

FannyLo_SVM

April 10, 2023

0.0.1 Importing Libraries

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import sys
```

0.0.2 Reading in Data

```
[2]: data=pd.read_csv('fetal_health.csv')
data
```

```
[2]:
```

| | baseline | value | accelerations | fetal_movement | uterine_contractions | \ |
|------|----------|-------|---------------|----------------|----------------------|---|
| 0 | | 120.0 | 0.000 | 0.000 | 0.000 | |
| 1 | | 132.0 | 0.006 | 0.000 | 0.006 | |
| 2 | | 133.0 | 0.003 | 0.000 | 0.008 | |
| 3 | | 134.0 | 0.003 | 0.000 | 0.008 | |
| 4 | | 132.0 | 0.007 | 0.000 | 0.008 | |
| ... | | ... | ... | ... | ... | |
| 2121 | | 140.0 | 0.000 | 0.000 | 0.007 | |
| 2122 | | 140.0 | 0.001 | 0.000 | 0.007 | |
| 2123 | | 140.0 | 0.001 | 0.000 | 0.007 | |
| 2124 | | 140.0 | 0.001 | 0.000 | 0.006 | |
| 2125 | | 142.0 | 0.002 | 0.002 | 0.008 | |

| | light_decelerations | severe_decelerations | prolongued_decelerations | \ |
|------|---------------------|----------------------|--------------------------|---|
| 0 | 0.000 | 0.0 | 0.0 | |
| 1 | 0.003 | 0.0 | 0.0 | |
| 2 | 0.003 | 0.0 | 0.0 | |
| 3 | 0.003 | 0.0 | 0.0 | |
| 4 | 0.000 | 0.0 | 0.0 | |
| ... | ... | ... | ... | |
| 2121 | 0.000 | 0.0 | 0.0 | |
| 2122 | 0.000 | 0.0 | 0.0 | |
| 2123 | 0.000 | 0.0 | 0.0 | |
| 2124 | 0.000 | 0.0 | 0.0 | |

| | | | |
|------|---------------------------------|--------------------------------------|-----|
| 2125 | 0.000 | 0.0 | 0.0 |
| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
| 0 | 73.0 | | 0.5 |
| 1 | 17.0 | | 2.1 |
| 2 | 16.0 | | 2.1 |
| 3 | 16.0 | | 2.4 |
| 4 | 16.0 | | 2.4 |
| ... | ... | | ... |
| 2121 | 79.0 | | 0.2 |
| 2122 | 78.0 | | 0.4 |
| 2123 | 79.0 | | 0.4 |
| 2124 | 78.0 | | 0.4 |
| 2125 | 74.0 | | 0.4 |

| | | | |
|------|--|------|-----|
| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
| 0 | | 43.0 | ... |
| 1 | | 0.0 | ... |
| 2 | | 0.0 | ... |
| 3 | | 0.0 | ... |
| 4 | | 0.0 | ... |
| ... | | ... | ... |
| 2121 | | 25.0 | ... |
| 2122 | | 22.0 | ... |
| 2123 | | 20.0 | ... |
| 2124 | | 27.0 | ... |
| 2125 | | 36.0 | ... |

| | | | | |
|------|---------------|---------------|---------------------------|---|
| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
| 0 | 62.0 | 126.0 | 2.0 | |
| 1 | 68.0 | 198.0 | 6.0 | |
| 2 | 68.0 | 198.0 | 5.0 | |
| 3 | 53.0 | 170.0 | 11.0 | |
| 4 | 53.0 | 170.0 | 9.0 | |
| ... | ... | ... | ... | |
| 2121 | 137.0 | 177.0 | 4.0 | |
| 2122 | 103.0 | 169.0 | 6.0 | |
| 2123 | 103.0 | 170.0 | 5.0 | |
| 2124 | 103.0 | 169.0 | 6.0 | |
| 2125 | 117.0 | 159.0 | 2.0 | |

| | | | | |
|---|----------------------------|----------------|----------------|-------|
| | histogram_number_of_zeroes | histogram_mode | histogram_mean | \ |
| 0 | | 0.0 | 120.0 | 137.0 |
| 1 | | 1.0 | 141.0 | 136.0 |
| 2 | | 1.0 | 141.0 | 135.0 |
| 3 | | 0.0 | 137.0 | 134.0 |
| 4 | | 0.0 | 137.0 | 136.0 |

| | | | |
|------|-----|-------|-------|
| ... | ... | ... | ... |
| 2121 | 0.0 | 153.0 | 150.0 |
| 2122 | 0.0 | 152.0 | 148.0 |
| 2123 | 0.0 | 153.0 | 148.0 |
| 2124 | 0.0 | 152.0 | 147.0 |
| 2125 | 1.0 | 145.0 | 143.0 |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 0 | 121.0 | 73.0 | 1.0 | 2.0 |
| 1 | 140.0 | 12.0 | 0.0 | 1.0 |
| 2 | 138.0 | 13.0 | 0.0 | 1.0 |
| 3 | 137.0 | 13.0 | 1.0 | 1.0 |
| 4 | 138.0 | 11.0 | 1.0 | 1.0 |
| ... | ... | ... | ... | ... |
| 2121 | 152.0 | 2.0 | 0.0 | 2.0 |
| 2122 | 151.0 | 3.0 | 1.0 | 2.0 |
| 2123 | 152.0 | 4.0 | 1.0 | 2.0 |
| 2124 | 151.0 | 4.0 | 1.0 | 2.0 |
| 2125 | 145.0 | 1.0 | 0.0 | 1.0 |

[2126 rows x 22 columns]

Converting 3 class problem into binary classification problem Originally: 1 = Normal ; 2 = Suspect; 3 = Pathological

Modification: 0 = non-pathological (previously 1 and 2); 1 = pathological

```
[3]: data['fetal_health'] = data['fetal_health'].replace(1.0,0)
data['fetal_health'] = data['fetal_health'].replace(2.0,0)
data['fetal_health'] = data['fetal_health'].replace(3.0,1)
data
```

