Analysis of Breath signal in Asthmatic Subjects

By Akansh Maurya

Submitted to Dr. Prasanta Kumar Ghosh

Agenda

1. Objective
2. Data Flow
3. Initial Experiment
4. Improved Experiments
5. Exhale-Inhale: htaerb
6. Before-After Classification
7. Gender Classification
8. Conclusion and Future Work

Objectives

O1: To know which segment of breath is responsible to give maximum accuracy for classification of Asthmatic patients.

      I.   If a segment of breath gives more accuracy than full breath signal then for classification one must use that segment.

   II.   Small part of data will be stored saving storage and increasing processing speed.

O2: To know whether random chunks from breath cycle gives comparable accuracy.

      I.   If random chunk gives high accuracy than we can omit the process of manual/automatic annotation of breath boundaries.

O3: To know difference between Before broncodil and after broncodil subjects.

O4: Gender classification from breath signal.

1. **NaAsthma**: Patient that are treated from Asthma from long time
2. Before: Before Broncodil
3. After: After Broncodil
4. Unique: Either Before or after is present
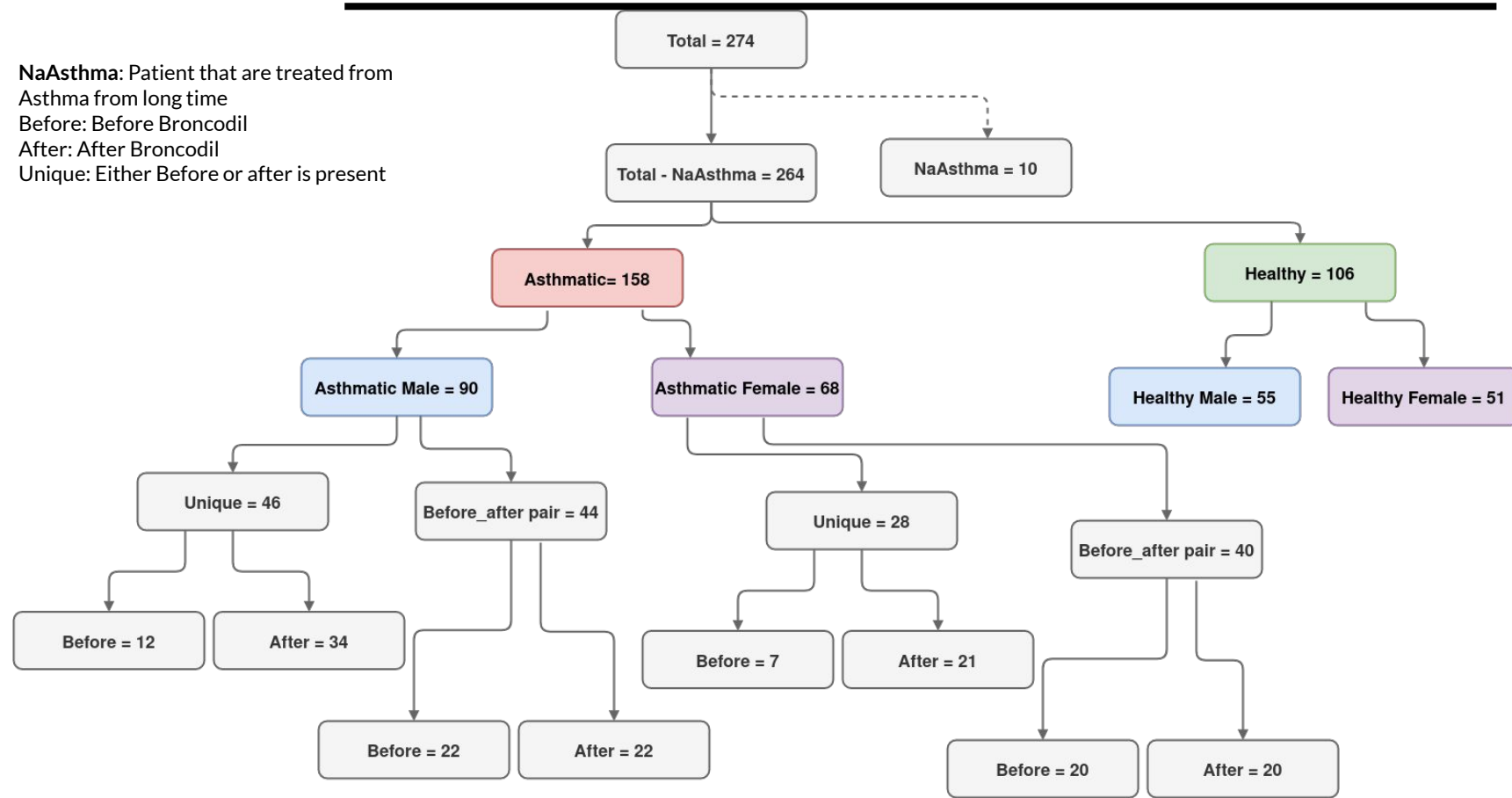
Feature Used

