Lee Silverman Voice Treatment (LSVT) system

Mahmoud Mosaad
SBME
Cairo University
Group 7

Habiba Ahmed
SBME
Cairo University
Group 7

Noha Shafik
SBME
Cairo University
Group 7

Ayatullah Hassan
SBME
Cairo University
Group 7

Nada Galal
SBME
Cairo University
Group 7

***Abstract*—One of the main symptoms of Parkinson's disease (PD) is vocal degardation, Thus a rehabilitaion overseeing with the doctor is needed. A computer program called Lee Silverman Voice Treatment (LVST) was developed to be able to distinguish between two separate classes (Train) and (Test) and judge on the test in a binary way whether it's (Accepted) or (Not-Accepted) according to the train and it's feature selection. By using MATLAB as a machine learning tool to do the above mentioned classification, We get a 56% accuracy which is not efficient enough to be operable but it's a base for further continuing developing study.**

I. Introduction

Parkinson's disease is a neurodegenerative disorder, which leads to progressive deterioration of motor function due to loss of dopamine-producing brain cells. The cause of Parkinsons Disease is unknown but researchers speculate that both genetic and environmental factors are involved; some genes have been linked to the disease. Primary symptoms include: 1-Tremor 2-Stiffness 3-Slowness 4-Impaired balance 5-Shuffling gait later in the disease 6- Vocal cord abductor paralysis (VCAP) Some secondary symptoms include: 1-Anxiety 2-Depression 3-Dementia. vocal cord abductor paralysis (VCAP) is rare in Parkinson's disease (PD), while it is frequent in multiple system atrophy (MSA). Although VCAP is a life-threatening complication it has not yet been clarified whether there is any difference in the mechanism of VCAP between PD and MSA. Examining 3 autopsy-proven PD patients who developed severe VCAP requiring tracheostomy, we found the following differences in the mechanism of VCAP between MSA and PD: (1) clinical and laryngofiberscopic examination showed that VCAP in PD was not exacerbated during sleep, unlike in MSA; (2) On histological examination of the intrinsic laryngeal muscles, the posterior cricoarytenoid muscle demonstrated no abnormalities in PD, while the muscle showed characteristic neurogenic atrophy in MSA. There seemed to be two types of VCAP, namely the non-paralytic type observed in PD, and the paralytic type observed in MSA. Severe dysphagia requiring tube-feeding was common among PD patients who presented with VCAP. Although the relationship between VCAP and dysphagia is unknown, one should be aware of the possibility of fatal VCAP in PD patients with severe dysphagia. By using LSVT computer program, Machine learning tools (MATLAB) and some statistical techniques and properties we get to respectively measure the approximation of all of the followings: Accuracy, specificity and Sensitivity. The features acquired through the dataset that were given made it possible to estimate the required properties mentioned above, And by using Naive Bayes classifier and K-nearest Neighbors algorithm (KNN) we were able to estimate and compare between the properties and their calculation techniques to acquire the most efficient and precise results. However comparing the results of our own data handling system is not quite applicable and it's still under development and research study.

II. Material & Methods

A. DATASET

Our data is set consisted of 127 sample each one has 310 feature excluding the (Accepted) & (Not-Accepted) feature and each feature descrips a certain property measured for this specific sample. In addition to the 311 feature we have four other features descriping

$$E = \sum x^2,$$
$$P = \sum X^4,$$
$$NE = \sum -Xi * Xi - 2 + (Xi - 2)^2,$$
$$CL = \sum Xi - Xi - 1$$

Where Xi is the feature and it's number, E is the energy equation, P is the forth power equation, NE is the Nonlinear Energy equation and CL is the Curve Length equation.

We can take the gender as a feature as well to get a better efficient results.