```
[3]:
```

| | baseline value | accelerations | fetal_movement | uterine_contractions | \ |
|------|----------------|---------------|----------------|----------------------|---|
| 0 | 120.0 | 0.000 | 0.000 | 0.000 | |
| 1 | 132.0 | 0.006 | 0.000 | 0.006 | |
| 2 | 133.0 | 0.003 | 0.000 | 0.008 | |
| 3 | 134.0 | 0.003 | 0.000 | 0.008 | |
| 4 | 132.0 | 0.007 | 0.000 | 0.008 | |
| ... | ... | ... | ... | ... | |
| 2121 | 140.0 | 0.000 | 0.000 | 0.007 | |
| 2122 | 140.0 | 0.001 | 0.000 | 0.007 | |
| 2123 | 140.0 | 0.001 | 0.000 | 0.007 | |
| 2124 | 140.0 | 0.001 | 0.000 | 0.006 | |
| 2125 | 142.0 | 0.002 | 0.002 | 0.008 | |

| | light_decelerations | severe_decelerations | prolonged_decelerations | \ |
|---|---------------------|----------------------|-------------------------|---|
| 0 | 0.000 | 0.0 | 0.0 | |

| | | | |
|------|-------|-----|-----|
| 1 | 0.003 | 0.0 | 0.0 |
| 2 | 0.003 | 0.0 | 0.0 |
| 3 | 0.003 | 0.0 | 0.0 |
| 4 | 0.000 | 0.0 | 0.0 |
| ... | ... | ... | ... |
| 2121 | 0.000 | 0.0 | 0.0 |
| 2122 | 0.000 | 0.0 | 0.0 |
| 2123 | 0.000 | 0.0 | 0.0 |
| 2124 | 0.000 | 0.0 | 0.0 |
| 2125 | 0.000 | 0.0 | 0.0 |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 0 | 73.0 | 0.5 | |
| 1 | 17.0 | 2.1 | |
| 2 | 16.0 | 2.1 | |
| 3 | 16.0 | 2.4 | |
| 4 | 16.0 | 2.4 | |
| ... | ... | ... | |
| 2121 | 79.0 | 0.2 | |
| 2122 | 78.0 | 0.4 | |
| 2123 | 79.0 | 0.4 | |
| 2124 | 78.0 | 0.4 | |
| 2125 | 74.0 | 0.4 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|------|--|-----|---|
| 0 | 43.0 | ... | |
| 1 | 0.0 | ... | |
| 2 | 0.0 | ... | |
| 3 | 0.0 | ... | |
| 4 | 0.0 | ... | |
| ... | ... | ... | |
| 2121 | 25.0 | ... | |
| 2122 | 22.0 | ... | |
| 2123 | 20.0 | ... | |
| 2124 | 27.0 | ... | |
| 2125 | 36.0 | ... | |

| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
|------|---------------|---------------|---------------------------|---|
| 0 | 62.0 | 126.0 | 2.0 | |
| 1 | 68.0 | 198.0 | 6.0 | |
| 2 | 68.0 | 198.0 | 5.0 | |
| 3 | 53.0 | 170.0 | 11.0 | |
| 4 | 53.0 | 170.0 | 9.0 | |
| ... | ... | ... | ... | |
| 2121 | 137.0 | 177.0 | 4.0 | |
| 2122 | 103.0 | 169.0 | 6.0 | |
| 2123 | 103.0 | 170.0 | 5.0 | |

| | | | |
|------|-------|-------|-----|
| 2124 | 103.0 | 169.0 | 6.0 |
| 2125 | 117.0 | 159.0 | 2.0 |

| | histogram_number_of_zeroes | histogram_mode | histogram_mean \ |
|------|----------------------------|----------------|------------------|
| 0 | 0.0 | 120.0 | 137.0 |
| 1 | 1.0 | 141.0 | 136.0 |
| 2 | 1.0 | 141.0 | 135.0 |
| 3 | 0.0 | 137.0 | 134.0 |
| 4 | 0.0 | 137.0 | 136.0 |
| ... | ... | ... | ... |
| 2121 | 0.0 | 153.0 | 150.0 |
| 2122 | 0.0 | 152.0 | 148.0 |
| 2123 | 0.0 | 153.0 | 148.0 |
| 2124 | 0.0 | 152.0 | 147.0 |
| 2125 | 1.0 | 145.0 | 143.0 |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 0 | 121.0 | 73.0 | 1.0 | 0.0 |
| 1 | 140.0 | 12.0 | 0.0 | 0.0 |
| 2 | 138.0 | 13.0 | 0.0 | 0.0 |
| 3 | 137.0 | 13.0 | 1.0 | 0.0 |
| 4 | 138.0 | 11.0 | 1.0 | 0.0 |
| ... | ... | ... | ... | ... |
| 2121 | 152.0 | 2.0 | 0.0 | 0.0 |
| 2122 | 151.0 | 3.0 | 1.0 | 0.0 |
| 2123 | 152.0 | 4.0 | 1.0 | 0.0 |
| 2124 | 151.0 | 4.0 | 1.0 | 0.0 |
| 2125 | 145.0 | 1.0 | 0.0 | 0.0 |

[2126 rows x 22 columns]

```
[4]: # to confirm that we have 0 and 1
data.iloc[0:10]
```

```
[4]: baseline value accelerations fetal_movement uterine_contractions \
0      120.0      0.000      0.0      0.000
1      132.0      0.006      0.0      0.006
2      133.0      0.003      0.0      0.008
3      134.0      0.003      0.0      0.008
4      132.0      0.007      0.0      0.008
5      134.0      0.001      0.0      0.010
6      134.0      0.001      0.0      0.013
7      122.0      0.000      0.0      0.000
8      122.0      0.000      0.0      0.002
9      122.0      0.000      0.0      0.003
```