- Number of MFCC coefficients: 12
- Statistical features used :
  - Mean
  - Mode
  - Median
  - RMS
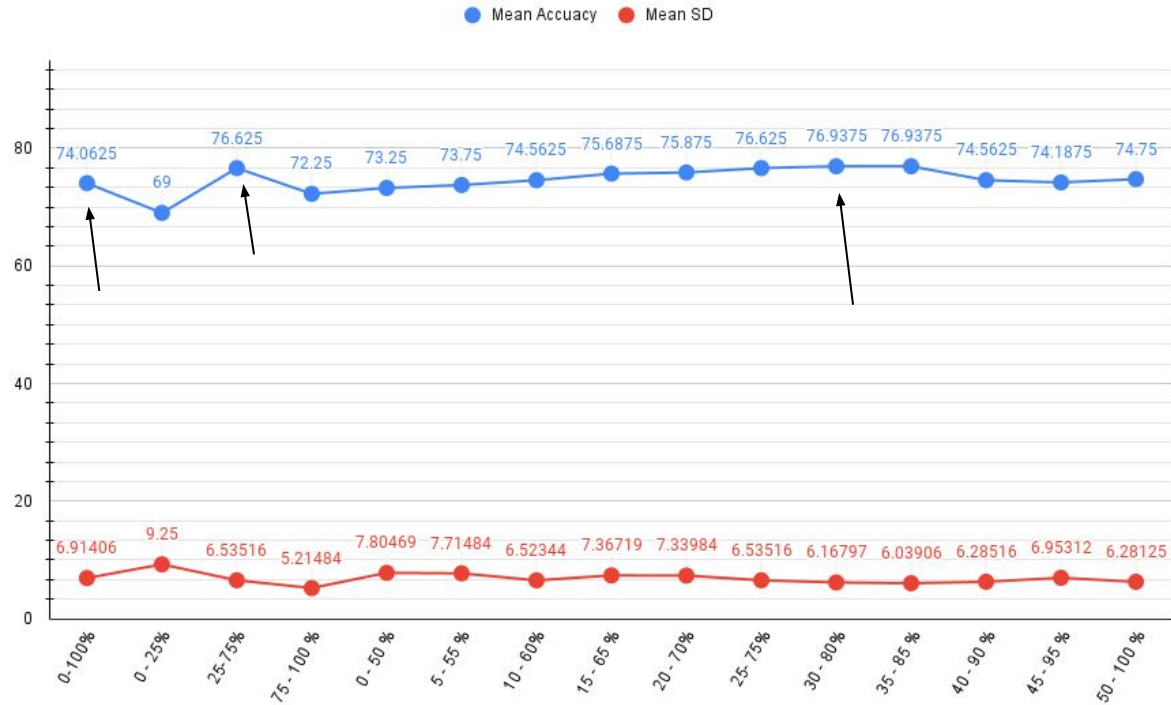 - Standard Deviation
  - Variance

O1: To know which segment of breath is responsible to give maximum accuracy for classification of Asthmatic patients.

O1.1: Experiment 1

The objective of this experiment is to find which segment of breath gives maximum accuracy.

Total Subjects	Train Set	Test set	Folds	Classifier
• Total = 93 • Asthmatic_male = 26 • Asthmatic_female = 19 • Healthy_Male = 24 • Healthy_female = 24	• Asthmatic_male = 21 • Asthmatic_female = 14 • Healthy_Male = 19 • Healthy_female = 19	• Asthmatic_male = 5 • Asthmatic_female = 5 • Healthy_Male = 5 • Healthy_female = 5	10	XGBClassifier

X-axis: Segments
Y-axis: Accuracy(%)

Conclusion of this Experiment:

- Highest accuracy is received for 30-80%, which is slightly higher than 25-75%.
- This accuracy is higher from full breath cycle that is 0-100%.
- 30-80% > 25-75% > 0-100% > 75-100% > 0-25%

O1: To know which segment of breath is responsible to give maximum accuracy for classification of Asthmatic patients.
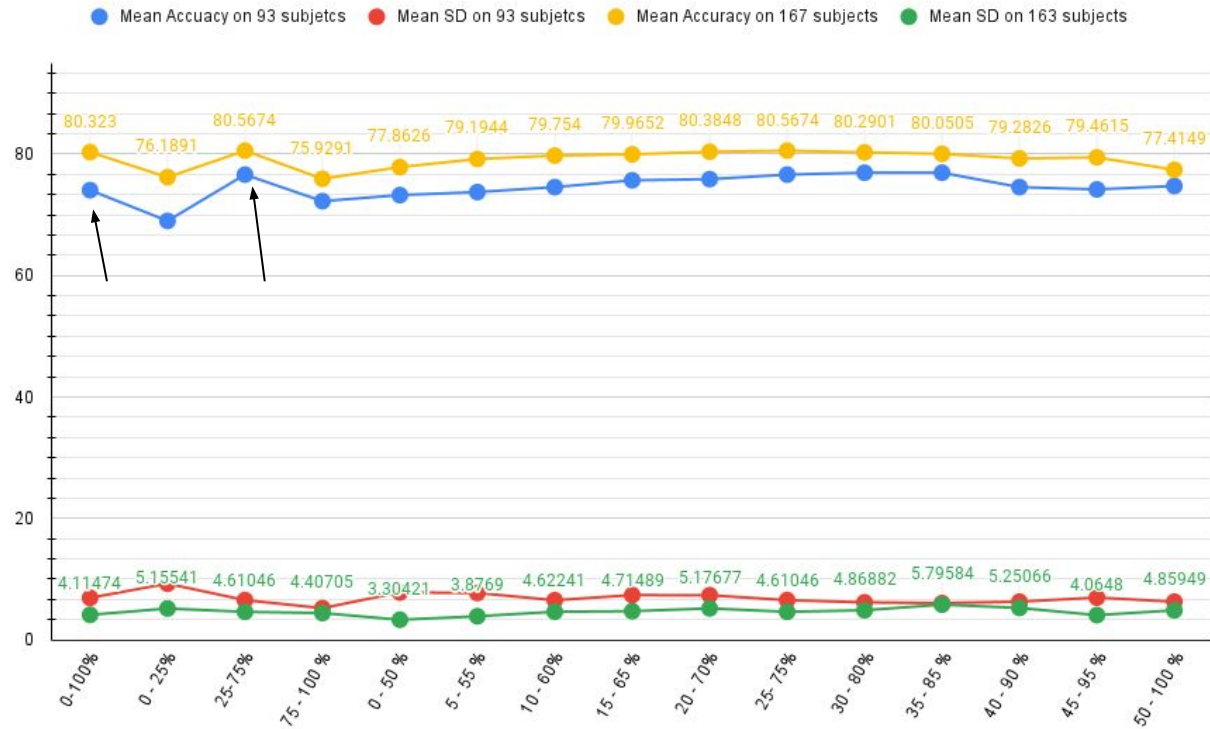
O1.2: Experiment 2

The objective of this experiment is to find whether the trend mentioned in aforesaid experiment continues when we increase the number of Datapoints.

