Indian Institute of Information Technology, Sricity

BRAIN COMPUTER INTERACTION

CLASSIFICATION OF HAND MOVEMENTS USING DECISION TREE ID3

August 15, 2021

1 Introduction

The Electroencephalogram (EEG) is a noninvasive test that measures the voltage fluctuations from the ionic current within the neurons of the brain, or simply records the brain's electrical activity over a period of time.

EEG signals contain a lot of information, and by using different feature extraction techniques, this useful information and characteristics can be acquired.

2 Problem Statement

Classification of EEG data from Hands movement using ID3 decision tree and Outputting as which of motor action the user has done with the same signals.

3 Decision Tree - ID3

Decision tree is a data structure used to classify and store data, it is a top-down, recursive and divide-and-conquer approach.

- The procedure is to choose an attribute based on a different metric and split it from a larger training set into smaller training sets which finally results in the form of a tree.
- Different algorithms have been proposed to take a good control over Choosing the best attribute to be split, and splitting criteria.
- Several algorithms have been proposed for the classifying the data in the form of a tree. Here the decision tree is implemented using the "ID3" algorithm.

3.1 ID-3 Algorithm

- ID3 - Iterative Dichotomizer 3 was introduced by Quinlan(1984).
- In ID3, Entropy is used as a metric to measure how informative a node is.
- ID3 algorithm defines a measurement of a splitting called Information Gain to determine the goodness of a split
- Information Gain is calculated using the entropy calculated from both individual data points and overall data points.

- The attribute with the largest value of information gain is chosen as the splitting attribute.
- The procedure continues for splitting of other nodes using information gain.

$$E = -\sum_{i=1}^{k} p_i log_2(p_i)$$

We consider the following symbols and terminologies to define information gain, which is denoted as α

D denotes the training set at any instant
|D| denotes the size of the training set D
E(D) denotes the entropy of the training set D
The entropy of the training set D
where the training set D has c_1, c_2, ... , c_k, the k number of distinct classes and pi, 0< pi1 is the probability that an arbitrary tuple in D belongs to class c_i (i = 1, 2, ... , k).

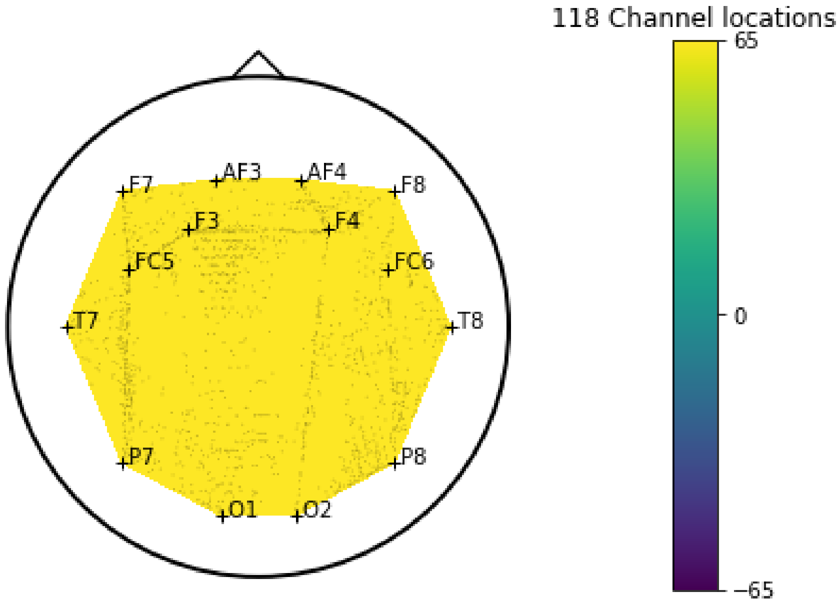
The weighted entropy denoted as $E_A(D)$ all partitions of D with respect to A is given by:

$$E_A(D) = \sum_{i=1}^{k} \frac{|D_j|}{|D|} E(D_j)$$

Information Gain: $\alpha(A.D) = E(D) - E_A(D))$

4 Data Pre-Analysis and Understanding

- EEG data has been extracted from 4 subjects namely user_a,user_b,user_c and user_d.
- For the capture of brain impulses of a subject, an EMOTIV EPOC + 14 Channel electroencephalogram from EMOTIV was used. This equipment has 128 Hz sampling frequency with 16-bit analog analog converter with 14 electrode channels: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4.