```
light_decelerations severe_decelerations prolonged_decelerations \
```

| | | | |
|---|-------|-----|-------|
| 0 | 0.000 | 0.0 | 0.000 |
| 1 | 0.003 | 0.0 | 0.000 |
| 2 | 0.003 | 0.0 | 0.000 |
| 3 | 0.003 | 0.0 | 0.000 |
| 4 | 0.000 | 0.0 | 0.000 |
| 5 | 0.009 | 0.0 | 0.002 |
| 6 | 0.008 | 0.0 | 0.003 |
| 7 | 0.000 | 0.0 | 0.000 |
| 8 | 0.000 | 0.0 | 0.000 |
| 9 | 0.000 | 0.0 | 0.000 |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|---|---------------------------------|--------------------------------------|---|
| 0 | 73.0 | 0.5 | |
| 1 | 17.0 | 2.1 | |
| 2 | 16.0 | 2.1 | |
| 3 | 16.0 | 2.4 | |
| 4 | 16.0 | 2.4 | |
| 5 | 26.0 | 5.9 | |
| 6 | 29.0 | 6.3 | |
| 7 | 83.0 | 0.5 | |
| 8 | 84.0 | 0.5 | |
| 9 | 86.0 | 0.3 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | histogram_min | \ |
|---|--|-----|---------------|---|
| 0 | 43.0 | ... | 62.0 | |
| 1 | 0.0 | ... | 68.0 | |
| 2 | 0.0 | ... | 68.0 | |
| 3 | 0.0 | ... | 53.0 | |
| 4 | 0.0 | ... | 53.0 | |
| 5 | 0.0 | ... | 50.0 | |
| 6 | 0.0 | ... | 50.0 | |
| 7 | 6.0 | ... | 62.0 | |
| 8 | 5.0 | ... | 62.0 | |
| 9 | 6.0 | ... | 62.0 | |

| | histogram_max | histogram_number_of_peaks | histogram_number_of_zeroes | \ |
|---|---------------|---------------------------|----------------------------|---|
| 0 | 126.0 | 2.0 | 0.0 | |
| 1 | 198.0 | 6.0 | 1.0 | |
| 2 | 198.0 | 5.0 | 1.0 | |
| 3 | 170.0 | 11.0 | 0.0 | |
| 4 | 170.0 | 9.0 | 0.0 | |
| 5 | 200.0 | 5.0 | 3.0 | |
| 6 | 200.0 | 6.0 | 3.0 | |
| 7 | 130.0 | 0.0 | 0.0 | |
| 8 | 130.0 | 0.0 | 0.0 | |
| 9 | 130.0 | 1.0 | 0.0 | |

| | histogram_mode | histogram_mean | histogram_median | histogram_variance \ |
|---|----------------|----------------|------------------|----------------------|
| 0 | 120.0 | 137.0 | 121.0 | 73.0 |
| 1 | 141.0 | 136.0 | 140.0 | 12.0 |
| 2 | 141.0 | 135.0 | 138.0 | 13.0 |
| 3 | 137.0 | 134.0 | 137.0 | 13.0 |
| 4 | 137.0 | 136.0 | 138.0 | 11.0 |
| 5 | 76.0 | 107.0 | 107.0 | 170.0 |
| 6 | 71.0 | 107.0 | 106.0 | 215.0 |
| 7 | 122.0 | 122.0 | 123.0 | 3.0 |
| 8 | 122.0 | 122.0 | 123.0 | 3.0 |
| 9 | 122.0 | 122.0 | 123.0 | 1.0 |

| | histogram_tendency | fetal_health |
|---|--------------------|--------------|
| 0 | 1.0 | 0.0 |
| 1 | 0.0 | 0.0 |
| 2 | 0.0 | 0.0 |
| 3 | 1.0 | 0.0 |
| 4 | 1.0 | 0.0 |
| 5 | 0.0 | 1.0 |
| 6 | 0.0 | 1.0 |
| 7 | 1.0 | 1.0 |
| 8 | 1.0 | 1.0 |
| 9 | 1.0 | 1.0 |

[10 rows x 22 columns]

```
[5]: X = data.drop('fetal_health', axis = 1)
X
```

```
[5]:
```

| | baseline | value | accelerations | fetal_movement | uterine_contractions \ |
|------|----------|-------|---------------|----------------|------------------------|
| 0 | | 120.0 | 0.000 | 0.000 | 0.000 |
| 1 | | 132.0 | 0.006 | 0.000 | 0.006 |
| 2 | | 133.0 | 0.003 | 0.000 | 0.008 |
| 3 | | 134.0 | 0.003 | 0.000 | 0.008 |
| 4 | | 132.0 | 0.007 | 0.000 | 0.008 |
| ... | | ... | ... | ... | ... |
| 2121 | | 140.0 | 0.000 | 0.000 | 0.007 |
| 2122 | | 140.0 | 0.001 | 0.000 | 0.007 |
| 2123 | | 140.0 | 0.001 | 0.000 | 0.007 |
| 2124 | | 140.0 | 0.001 | 0.000 | 0.006 |
| 2125 | | 142.0 | 0.002 | 0.002 | 0.008 |

| | light_decelerations | severe_decelerations | prolongued_decelerations \ |
|---|---------------------|----------------------|----------------------------|
| 0 | 0.000 | 0.0 | 0.0 |
| 1 | 0.003 | 0.0 | 0.0 |
| 2 | 0.003 | 0.0 | 0.0 |
| 3 | 0.003 | 0.0 | 0.0 |

| | | | |
|------|-------|-----|-----|
| 4 | 0.000 | 0.0 | 0.0 |
| ... | ... | ... | ... |
| 2121 | 0.000 | 0.0 | 0.0 |
| 2122 | 0.000 | 0.0 | 0.0 |
| 2123 | 0.000 | 0.0 | 0.0 |
| 2124 | 0.000 | 0.0 | 0.0 |
| 2125 | 0.000 | 0.0 | 0.0 |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 0 | 73.0 | 0.5 | |
| 1 | 17.0 | 2.1 | |
| 2 | 16.0 | 2.1 | |
| 3 | 16.0 | 2.4 | |
| 4 | 16.0 | 2.4 | |
| ... | ... | ... | |
| 2121 | 79.0 | 0.2 | |
| 2122 | 78.0 | 0.4 | |
| 2123 | 79.0 | 0.4 | |
| 2124 | 78.0 | 0.4 | |
| 2125 | 74.0 | 0.4 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|------|--|-----|---|
| 0 | 43.0 | ... | |
| 1 | 0.0 | ... | |
| 2 | 0.0 | ... | |
| 3 | 0.0 | ... | |
| 4 | 0.0 | ... | |
| ... | ... | ... | |
| 2121 | 25.0 | ... | |
| 2122 | 22.0 | ... | |
| 2123 | 20.0 | ... | |
| 2124 | 27.0 | ... | |
| 2125 | 36.0 | ... | |

| | histogram_width | histogram_min | histogram_max | \ |
|------|-----------------|---------------|---------------|---|
| 0 | 64.0 | 62.0 | 126.0 | |
| 1 | 130.0 | 68.0 | 198.0 | |
| 2 | 130.0 | 68.0 | 198.0 | |
| 3 | 117.0 | 53.0 | 170.0 | |
| 4 | 117.0 | 53.0 | 170.0 | |
| ... | ... | ... | ... | |
| 2121 | 40.0 | 137.0 | 177.0 | |
| 2122 | 66.0 | 103.0 | 169.0 | |
| 2123 | 67.0 | 103.0 | 170.0 | |
| 2124 | 66.0 | 103.0 | 169.0 | |
| 2125 | 42.0 | 117.0 | 159.0 | |

| | histogram_number_of_peaks | histogram_number_of_zeroes | histogram_mode | \ |
|------|---------------------------|----------------------------|----------------|---|
| 0 | 2.0 | 0.0 | 120.0 | |
| 1 | 6.0 | 1.0 | 141.0 | |
| 2 | 5.0 | 1.0 | 141.0 | |
| 3 | 11.0 | 0.0 | 137.0 | |
| 4 | 9.0 | 0.0 | 137.0 | |
| ... | ... | ... | ... | |
| 2121 | 4.0 | 0.0 | 153.0 | |
| 2122 | 6.0 | 0.0 | 152.0 | |
| 2123 | 5.0 | 0.0 | 153.0 | |
| 2124 | 6.0 | 0.0 | 152.0 | |
| 2125 | 2.0 | 1.0 | 145.0 | |

| | histogram_mean | histogram_median | histogram_variance | histogram_tendency |
|------|----------------|------------------|--------------------|--------------------|
| 0 | 137.0 | 121.0 | 73.0 | 1.0 |
| 1 | 136.0 | 140.0 | 12.0 | 0.0 |
| 2 | 135.0 | 138.0 | 13.0 | 0.0 |
| 3 | 134.0 | 137.0 | 13.0 | 1.0 |
| 4 | 136.0 | 138.0 | 11.0 | 1.0 |
| ... | ... | ... | ... | ... |
| 2121 | 150.0 | 152.0 | 2.0 | 0.0 |
| 2122 | 148.0 | 151.0 | 3.0 | 1.0 |
| 2123 | 148.0 | 152.0 | 4.0 | 1.0 |
| 2124 | 147.0 | 151.0 | 4.0 | 1.0 |
| 2125 | 143.0 | 145.0 | 1.0 | 0.0 |

[2126 rows x 21 columns]