B. PROCESSING

1) NAIVE BAYES: Through the data given in the dataset we partition the whole samples into two classes train and test by using the K-Fold technique to obtain the best accuracy possible where train is for the samples that used to build the comparison system and the test is used for testing how the developed system efficient and accurate is. The train is divided into two subclasses Accepted and Non-Accepted, Afterwords we need to calculate the desired features from all the samples of which the processing will take action upon, and in our processing technique we're using the features Energy, Power, Nonlinear Energy and the Curve Length equations, By calculating each feature we're forwarded to the statistical mapping to get some of the statistical data needed for the analysis, The mean of the four main features(E, P, NE, CL) respectively, the probability density function (PDF) of each two classes, the Accepted and the Non-Accepted, then we calculate the probability (P) of all the samples in the test class afterwords we compare the actual test with the one we obtained in the K-Fold technique, hence we got a final resultant accuracy of (46%), final sensitivity of (36%) and a specifity of (75%)

2) KNN: The KNN is easier in the implementation and easier as a processing mechanism. As the naive base we partition the the samples into two separate classes train and test and we obtain the features (E, P, NE, CL) needed from the train class to forward them to the KNN processing mechanism with respectively with every sample in the test class, Hence we get an accuracy of (53%), sensitivity of (16%) and specifity of (65%)

C. DISCTRIBTIVE CODE HANDELING

In our software code handeling we use MATLAB as our query injection to perform the desired processings on the dataset given.

1) NAIVE BAYES: To partition the samples we use K-Fold.

```
%feature with classification
All=[features classification];
CV=cvpartition(classification,'KFold', num);
```

Calculating the Specificity, Sensitivity and Accuracy.

```
[confusionMatix, order] =
    confusionmat(actual,predicted);
 %CONFUSIONMAT OUTPUT
   %%%%%%%%%%%%%%%%%%%%%%
   %       predicted
   %       | 1   0
   %   r -----------
   %   e 1 | TP FP
   %   a 0 | FN TN
   %   l
   %%%%%%%%%%%%%%%%%%%%%%
   %Specifity(o,:) = (trueNegative)/
       (trueNegative + falsePositive) ;
   %Senstivity(o,:) = (truePositive) /
       (truePositive + falseNegative) ;
   %https://en.wikipedia.org/wiki
   %/Confusion_matrix

   Accuracy(o,:) = (confusionMatix(1,1) +
       confusionMatix(2,2) )
       /(confusionMatix(1,1)+confusionMatix(1,2)
       +confusionMatix(2,1) +
       confusionMatix(2,2));
   Senstivity(o,:) = (confusionMatix(1,1))/
       (confusionMatix(1,1) +
       confusionMatix(2,1)) ;
   Specifity(o,:) = (confusionMatix(2,2)) /
       (confusionMatix(1,2) +
       confusionMatix(2,2)) ;
```

Final values.

```
accuracyFinal = sum(Accuracy)/num;
sentivityFinal = sum(Senstivity)/ num;
specifityFinal = sum(Specifity) / num;
```

2) KNN: Partitioning the samples.

```
%to divide my matrix into 2 categories (train
    and test)
output = mat2cell(features,[100 26]);
```

To begin the KNN calssification.

```
predicted = knnclassify(output{2,1} ,
    output{1,1} , classification(1:100,:));
actual = classification(101:126,:);
```

The final results.

```
Accuracy = (confusionMatix(1,1) +
    confusionMatix(2,2) ) /
    (confusionMatix(1,1)+confusionMatix(1,2)
    +confusionMatix(2,1) +
    confusionMatix(2,2));
Senstivity = (confusionMatix(1,1))/
    (confusionMatix(1,1) +
    confusionMatix(2,1)) ;
Specifity = (confusionMatix(2,2)) /
    (confusionMatix(1,2) +
    confusionMatix(2,2)) ;
```

III. RESULTS AND COMPARISON

The results we had compared with each of the two techniques are.

Results		
Specification	Naive Bayes	KNN
Accuracy	46%	53%
Sensitivity	36%	16%
Specifity	75%	65%

IV. CONCLUSION

One of the most important symptoms of the PD is vocal degardation, By using LSVT binary classification system we were able to an accuracy of 53%, sensitivity of 36% and specificity of 75% from the both used techniques (Naive Bayes) and (KNN) by comparing the testing samples with their real values, Which might not be applicable But by further development and study we could get a better accuracy reaching up to 90%

REFERENCES

[1] https://people.maths.ox.ac.uk/tsanas/Preprints/TNSRE2013.pdf
[2] https://www.ncbi.nlm.nih.gov/pubmed/8586986
[3] https://www.medicinenet.com/parkinsonsdisease/article.html
[4] http://archive.ics.uci.edu/ml/datasets/LSVT+Voice+Rehabilitation