(a) Electrode placements

- EEG data has been recorded for approximately 25 mins.In that time the subject is exposed to certain visual cues that instructs him to think of an activity(right arrow to think about moving right hand,left arrow think about moving left hand)
- A circle cue represents no motor action. The only exception is the situation in which the participant is with his eyes closed, where a beep is given to the participant to open the eye only at specific times to view the images(visual cues).

5 Preprocessing

The preprocessing stage of the EEG signals is essential due to the noise present in the signal. Several filters are used to denoise the signals. This stage is followed by the stages feature extraction and classification.

As the brainwaves of interest are in the range of 0 to 30 Hz , this information collected is passed by the Fast Fourier Transform (FFT) algorithm in the frequency range 0 to 30 Hz. , resulting in a 30x14 matrix in the frequency domain.

After the transformation of the data collected for the frequency domain, the weighted and arithmetic mean of each wave was performed, so that the resulting matrix has the dimensions 14x4x2. Thus, each instant of data collected is represented by the weighted and arithmetic mean for each of the 14 device channels and the 4 wave classifications.

6 Methodology

- Downloaded data from provided reference (Kaggle).
- Identified files of 4 users given and classes of motor action as 0 - No activity, 1 - Left and 2 - Right.
- Read the data as data frame into python module.
- Extra analysis: Dropped data subjecting to 0 class - No activity.
- Data is split into Test and Train (20% and 80%)
- Extra analysis: Feature extraction - Applied Non Linear transformation.
- Extra analysis: Applied PCA
- Trained ID3 decision tree with different properties.
- Analysed optimized subject of data and Tree properties.

6.1 Data Division

For each of the 4 users, 4 individual data frames have been created to store data and to process. Data frames are the most useful modules of python to play with labelled and 2D data.

6.2 Feature Extraction

It is clear from the data provided that the features included values of standard deviation and mean of each wave and of each electrodes. The standard deviation provides an understanding of variance in data points: standard deviation = sqrt(variance). Mean is simply the average of all the data points in the particular epoch-ed data.

With this features, the relation between them can be directly identified as bringing up the Second Moments of each wave and electrodes.

$$SecondMoments = std^2 + mean^2$$

With this non-linear transformation onto data, size reduction can be brought into analysis for faster computations and performance improvement.

6.3 Classification

For the classification, Python implementation of Decision Tree - ID3 has been used onto the processed data, to understand the performance and predictions.

$$Tree = Classifier.fit(X\_train, y\_train, parameter =' entropy')$$

7 Results and Discussion

With the classification on User_A data the following analysis has been observed.