```
[6]: # Known output Y = classification result
y = data['fetal_health']
y
```

```
[6]: 0      0.0
      1      0.0
      2      0.0
      3      0.0
      4      0.0
      ...
      2121    0.0
      2122    0.0
      2123    0.0
      2124    0.0
      2125    0.0
      Name: fetal_health, Length: 2126, dtype: float64
```

```
[7]: from sklearn.model_selection import train_test_split
```

```
x_train,x_test,y_train,y_test = train_test_split(X,y,test_size=0.2,
↳random_state=142)
x_train
```

```
[7]:
```

| | baseline value | accelerations | fetal_movement | uterine_contractions | \ |
|------|----------------|---------------|----------------|----------------------|---|
| 199 | 120.0 | 0.000 | 0.013 | 0.000 | |
| 340 | 133.0 | 0.000 | 0.000 | 0.000 | |
| 440 | 142.0 | 0.001 | 0.003 | 0.001 | |
| 1427 | 144.0 | 0.006 | 0.000 | 0.004 | |
| 1015 | 139.0 | 0.007 | 0.000 | 0.005 | |
| ... | ... | ... | ... | ... | |
| 1420 | 142.0 | 0.006 | 0.000 | 0.007 | |
| 1616 | 144.0 | 0.003 | 0.049 | 0.002 | |
| 1050 | 125.0 | 0.007 | 0.000 | 0.005 | |
| 511 | 154.0 | 0.007 | 0.001 | 0.002 | |
| 277 | 123.0 | 0.001 | 0.000 | 0.004 | |

| | light_decelerations | severe_decelerations | prolongued_decelerations | \ |
|------|---------------------|----------------------|--------------------------|---|
| 199 | 0.001 | 0.0 | 0.000 | |
| 340 | 0.000 | 0.0 | 0.000 | |
| 440 | 0.002 | 0.0 | 0.000 | |
| 1427 | 0.000 | 0.0 | 0.000 | |
| 1015 | 0.000 | 0.0 | 0.000 | |
| ... | ... | ... | ... | |
| 1420 | 0.000 | 0.0 | 0.000 | |
| 1616 | 0.006 | 0.0 | 0.001 | |
| 1050 | 0.000 | 0.0 | 0.000 | |
| 511 | 0.000 | 0.0 | 0.000 | |
| 277 | 0.000 | 0.0 | 0.000 | |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 199 | 53.0 | 0.7 | |
| 340 | 75.0 | 0.2 | |
| 440 | 55.0 | 1.3 | |
| 1427 | 39.0 | 1.0 | |
| 1015 | 38.0 | 0.9 | |
| ... | ... | ... | |
| 1420 | 42.0 | 1.0 | |
| 1616 | 66.0 | 3.4 | |
| 1050 | 26.0 | 1.3 | |
| 511 | 45.0 | 0.8 | |
| 277 | 54.0 | 0.5 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|-----|--|-----|---|
| 199 | 7.0 | ... | |
| 340 | 91.0 | ... | |
| 440 | 10.0 | ... | |

| | | |
|------|-----|-----|
| 1427 | 5.0 | ... |
| 1015 | 0.0 | ... |
| ... | ... | ... |
| 1420 | 0.0 | ... |
| 1616 | 0.0 | ... |
| 1050 | 0.0 | ... |
| 511 | 0.0 | ... |
| 277 | 9.0 | ... |

| | histogram_width | histogram_min | histogram_max | \ |
|------|-----------------|---------------|---------------|---|
| 199 | 77.0 | 56.0 | 133.0 | |
| 340 | 7.0 | 131.0 | 138.0 | |
| 440 | 115.0 | 52.0 | 167.0 | |
| 1427 | 43.0 | 136.0 | 179.0 | |
| 1015 | 34.0 | 136.0 | 170.0 | |
| ... | ... | ... | ... | |
| 1420 | 97.0 | 74.0 | 171.0 | |
| 1616 | 113.0 | 67.0 | 180.0 | |
| 1050 | 87.0 | 79.0 | 166.0 | |
| 511 | 47.0 | 142.0 | 189.0 | |
| 277 | 29.0 | 116.0 | 145.0 | |

| | histogram_number_of_peaks | histogram_number_of_zeroes | histogram_mode | \ |
|------|---------------------------|----------------------------|----------------|---|
| 199 | 6.0 | 0.0 | 123.0 | |
| 340 | 1.0 | 0.0 | 133.0 | |
| 440 | 12.0 | 3.0 | 148.0 | |
| 1427 | 1.0 | 0.0 | 157.0 | |
| 1015 | 1.0 | 0.0 | 144.0 | |
| ... | ... | ... | ... | |
| 1420 | 6.0 | 0.0 | 148.0 | |
| 1616 | 7.0 | 0.0 | 141.0 | |
| 1050 | 4.0 | 0.0 | 131.0 | |
| 511 | 2.0 | 1.0 | 161.0 | |
| 277 | 4.0 | 0.0 | 126.0 | |

| | histogram_mean | histogram_median | histogram_variance | histogram_tendency |
|------|----------------|------------------|--------------------|--------------------|
| 199 | 121.0 | 123.0 | 3.0 | 1.0 |
| 340 | 134.0 | 135.0 | 0.0 | 0.0 |
| 440 | 143.0 | 147.0 | 17.0 | 1.0 |
| 1427 | 157.0 | 157.0 | 4.0 | 0.0 |
| 1015 | 146.0 | 146.0 | 4.0 | -1.0 |
| ... | ... | ... | ... | ... |
| 1420 | 148.0 | 149.0 | 5.0 | 1.0 |
| 1616 | 110.0 | 147.0 | 89.0 | 1.0 |
| 1050 | 133.0 | 132.0 | 11.0 | 0.0 |
| 511 | 166.0 | 165.0 | 10.0 | 0.0 |
| 277 | 128.0 | 129.0 | 3.0 | 0.0 |

[1700 rows x 21 columns]

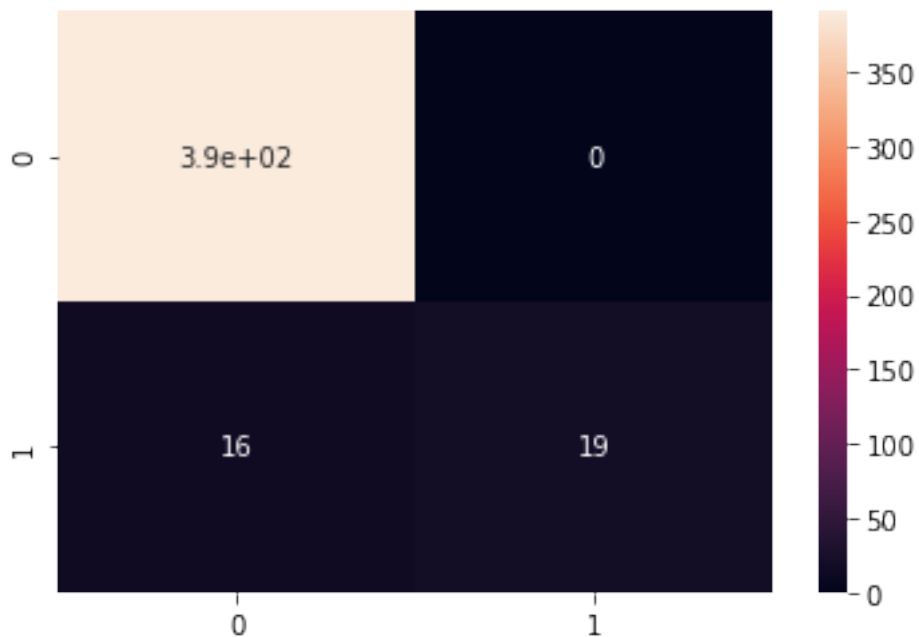
```
[8]: from sklearn import svm  
svm = svm.SVC()
```