Total Subjects	Train Set	Test set	Folds	Classifier
• Total = 167 • Asthmatic_male = 34 • Asthmatic_female = 27 • Healthy_Male = 55 • Healthy_female = 51	• Asthmatic_male = 27 • Asthmatic_female = 20 • Healthy_Male = 48 • Healthy_female = 44	• Asthmatic_male = 7 • Asthmatic_female = 7 • Healthy_Male = 7 • Healthy_female = 7	10	XGBClassifier

- Only before broncodil are present.
- Sets are not similar as previous experiments as the number of Datapoints increases, both train and test dataset size changes, test set should have 15-20% data.
- **Train Data is skewed**!

X-axis: Segments
Y-axis: Accuracy(%)

Legend: Mean Accuacy on 93 subjetcs · Mean SD on 93 subjetcs · Mean Accuracy on 167 subjects · Mean SD on 163 subjects

*Typo: 167 subjects

Yellow line values: 80.323, 76.1891, 80.5674, 75.9291, 77.8626, 79.1944, 79.754, 79.9652, 80.3848, 80.5674, 80.2901, 80.0505, 79.2826, 79.4615, 77.4149

Green line values: 4.11474, 5.15541, 4.61046, 4.40705, 3.30421, 3.8769, 4.62241, 4.71489, 5.17677, 4.61046, 4.86882, 5.79584, 5.25066, 4.0648, 4.85949

X-axis labels: 0-100%, 0-25%, 25-75%, 75-100%, 0-50%, 5-55%, 10-60%, 15-65%, 20-70%, 25-75%, 30-80%, 35-85%, 40-90%, 45-95%, 50-100%

Conclusion of this Experiment:

- Highest accuracy is received for 25-75% which is slightly higher than 0-100%.
- The trend as compared to previous experiment remains same. The accuracy for each points improves and standard deviation decreases.
- 25-75% > 0-100% > 25-75% > 75-100%

O1: To know which segment of breath is responsible to give maximum accuracy for classification of Asthmatic patients.
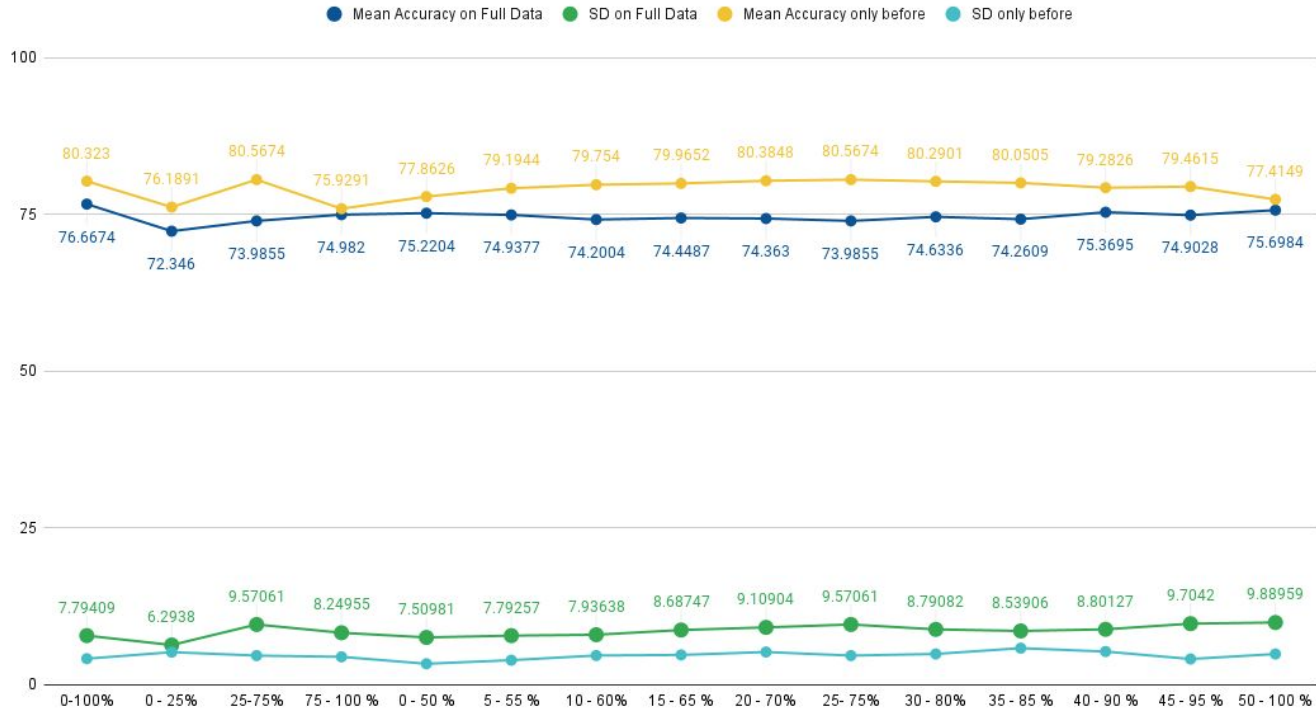
O1.3: Experiment 3

The objective of this experiment is to find whether the trend mentioned in aforesaid experiment continues when we increase the number of Datapoints. The key difference between this and previous experiment is the addition of after broncodil.