Table 1: Performance of Decision Tree under processing conditions

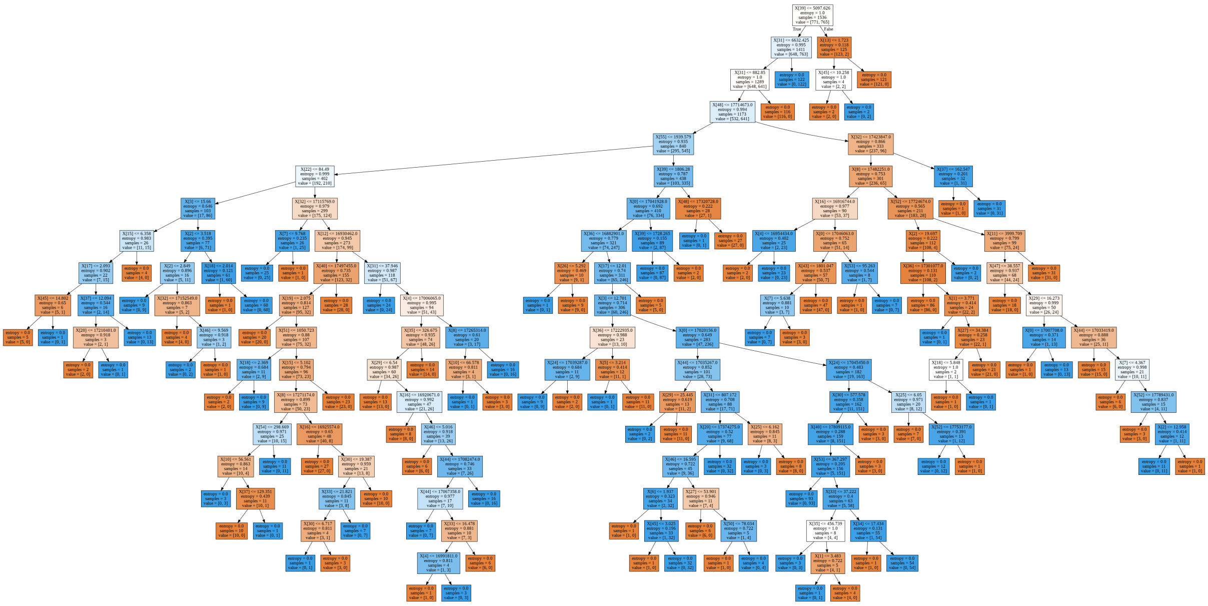
Processing technique	DataSize	Accuaracy	Comments
No processing on data	2880 x 112	64.97%	Failing performance
3 Class + Non Linear Transformation	2880 x 56	78.56%	Better performance
2 Class + Non Linear Transformation	1920 x 56	86.19%	Best
2 Class + PCA	1920 x 46	67.185%	Slightly good

From the above table, Non linear Transformation helps in reduction of Data Size which in turn leads in getting good accuracy and better performance compared to original data. So, For the accuracy of remaining users in the dataset is calculated using non linear transformation in the below table. To represent it decision trees and confusion matrices too have been analysed