```
[9]: svm.fit(x_train,y_train)  
y_pred_train = svm.predict(x_train)  
y_pred_test = svm.predict(x_test)
```

```
[10]: # true_val = y_train  
# pred_val = y_pred_train
```

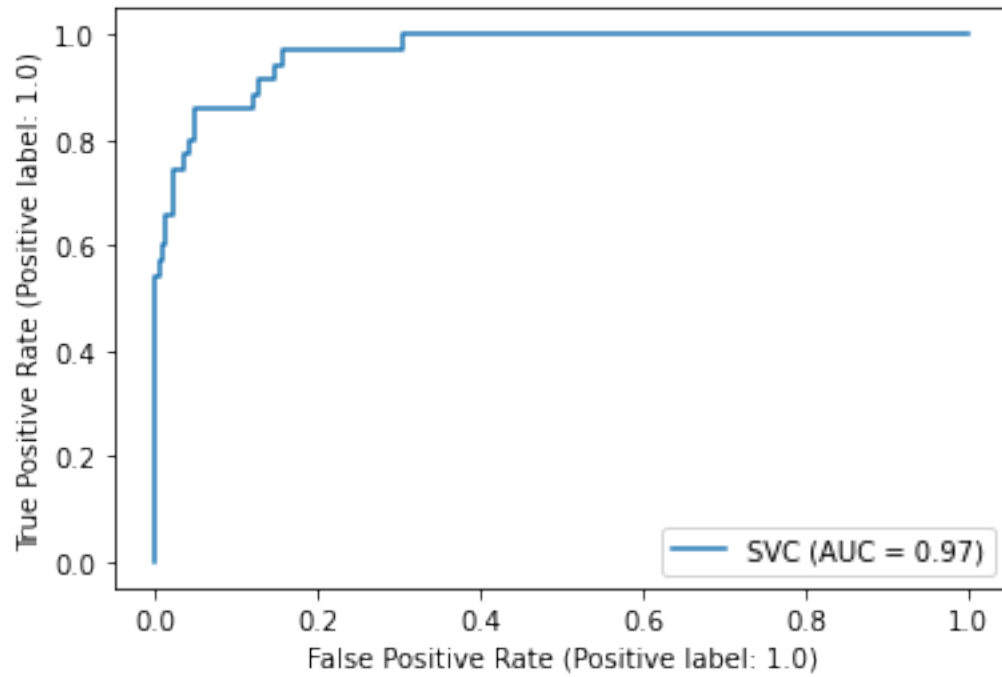
```
[11]: #confusion matrix  
from sklearn.metrics import confusion_matrix  
from sklearn.metrics import accuracy_score  
  
cm = confusion_matrix(y_test, y_pred_test)  
  
sns.heatmap(cm, annot=True)  
print(accuracy_score(y_test, y_pred_test))
```

0.9624413145539906



