion were given a ‘?’ value).These ‘?’were converted to ‘0’ value because markers were not occluded from the camera.

**Percentage of Null Values:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Feature** | **%Null** | **Feature** | **%Null** | **Feature** | **%Null** |
| X0 | 0.00 | X1 | 0.00 | X2 | 0.00 |
| Y0 | 0.00 | Y1 | 0.00 | Y2 | 0.00 |
| Z0 | 0.00 | Z1 | 0.00 | Z2 | 0.00 |
| X3 | 0.88 | X4 | 3.99 | X5 | 16.67 |
| Y3 | 0.88 | Y4 | 3.99 | Y5 | 16.67 |
| Z3 | 0.88 | Z4 | 3.99 | Z5 | 16.67 |
| X6 | 33.09 | X7 | 50.13 | X8 | 60.86 |
| Y6 | 33.09 | Y7 | 50.13 | Y8 | 60.86 |
| Z6 | 33.09 | Z7 | 50.13 | Z8 | 60.86 |
| X9 | 69.31 | X10 | 81.11 | X11 | 99.96 |
| Y9 | 69.31 | Y10 | 81.11 | Y11 | 99.96 |
| Z9 | 69.31 | Z10 | 81.11 | Z11 | 99.96 |

**Inferences:-**

1. Features X8 onwards have a large number of NULL values( more than 50% ) but we will not drop these because of the vital information loss it may cause for a particular class.
2. Instead, we will replace NULL with 0, inherently transferring all the NULL data points(a trio of Xi,Yi, Zi) to origin of the local coordinate axes, nullifying the effect without any information loss.
3. Rest of the values are pre-processed and hence clean enough to perform EDA and modelling on.
4. Scaling of data was not done because each of the features was already on a similar scale.

II Exploratory Data Analysis:-

The data set was loaded in a pandas DataFrame object and a thorough investigation was done to unearth vital information about the data.Then we proceeded with the Univariate analysis on features (performed on the trio trio of Xi,Yi, Zifor a better understanding). Few of the plots generated are presented.

**Univariate Analysis:-**

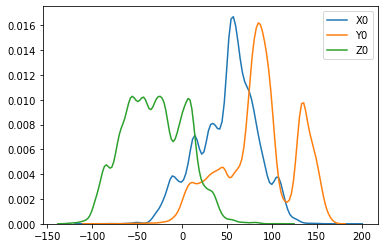


Figure 2: kde plot of the features X0, Y0 and Z0

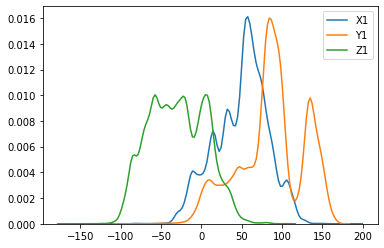


Figure 3: kde plot of the features X1, Y1 and Z1

**Inferences:-**

* The X0 feature has values in the middle range i.e. it ranges from -50 to 150.
* The Y0 feature has values in the highest range i.e. it ranges from -50 to 200.
* The Z0 feature has values in the lowest range i.e. it ranges from -150 to 100.(It has high number of negative values)
* When compared with the above graph which shows the distribution of X0,Y0,Z0 features we can come to a conclusion that (Xi,Yi,Zi) features are very similar.