Total Subjects	Train Set	Test set	Folds	Classifier
- Total: 264 - Asthmatic_male = 90 - Asthmatic_female = 68 - Healthy_Male = 55 - Healthy_female = 51	- Asthmatic_male = 84 - Asthmatic_female = 61 - Healthy_Male = 48 - Healthy_female = 44	- Asthmatic_male = 7 - Asthmatic_female = 7 - Healthy_Male = 7 - Healthy_female = 7	10	XGBClassifier

- Sets are not similar as previous experiments as the number of Datapoints increases, both train and test dataset size changes, test set should have 15-20% data.

X-axis: Segment
Y-axis: Accuracy(%)

Conclusion of this Experiment:

- The trend does not remain same. Overall accuracy at each point decreases. Standard deviation also increases.
- Highest accuracy in this experiment is achieved for 0-100% > 75-100% .

O1: To know which segment of breath is responsible to give maximum accuracy for classification of Asthmatic patients.

O1.4: Experiment 4

The objective of this experiment is to find the effect of addition of only **After broncodil** subjects.

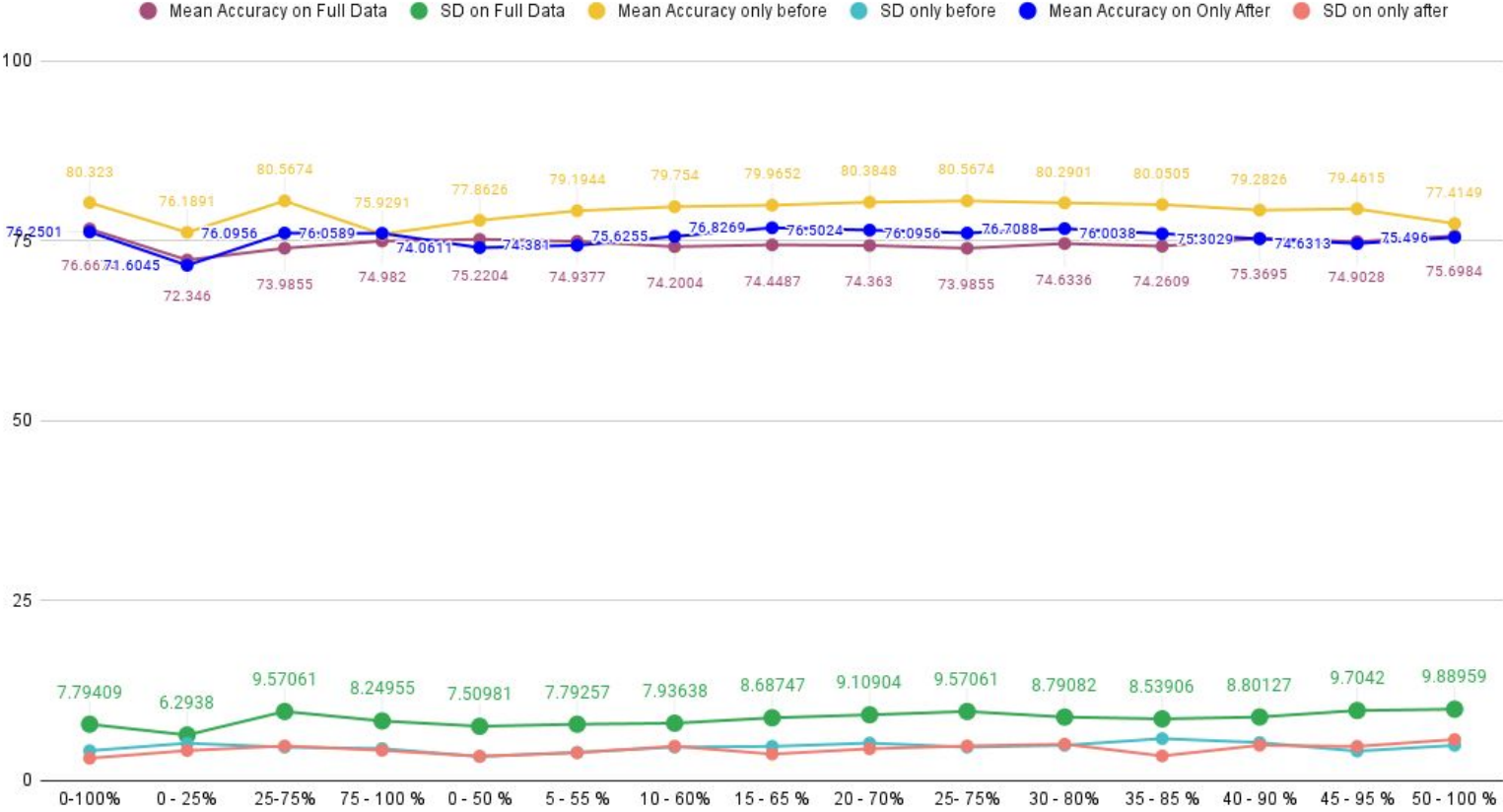| Total Subjects	Train Set	Test set	Folds	Classifier
- Total = 201 - Asthmatic_male = 55 - Asthmatic_female = 40 - Healthy_Male = 55 - Healthy_female = 51	- TRAIN SET = 173 - Asthmatic_male = 48 - Asthmatic_female = 33 - Healthy_Male = 48 - Healthy_female = 44	- Asthmatic_male = 7 - Asthmatic_female = 7 - Healthy_Male = 7 - Healthy_female = 7	10	XGBClassifier

- Sets are not similar as previous experiments as the number of Datapoints increases, both train and test dataset size changes, test set should have 15-20% data.

Conclusion of this Experiment:

- Beside having more subjects, 201 in total, the accuracy by addition of After broncodil decreases.
- The trend as shown in previous experiment is also not extended to this experiment.
- Highest accuracy is for 15-65%.
- After Broncodil Asthtmatic patients may be different from Before Broncodil Asthmatic patients. This study performed later in other experiment.