```
[12]: from sklearn import metrics  
metrics.plot_roc_curve(svm, x_test, y_test)
```

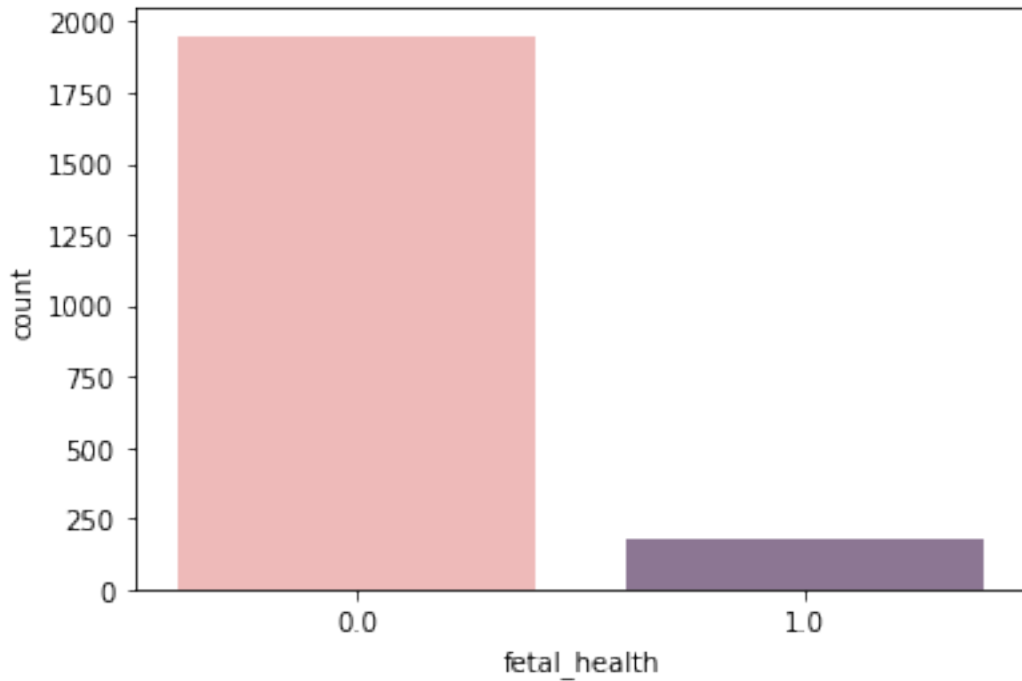
```
plt.show()
```



```
[13]: #Because the merging of two groups, our data is not balanced as a result. We
      → have way more class 0 than class 1
```

```
colours=["#f7b2b0", "#8f7198", "#003f5c"]
sns.countplot(data= data, x="fetal_health", palette=colours)
```

```
[13]: <AxesSubplot:xlabel='fetal_health', ylabel='count'>
```



```
[14]: from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred_test))
```

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0.0 | 0.96 | 1.00 | 0.98 | 391 |
| 1.0 | 1.00 | 0.54 | 0.70 | 35 |
| accuracy | | | 0.96 | 426 |
| macro avg | 0.98 | 0.77 | 0.84 | 426 |
| weighted avg | 0.96 | 0.96 | 0.96 | 426 |

0.0.3 Class Imbalance

We have issue with class imbalance which can hinder our accuracy of model. Here are some methods that we can try - Undersampling of class 0 - Upsampling of class 1

Undersampling

```
[15]: #undersampling
class_0 = data[data['fetal_health'] == 0]
class_0
```

```
[15]: baseline value  accelerations  fetal_movement  uterine_contractions  \
0          120.0          0.000          0.000          0.000
```

| | | | | |
|------|-------|-------|-------|-------|
| 1 | 132.0 | 0.006 | 0.000 | 0.006 |
| 2 | 133.0 | 0.003 | 0.000 | 0.008 |
| 3 | 134.0 | 0.003 | 0.000 | 0.008 |
| 4 | 132.0 | 0.007 | 0.000 | 0.008 |
| ... | ... | ... | ... | ... |
| 2121 | 140.0 | 0.000 | 0.000 | 0.007 |
| 2122 | 140.0 | 0.001 | 0.000 | 0.007 |
| 2123 | 140.0 | 0.001 | 0.000 | 0.007 |
| 2124 | 140.0 | 0.001 | 0.000 | 0.006 |
| 2125 | 142.0 | 0.002 | 0.002 | 0.008 |

| | light_decelerations | severe_decelerations | prolongued_decelerations | \ |
|------|---------------------|----------------------|--------------------------|---|
| 0 | 0.000 | 0.0 | 0.0 | |
| 1 | 0.003 | 0.0 | 0.0 | |
| 2 | 0.003 | 0.0 | 0.0 | |
| 3 | 0.003 | 0.0 | 0.0 | |
| 4 | 0.000 | 0.0 | 0.0 | |
| ... | ... | ... | ... | |
| 2121 | 0.000 | 0.0 | 0.0 | |
| 2122 | 0.000 | 0.0 | 0.0 | |
| 2123 | 0.000 | 0.0 | 0.0 | |
| 2124 | 0.000 | 0.0 | 0.0 | |
| 2125 | 0.000 | 0.0 | 0.0 | |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 0 | 73.0 | 0.5 | |
| 1 | 17.0 | 2.1 | |
| 2 | 16.0 | 2.1 | |
| 3 | 16.0 | 2.4 | |
| 4 | 16.0 | 2.4 | |
| ... | ... | ... | |
| 2121 | 79.0 | 0.2 | |
| 2122 | 78.0 | 0.4 | |
| 2123 | 79.0 | 0.4 | |
| 2124 | 78.0 | 0.4 | |
| 2125 | 74.0 | 0.4 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|------|--|-----|---|
| 0 | 43.0 | ... | |
| 1 | 0.0 | ... | |
| 2 | 0.0 | ... | |
| 3 | 0.0 | ... | |
| 4 | 0.0 | ... | |
| ... | ... | ... | |
| 2121 | 25.0 | ... | |
| 2122 | 22.0 | ... | |
| 2123 | 20.0 | ... | |

| | | |
|------|------|-----|
| 2124 | 27.0 | ... |
| 2125 | 36.0 | ... |

| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
|------|---------------|---------------|---------------------------|---|
| 0 | 62.0 | 126.0 | 2.0 | |
| 1 | 68.0 | 198.0 | 6.0 | |
| 2 | 68.0 | 198.0 | 5.0 | |
| 3 | 53.0 | 170.0 | 11.0 | |
| 4 | 53.0 | 170.0 | 9.0 | |
| ... | ... | ... | ... | |
| 2121 | 137.0 | 177.0 | 4.0 | |
| 2122 | 103.0 | 169.0 | 6.0 | |
| 2123 | 103.0 | 170.0 | 5.0 | |
| 2124 | 103.0 | 169.0 | 6.0 | |
| 2125 | 117.0 | 159.0 | 2.0 | |

| | histogram_number_of_zeroes | histogram_mode | histogram_mean | \ |
|------|----------------------------|----------------|----------------|---|
| 0 | 0.0 | 120.0 | 137.0 | |
| 1 | 1.0 | 141.0 | 136.0 | |
| 2 | 1.0 | 141.0 | 135.0 | |
| 3 | 0.0 | 137.0 | 134.0 | |
| 4 | 0.0 | 137.0 | 136.0 | |
| ... | ... | ... | ... | |
| 2121 | 0.0 | 153.0 | 150.0 | |
| 2122 | 0.0 | 152.0 | 148.0 | |
| 2123 | 0.0 | 153.0 | 148.0 | |
| 2124 | 0.0 | 152.0 | 147.0 | |
| 2125 | 1.0 | 145.0 | 143.0 | |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 0 | 121.0 | 73.0 | 1.0 | 0.0 |
| 1 | 140.0 | 12.0 | 0.0 | 0.0 |
| 2 | 138.0 | 13.0 | 0.0 | 0.0 |
| 3 | 137.0 | 13.0 | 1.0 | 0.0 |
| 4 | 138.0 | 11.0 | 1.0 | 0.0 |
| ... | ... | ... | ... | ... |
| 2121 | 152.0 | 2.0 | 0.0 | 0.0 |
| 2122 | 151.0 | 3.0 | 1.0 | 0.0 |
| 2123 | 152.0 | 4.0 | 1.0 | 0.0 |
| 2124 | 151.0 | 4.0 | 1.0 | 0.0 |
| 2125 | 145.0 | 1.0 | 0.0 | 0.0 |

[1950 rows x 22 columns]