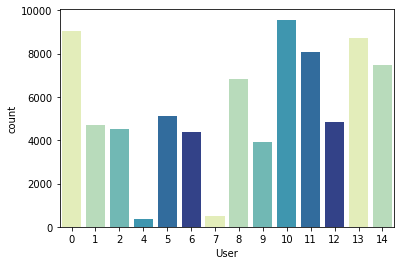
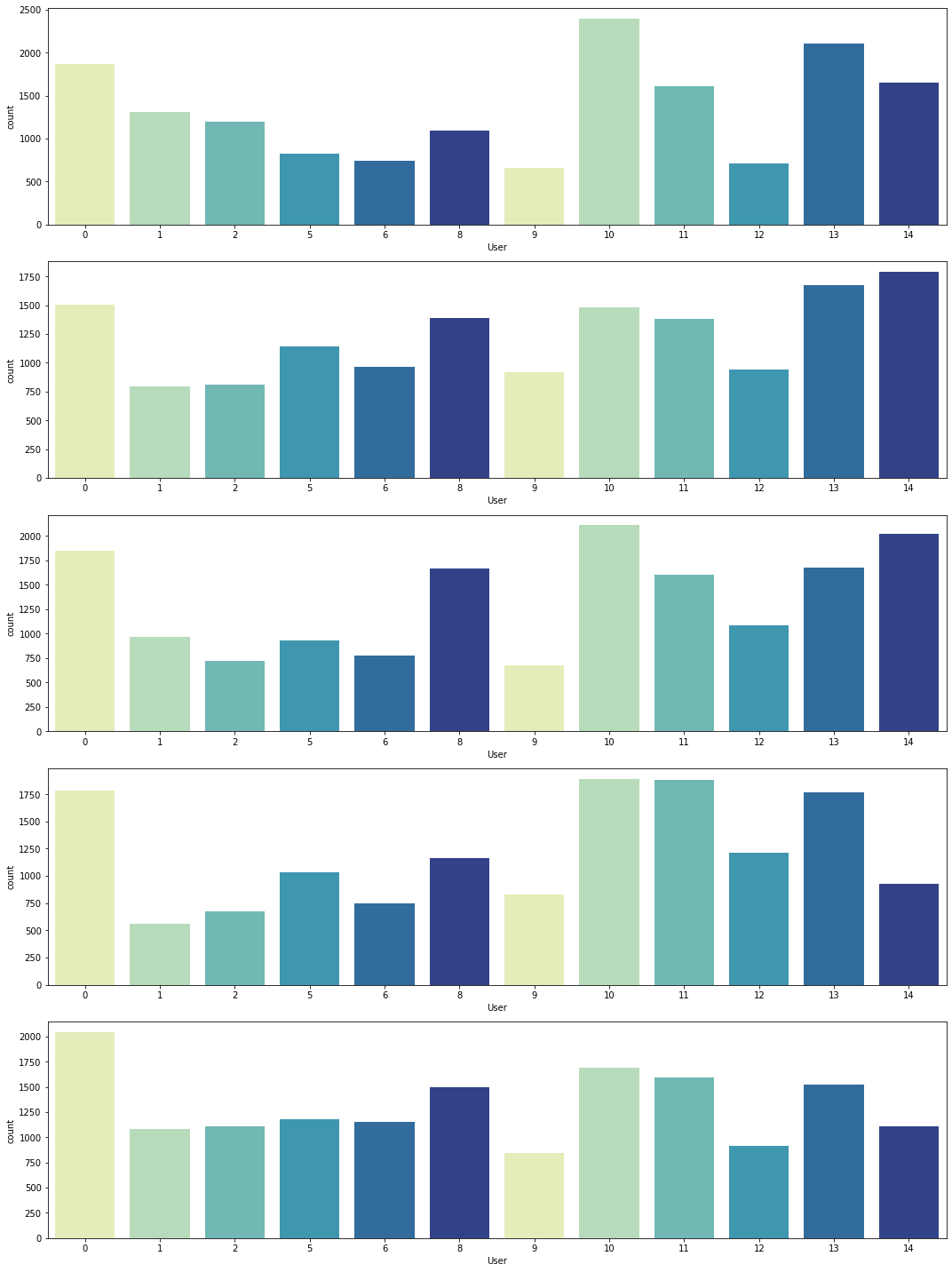
****

Figure 4: Distribution of User using Countplot

**Bivariate Analysis:-**

**** Figure 5:Class-Wise Distribution of User

**Inferences:-**

* As it is evident from the subplots above that Users 7 and 4 have only 4 postures namely 1,2,3,4 and 5th posture is missing.
* Also User 7 and User 4 have very less amount of values as compared to others.So we will drop them.

III Modelling Technique:-

Our main objective is to predict the hand postures using classification.

* Our first approach was to normally split the data into train and test.
* Then we applied Logistic Regression(One Vs Rest Technique) because our data had 5 classes so we opted for Logistic Regression(One Vs Rest Technique). We opted for Logistic Regression because our data had numerical values also it is always advised to start model building using simple models.
* After trying the Logistic Regression model we went for Decision Tree because of its capability to handle both unordered and unlabelled data efficiently. Here we splitted the data using normal train and test split.
* We used f1\_score as our evaluation parameter because it is multi-class classification technique.On using the above two approaches we find that we get considerably good score. But it raises a point that if we do train-test split the traditional way then it might happen that when our model is tetsed on an unknown set it might not perform well.

* In-order to overcome the above issue we decided to do the train-test split on the basis of one-user out and tested our model by using Logistic Regression(One Vs Rest Technique) and Decision Tree algorithms.On comparing the f1\_score we found that our data was performing well on one-user out basis and decision tree gave more efficient results. So we decided to keep Decision Tree as our base model.
* Another approach we followed for modelling was by transforming the dataset.
* We created 12 records in transformed data set for 1 row in original data set.
* We captured X,Y,Z coordinates along with respective class and row\_label .
* We trained our model and made predictions on transformed data set .
* Then for each 12 records in transformed data set we took voting measure(mode) for final prediction of original dataset but it was giving unsatisfactory results .
* So we stick up with our last approch of leave one user out with missing value imputation as 0.

Future Scope:-

* Feature Engineering:-
* The future scope of the project lies in the method where we will do feature extraction that do not depend on the order of the data.
* Moving forward in the project we might proceed with this data for posture recognition using Classification.

Relevant Papers:-

A. Gardner, J. Kanno, C. A. Duncan, and R. Selmic. 'Measuring distance between unordered sets of different sizes,' in 2014 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 2014, pp. 137-143.

A. Gardner, C. A. Duncan, J. Kanno, and R. Selmic. '3D hand posture recognition from small unlabeled point sets,' in 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC), Oct 2014, pp. 164-169.