X-axis: Segment
Y-axis: Accuracy(%)

● Mean Accuracy on Full Data ● SD on Full Data ● Mean Accuracy only before ● SD only before ● Mean Accuracy on Only After ● SD on only after

O1: To know which segment of breath is responsible to give maximum accuracy for classification of Asthmatic patients.
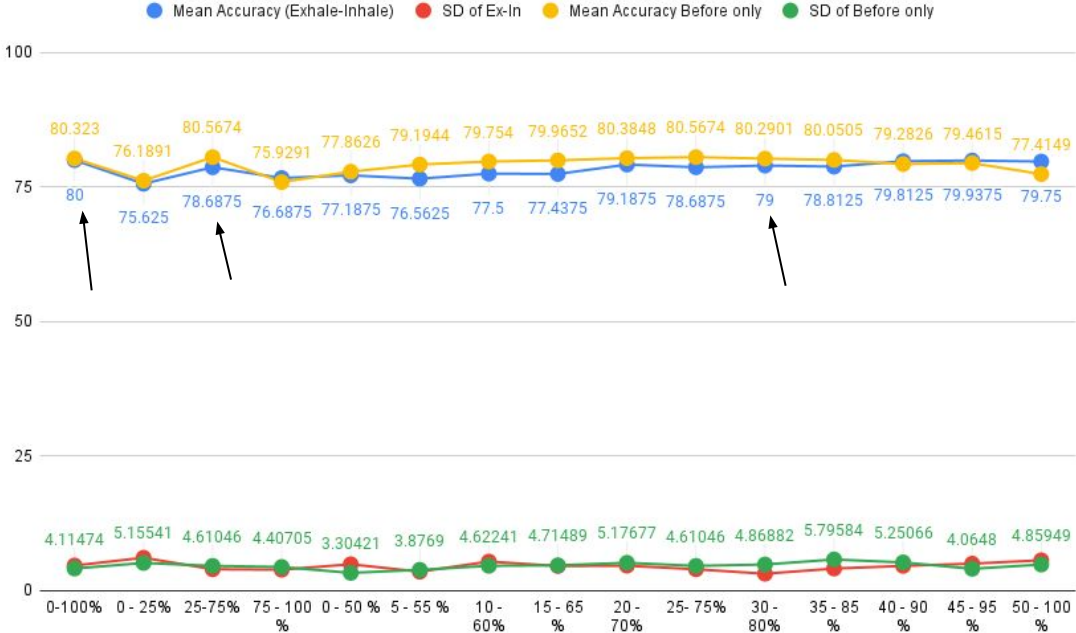
O1.5: Experiment 5

The objective of this experiment is to find the effect of transition. In previous experiments we found that middle portion i.e. 25-75% shows higher accuracy. In this experiment we will configure our data such that exhale is first than Inhale.

Total Subjects	Train Set	Test set	Folds	Classifier
- Total = 167 - Asthmatic_male = 34 - Asthmatic_female = 27 - Healthy_Male = 55 - Healthy_female = 51	- Asthmatic_male = 27 - Asthmatic_female = 20 - Healthy_Male = 48 - Healthy_female = 44	- Asthmatic_male = 7 - Asthmatic_female = 7 - Healthy_Male = 7 - Healthy_female = 7	10	XGBClassifier

- Sets are similar

X-axis: Segment
Y-axis: Accuracy(%)

Conclusion of this Experiment:

- Highest accuracy for this Experiment is for 0-100% followed by 30-80% then 25-75%.
- Trend almost looks similar.

O2: To know whether random chunks from breath cycle gives comparable accuracy.

O2.1: Experiment 1

The objective of this experiment is to find the accuracy on chunks.
Number of chunks = [10,20,30]
Length of chunks in seconds = [1,2,3,4,5]

Total Subjects	Train Set	Test set	Folds	Classifier
- Total: 264 - Asthmatic_male = 90 - Asthmatic_female = 68 - Healthy_Male = 55 - Healthy_female = 51	- Asthmatic_male = 84 - Asthmatic_female = 61 - Healthy_Male = 48 - Healthy_female = 44	- Asthmatic_male = 7 - Asthmatic_female = 7 - Healthy_Male = 7 - Healthy_female = 7	10	SVM

O2: To know whether random chunks from breath cycle gives comparable accuracy.

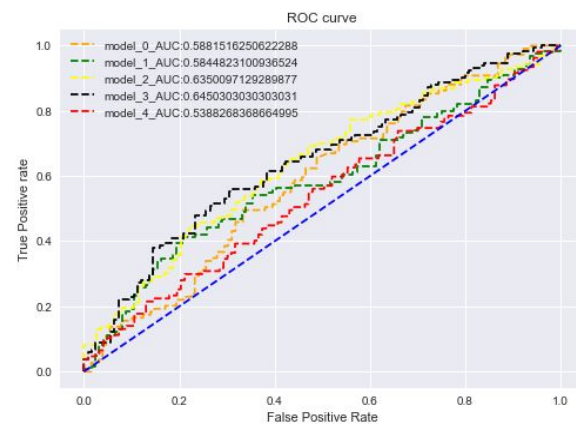| Chunk→	10	20	30
Length(in sec)			
1	71.58 ± 7.142	72.1±6.63	70.75±6.18
2	71.8±5.395	71.3±4.965	73.2±5.246
3	72.5±4.895	73.1±5.49	73.2±5.01
4	72.6±5.777	73.06±5.55	71.4±5.16
5	72.06±5.79	71.56±5.547	71.94±6.055