```
[16]: class_1 = data[data['fetal_health'] == 1]
      class_1
```



```

[16]: baseline value accelerations fetal_movement uterine_contractions \
5      134.0      0.001      0.000      0.010
6      134.0      0.001      0.000      0.013
7      122.0      0.000      0.000      0.000
8      122.0      0.000      0.000      0.002
9      122.0      0.000      0.000      0.003
...      ...      ...      ...      ...
2036    128.0      0.000      0.002      0.007
2037    128.0      0.000      0.002      0.007
2038    128.0      0.000      0.002      0.007
2039    128.0      0.000      0.003      0.007
2048    128.0      0.000      0.000      0.008

light_decelerations severe_decelerations prolonged_decelerations \
5      0.009      0.0      0.002
6      0.008      0.0      0.003
7      0.000      0.0      0.000
8      0.000      0.0      0.000
9      0.000      0.0      0.000
...      ...      ...      ...
2036    0.005      0.0      0.002
2037    0.005      0.0      0.002
2038    0.006      0.0      0.002
2039    0.006      0.0      0.002
2048    0.010      0.0      0.000

abnormal_short_term_variability mean_value_of_short_term_variability \
5      26.0      5.9
6      29.0      6.3
7      83.0      0.5
8      84.0      0.5
9      86.0      0.3
...      ...      ...
2036    65.0      2.6
2037    66.0      2.7
2038    65.0      2.5
2039    65.0      2.5
2048    63.0      4.2

percentage_of_time_with_abnormal_long_term_variability ... \
5      0.0      ...
6      0.0      ...
7      6.0      ...
8      5.0      ...
9      6.0      ...
...      ...      ...
2036    0.0      ...

```

| | | |
|------|-----|-----|
| 2037 | 0.0 | ... |
| 2038 | 0.0 | ... |
| 2039 | 0.0 | ... |
| 2048 | 0.0 | ... |

| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
|------|---------------|---------------|---------------------------|---|
| 5 | 50.0 | 200.0 | 5.0 | |
| 6 | 50.0 | 200.0 | 6.0 | |
| 7 | 62.0 | 130.0 | 0.0 | |
| 8 | 62.0 | 130.0 | 0.0 | |
| 9 | 62.0 | 130.0 | 1.0 | |
| ... | ... | ... | ... | |
| 2036 | 63.0 | 192.0 | 6.0 | |
| 2037 | 63.0 | 192.0 | 4.0 | |
| 2038 | 63.0 | 192.0 | 6.0 | |
| 2039 | 63.0 | 192.0 | 5.0 | |
| 2048 | 66.0 | 156.0 | 5.0 | |

| | histogram_number_of_zeroes | histogram_mode | histogram_mean | \ |
|------|----------------------------|----------------|----------------|---|
| 5 | 3.0 | 76.0 | 107.0 | |
| 6 | 3.0 | 71.0 | 107.0 | |
| 7 | 0.0 | 122.0 | 122.0 | |
| 8 | 0.0 | 122.0 | 122.0 | |
| 9 | 0.0 | 122.0 | 122.0 | |
| ... | ... | ... | ... | |
| 2036 | 0.0 | 103.0 | 92.0 | |
| 2037 | 0.0 | 105.0 | 89.0 | |
| 2038 | 0.0 | 104.0 | 94.0 | |
| 2039 | 0.0 | 114.0 | 97.0 | |
| 2048 | 0.0 | 69.0 | 73.0 | |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 5 | 107.0 | 170.0 | 0.0 | 1.0 |
| 6 | 106.0 | 215.0 | 0.0 | 1.0 |
| 7 | 123.0 | 3.0 | 1.0 | 1.0 |
| 8 | 123.0 | 3.0 | 1.0 | 1.0 |
| 9 | 123.0 | 1.0 | 1.0 | 1.0 |
| ... | ... | ... | ... | ... |
| 2036 | 114.0 | 28.0 | 0.0 | 1.0 |
| 2037 | 112.0 | 21.0 | 0.0 | 1.0 |
| 2038 | 115.0 | 28.0 | 0.0 | 1.0 |
| 2039 | 116.0 | 25.0 | 0.0 | 1.0 |
| 2048 | 118.0 | 128.0 | 0.0 | 1.0 |

[176 rows x 22 columns]

```
[17]: class_count_0, class_count_1 = data['fetal_health'].value_counts()
      class_0_under = class_0.sample(class_count_1)
      class_0_under #now class 0 and 1 have the same size
```