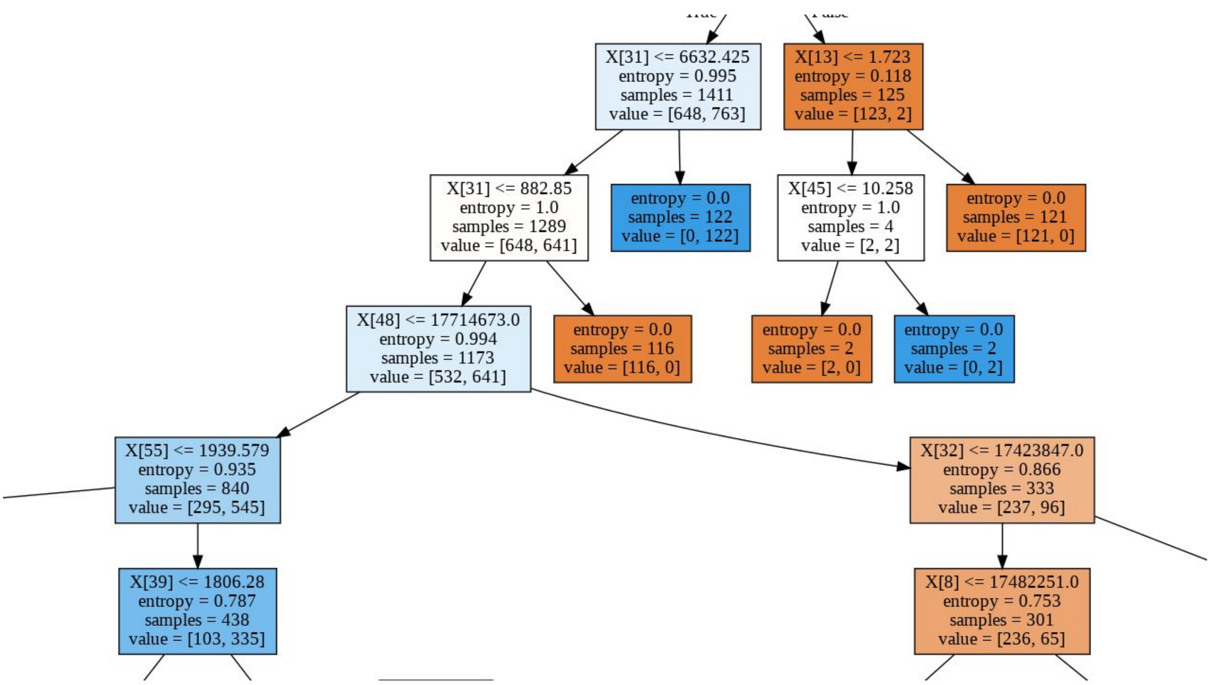
Table 2: Performance on Different users

Classification Type - User	User_a	User_b	User_c	User_d
3 Class + Non Linear Transformation	78.56%	76%	58.50%	62.67%
2 Class + Non Linear Transformation	86.19%	84.11%	67.7%	76.3%
2 Class + PCA	67.185%	72.39%	56.77%	63.28%

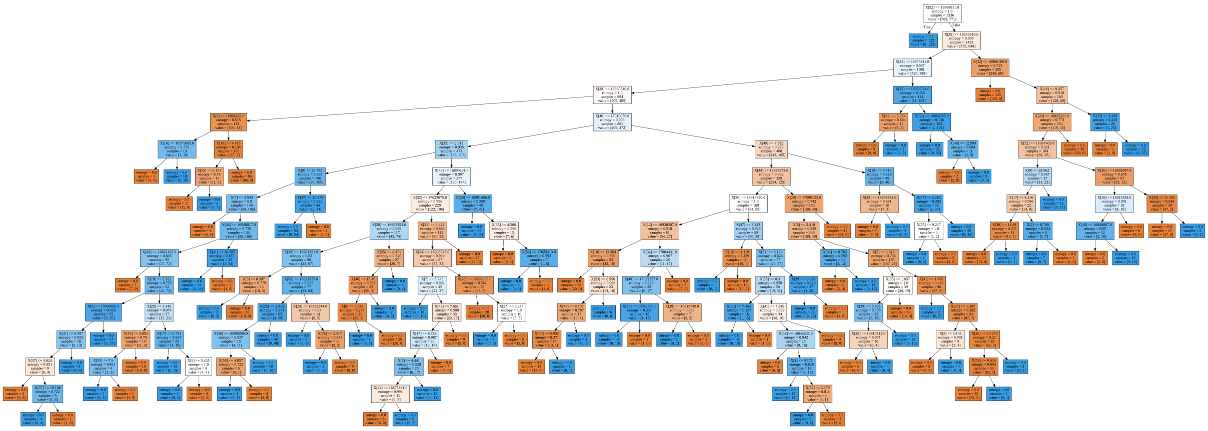
User-a and User-b have high accuracies compared to User-c and User-d using 2 Class Non Linear Transformation, and have better accuracies compared to classifiation using 3 Class Non Linear Transformation and Principal Component Analysis