O3.1: Before and After Classification

The objective of this experiment is to find if the classification of Before and After Broncodil is possible or not. If the experiment results in significant accuracy then there is difference between them.
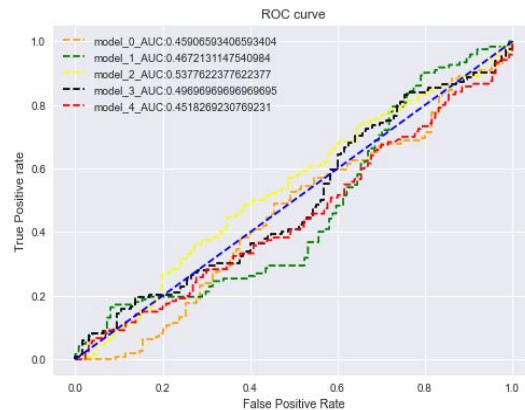
Total Subjects	Train Set	Test set	Folds	Classifier
- Total = 168 - Before = 67 - After = 101 - Male_before = 36 - Female_before = 31 - After_male = 58 - Before_female = 43	- Before = 53 - After = 87 - Male_before = 29 - Female_before = 24 - After_male = 51 - Before_female = 36	- Before_male = 7 - Before_female = 7 - After_male = 7 - After_female = 7	5	SVM

ROC curve (Exp 5.1: Breathe)
- model_0_AUC:0.4348565211478308
- model_1_AUC:0.4584583645911478
- model_2_AUC:0.6434579097491407
- model_3_AUC:0.5935757575757576
- model_4_AUC:0.5100062539086929
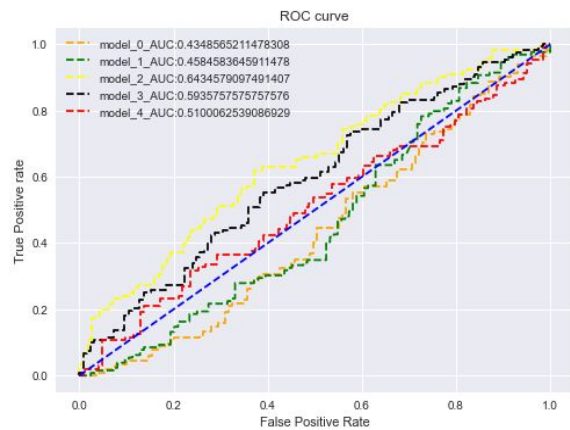
Mean Test Accuracy:52.153734603500766% with SD: 4.699313872999039%
Mean Train Accuracy:100.0%
Mean Test F1:58.248035950303986%
Mean Test AUC:52.8070961394514% with SD: 7.9329950629126635%

ROC curve (Exp 5.2: Inhale)
- model_0_AUC:0.5881516250622288
- model_1_AUC:0.5844823100936524
- model_2_AUC:0.6350097129289877
- model_3_AUC:0.6450303030303031
- model_4_AUC:0.5388268368664995

Mean Test Accuracy:53.02527280055214% with SD: 3.662549944169078%
Mean Train Accuracy:81.02364413986697%
Mean Test F1:64.31355910479478%
Mean Test AUC:59.83001575963344% with SD: 3.837264506347714%

ROC curve (Exp 5.3: Exhale)
- model_0_AUC:0.45906593406593404
- model_1_AUC:0.4672131147540984
- model_2_AUC:0.5377622377622377
- model_3_AUC:0.49696969696969695
- model_4_AUC:0.45182692307692931

Mean Test Accuracy:50.37632233290783% with SD: 3.51990199454
Mean Train Accuracy:87.65553771000558%
Mean Test F1:60.24694197115008%
Mean Test AUC:48.2567581325778% with SD: 3.158146979277402%

Exp 5.1: Breathe

Exp 5.2: Inhale

Exp 5.3: Exhale

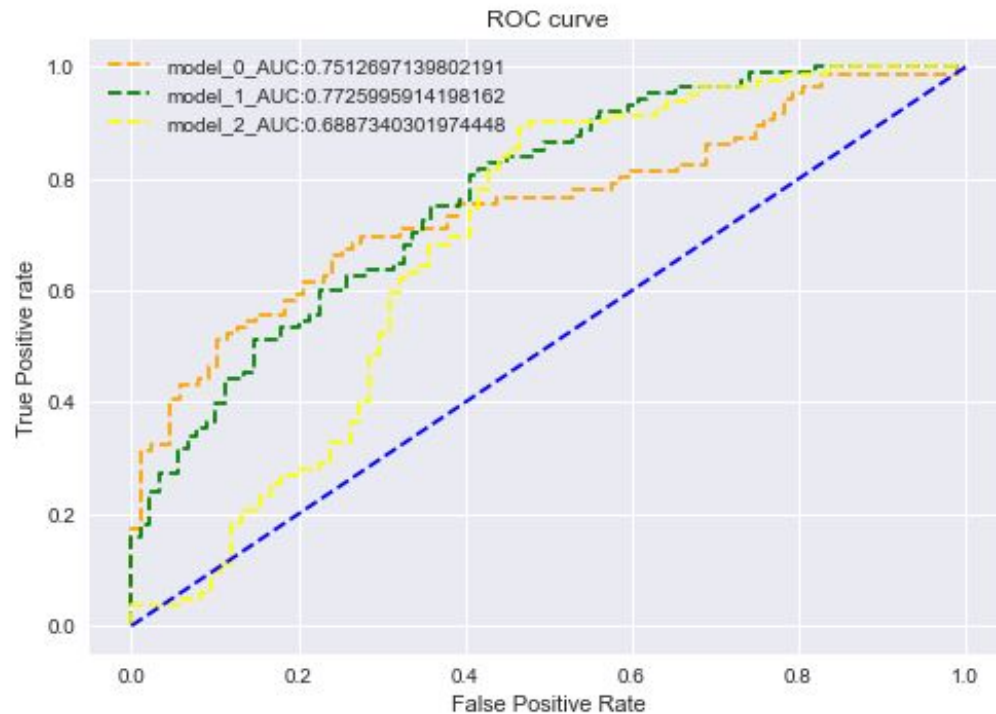| Experiment	Mean Test Accuracy (In %)	Mean F1 Score (In %)	Mean AUC (In %)
Exp 5.1: Breathe	52.15 ± 4.69	58.24± 4.767	52.83 ±7.93
Exp 5.2: Inhale	53.02 ± 3.66	64.31± 3.8477	51.83 ±3.83
Exp 5.3: Exhale	50.40 ± 3.51	60.24 ± 4.477	48.31 ±3.15

O4.1: Gender Classification[Only before]

The objective of this experiment is to find if the classification of Gender from Breath cycle. If it is not possible with significant accuracy then there is no requirement of splitting data according to gender.