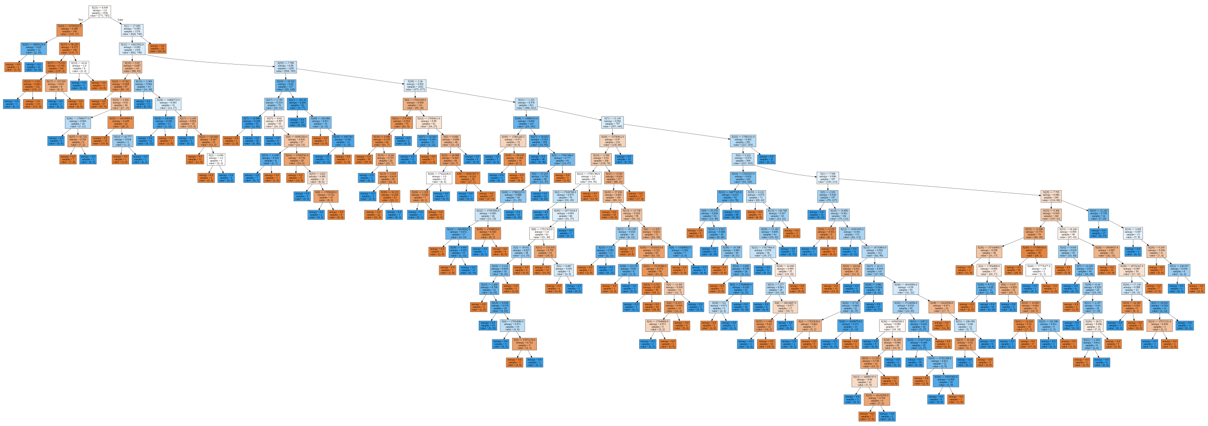
(a) ID-3 Decision Tree of User_a



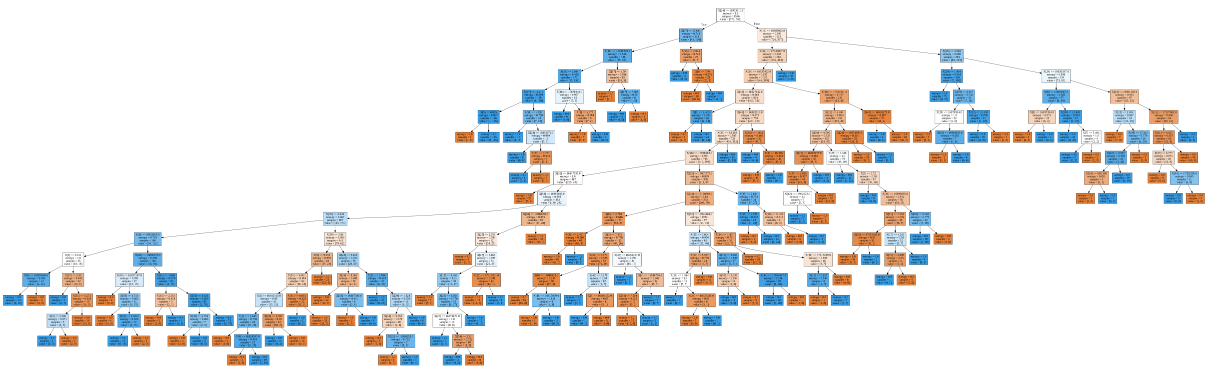
(a) Zoomed and Detailed: Mid part of ID-3 Decision Tree of User_a



(a) ID-3 Decision Tree of User_b



(a) ID-3 Decision Tree of User_c



(a) ID-3 Decision Tree of User_d

User A Confusion matrix:
([[168, 23], [30, 163]])
User B Confusion matrix:
([[168, 24], [37, 155]])
User C Confusion matrix:
([[131, 64], [60, 129]])
User D Confusion matrix:
([[152, 49], [42, 141]])

8 Conclusion

Applying some transformations to the data before classification helps in achieving good accuracy and performance. Considering the classification technique Decision Tree using ID3 algorithm, uses entropy and information gain as metrics for classification and gives the output in multi-level tree which helps in visualization of data in a better way compared to other classification techniques.

9 References

1. https://www.kaggle.com/fabriciotorquato/eeg-data-from-hands-movement
2. https://colab.research.google.com/drive/1aRI0BZHfHGGDBqB_vKqDu0pXqH_L4QpK?usp=sharing

10 Document Notes

This work is adhered to the coursework of BCI (Brain Computer Interaction) to aim at improving the existing solutions to achieve novelty in ML cases. Authors of work:
Vishnu Vamsi. B, Sai Srinivas. B, Khaleel. B, Deepak Sharma. D, Adithya. D
Authors mentioned of whose contirbution is equally divided in developing this work.