Total Subjects	Train Set	Train set	Folds	Classifier
- Total = 67 - Asthmatic_male = 36 - Asthmatic_female = 31	- TRAIN SET = 47 - Asthmatic_male = 26 - Asthmatic_female = 21	- Asthmatic_male = 10 - Asthmatic_female = 10	3	SVM

ROC curve

model_0_AUC:0.7512697139802191
model_1_AUC:0.7725995914198162
model_2_AUC:0.6887340301974448

Mean Test Accuracy:67.80425644316595% with SD: 1.1020899474
Mean Train Accuracy:99.83837220236977%
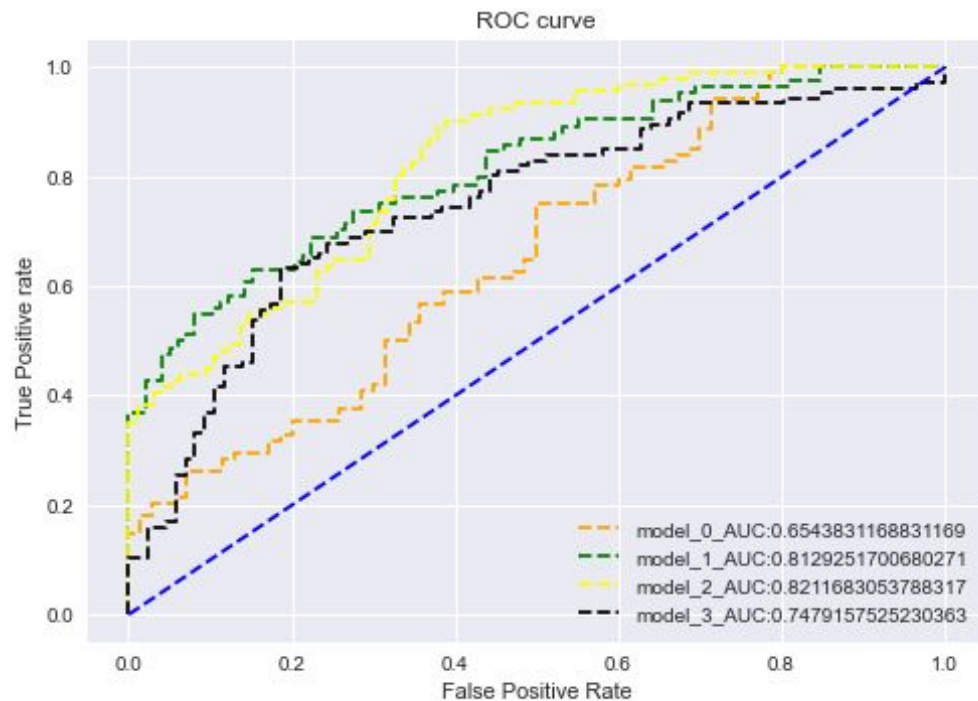Mean Test F1:70.01633986928105%
Mean Test AUC:73.753444519916%

O4.2: Gender Classification[Only after]

The objective of this experiment is to find if the classification of Gender from Breath cycle. If it is not possible with significant accuracy then there is no requirement of splitting data according to gender.

Total Subjects	Train Set	Test set	Folds	Classifier
• Total = 101 • Asthmatic_male = 58 • Asthmatic_female = 43	• TRAIN SET = 81 • Asthmatic_male = 48 • Asthmatic_female = 33	• Asthmatic_male = 10 • Asthmatic_female = 10	4	SVM

ROC curve

model_0_AUC:0.6543831168831169
model_1_AUC:0.8129251700680271
model_2_AUC:0.8211683053788317
model_3_AUC:0.7479157525230363

Mean Test Accuracy:65.8744734151788% with SD: 3.18191171577
Mean Train Accuracy:96.62077255626129%
Mean Test F1:69.41178635550256%
Mean Test AUC:75.9098086213253%

O4.3: Gender Classification[All Asthmatic]

The objective of this experiment is to find if the classification of Gender from Breath cycle. If it is not possible with significant accuracy then there is no requirement of splitting data according to gender.

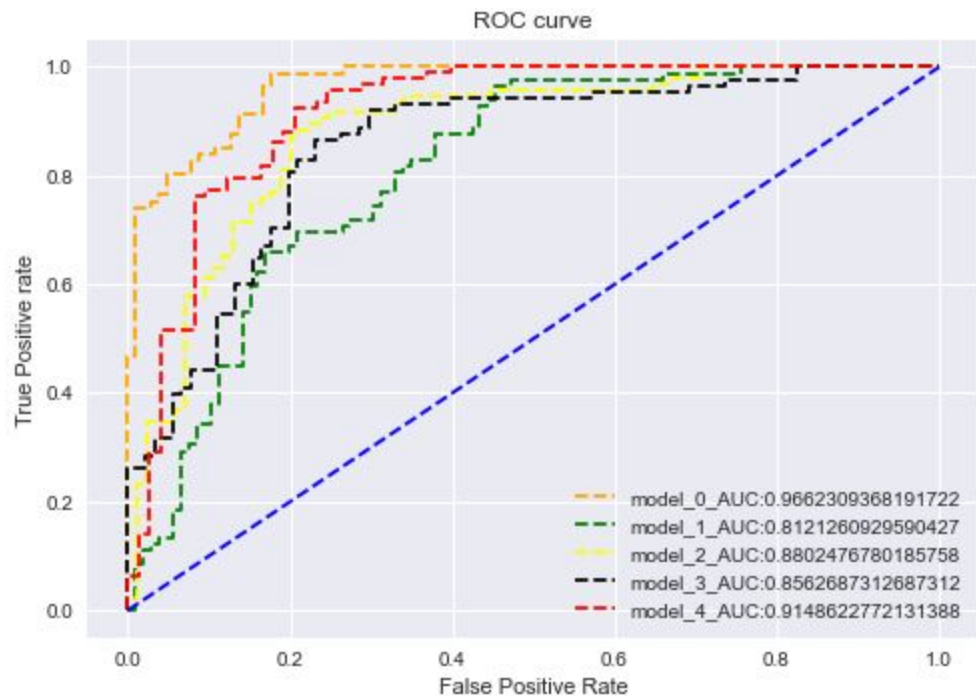| Total Subjects	Train Set	Test set	Folds	Classifier
• Total = 168 • Asthmatic_male = 94 • Asthmatic_female = 74	• TRAIN SET = 138 • Asthmatic_male = 79 • Asthmatic_female = 59	• TEST SET = 30 • Asthmatic_male = 15 • Asthmatic_female = 15	4	SVM

ROC curve
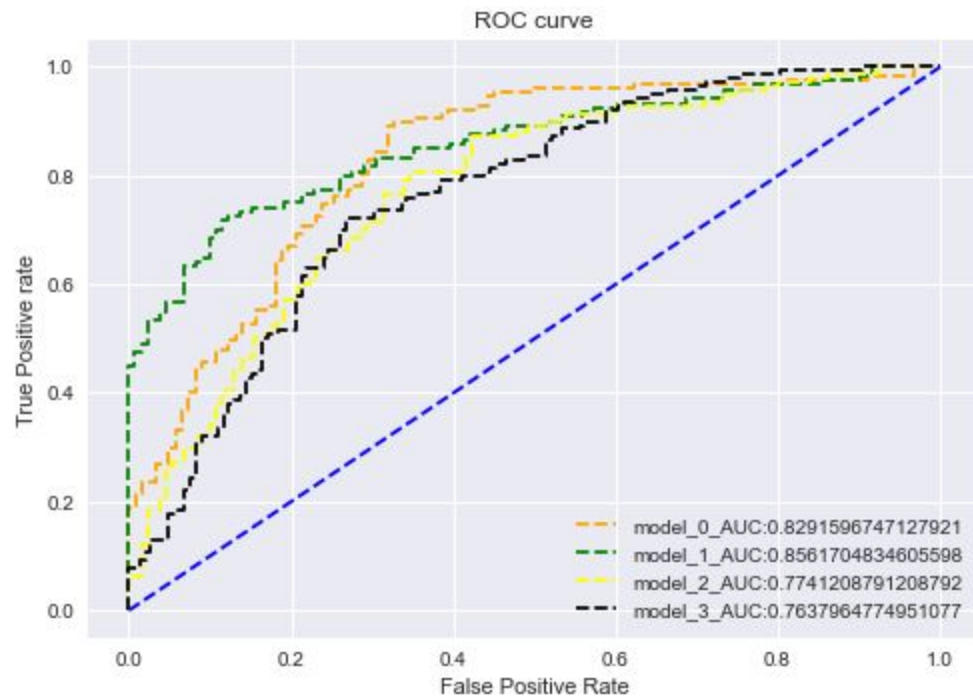
Mean Test Accuracy:72.51276853374486% with SD: 3.0402447500⟨
Mean Train Accuracy:95.59639019536252%
Mean Test F1:74.81885386495829%
Mean Test AUC:80.58118786973345%

O4.4: Gender Classification[All Healthy]

The objective of this experiment is to find if the classification of Gender from Breath cycle. If it is not possible with significant accuracy then there is no requirement of splitting data according to gender.

Total Subjects	Train Set	Test set	Folds	Classifier
• Total = 106 • Healthy_Male = 55 • Healthy_female = 51	• TRAIN SET = 86 • Healthy_Male = 45 • Healthy_female = 41	• TEST SET = 20 • Healthy_Male = 10 • Healthy_female = 10	5	SVM

ROC curve

True Positive rate / False Positive Rate

model_0_AUC:0.9662309368191722
model_1_AUC:0.8121260929590427
model_2_AUC:0.8802476780185758
model_3_AUC:0.85626873312687312
model_4_AUC:0.9148622772131388

Mean Test Accuracy:81.11590078173225% with SD: 6.13160215286
Mean Train Accuracy:94.76119244484835%
Mean Test F1:81.38861082513887%
Mean Test AUC:88.59471432557321%

O4: Gender classification from breath signal.

Experiment	Data Distribution	Mean Accuracy (In %)	Mean F1(In %)	Mean AUC(In %)
Only before	T:67 M:36 F:31	65.80 ± 1.10	70.01 ± 10.291	73.75 ± 9.40
Only After	T:101 M:58 F:43	69.36 ± 3.98	73.19 ± 4.412	77.43 ± 4.72
All Asthmatic	T:168 M:94 F:74	72.12 ± 3.04	74.81 ± 5.014	80.58 ± 9.96
All Healthy	T:106 M:55 F:51	81.15 ± 6.13	81.38 ± 4.23	88.59 ± 7.49

T: Total M:Male F: Female

Thank you!