```
[17]:
```

| | baseline | value | accelerations | fetal_movement | uterine_contractions | \ |
|------|----------|-------|---------------|----------------|----------------------|---|
| 836 | | 128.0 | 0.004 | 0.000 | 0.002 | |
| 1185 | | 120.0 | 0.000 | 0.000 | 0.012 | |
| 864 | | 142.0 | 0.002 | 0.000 | 0.003 | |
| 1614 | | 133.0 | 0.006 | 0.035 | 0.003 | |
| 1195 | | 138.0 | 0.012 | 0.000 | 0.005 | |
| ... | | ... | ... | ... | ... | |
| 1374 | | 121.0 | 0.004 | 0.000 | 0.009 | |
| 2011 | | 125.0 | 0.000 | 0.001 | 0.009 | |
| 1663 | | 106.0 | 0.000 | 0.000 | 0.010 | |
| 530 | | 142.0 | 0.016 | 0.071 | 0.000 | |
| 218 | | 127.0 | 0.005 | 0.008 | 0.001 | |

| | light_decelerations | severe_decelerations | prolongued_decelerations | \ |
|------|---------------------|----------------------|--------------------------|---|
| 836 | 0.007 | 0.0 | 0.000 | |
| 1185 | 0.000 | 0.0 | 0.000 | |
| 864 | 0.000 | 0.0 | 0.000 | |
| 1614 | 0.001 | 0.0 | 0.000 | |
| 1195 | 0.000 | 0.0 | 0.000 | |
| ... | ... | ... | ... | |
| 1374 | 0.002 | 0.0 | 0.000 | |
| 2011 | 0.008 | 0.0 | 0.001 | |
| 1663 | 0.000 | 0.0 | 0.000 | |
| 530 | 0.002 | 0.0 | 0.000 | |
| 218 | 0.001 | 0.0 | 0.000 | |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 836 | 20.0 | 2.6 | |
| 1185 | 53.0 | 0.7 | |
| 864 | 45.0 | 0.7 | |
| 1614 | 65.0 | 2.7 | |
| 1195 | 37.0 | 1.2 | |
| ... | ... | ... | |
| 1374 | 24.0 | 1.5 | |
| 2011 | 64.0 | 1.1 | |
| 1663 | 63.0 | 0.6 | |
| 530 | 32.0 | 3.1 | |
| 218 | 54.0 | 1.0 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|------|--|-----|---|
| 836 | 0.0 | ... | |
| 1185 | 2.0 | ... | |
| 864 | 16.0 | ... | |

| | | |
|------|-----|-----|
| 1614 | 0.0 | ... |
| 1195 | 0.0 | ... |
| ... | ... | ... |
| 1374 | 0.0 | ... |
| 2011 | 0.0 | ... |
| 1663 | 0.0 | ... |
| 530 | 0.0 | ... |
| 218 | 5.0 | ... |

| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
|------|---------------|---------------|---------------------------|---|
| 836 | 79.0 | 171.0 | 2.0 | |
| 1185 | 107.0 | 127.0 | 2.0 | |
| 864 | 135.0 | 159.0 | 0.0 | |
| 1614 | 126.0 | 161.0 | 2.0 | |
| 1195 | 120.0 | 168.0 | 0.0 | |
| ... | ... | ... | ... | |
| 1374 | 90.0 | 157.0 | 2.0 | |
| 2011 | 64.0 | 146.0 | 5.0 | |
| 1663 | 95.0 | 122.0 | 1.0 | |
| 530 | 51.0 | 200.0 | 10.0 | |
| 218 | 63.0 | 146.0 | 9.0 | |

| | histogram_number_of_zeroes | histogram_mode | histogram_mean | \ |
|------|----------------------------|----------------|----------------|---|
| 836 | 1.0 | 151.0 | 137.0 | |
| 1185 | 0.0 | 122.0 | 119.0 | |
| 864 | 0.0 | 150.0 | 149.0 | |
| 1614 | 0.0 | 138.0 | 126.0 | |
| 1195 | 1.0 | 146.0 | 147.0 | |
| ... | ... | ... | ... | |
| 1374 | 0.0 | 120.0 | 125.0 | |
| 2011 | 2.0 | 116.0 | 110.0 | |
| 1663 | 0.0 | 112.0 | 110.0 | |
| 530 | 0.0 | 167.0 | 154.0 | |
| 218 | 1.0 | 140.0 | 135.0 | |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 836 | 142.0 | 39.0 | 1.0 | 0.0 |
| 1185 | 121.0 | 2.0 | 1.0 | 0.0 |
| 864 | 150.0 | 1.0 | 0.0 | 0.0 |
| 1614 | 140.0 | 3.0 | 0.0 | 0.0 |
| 1195 | 148.0 | 4.0 | 0.0 | 0.0 |
| ... | ... | ... | ... | ... |
| 1374 | 123.0 | 16.0 | 0.0 | 0.0 |
| 2011 | 115.0 | 10.0 | 0.0 | 0.0 |
| 1663 | 112.0 | 1.0 | 0.0 | 0.0 |
| 530 | 160.0 | 55.0 | 1.0 | 0.0 |
| 218 | 138.0 | 14.0 | 1.0 | 0.0 |

[176 rows x 22 columns]

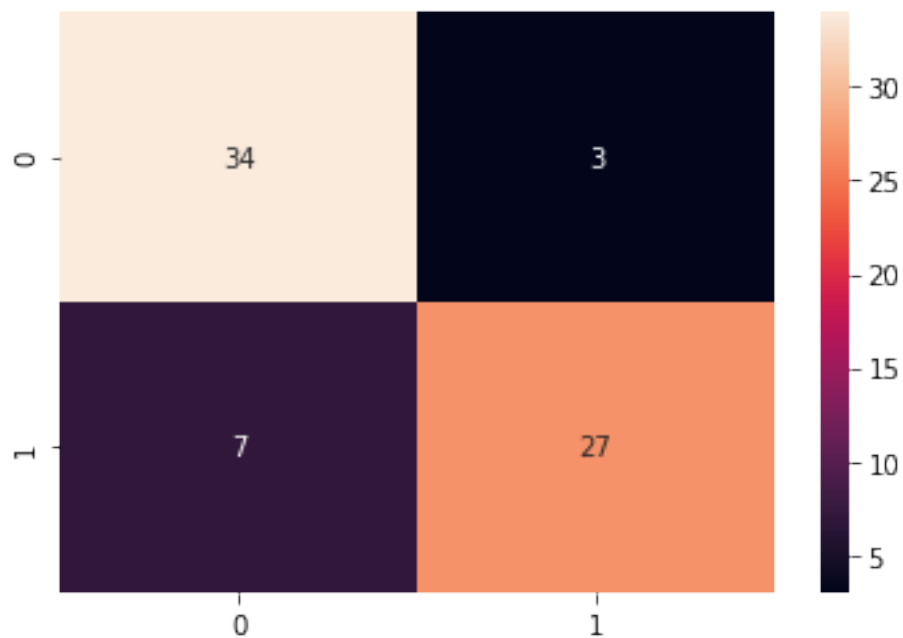
```
[18]: from sklearn import svm
data_under = pd.concat([class_0_under, class_1], axis=0)
X_under = data_under.drop('fetal_health', axis = 1)
y_under = data_under['fetal_health']
x_train_u, x_test_u, y_train_u, y_test_u = \
    train_test_split(X_under, y_under, test_size=0.2, random_state=142)
svm_under = svm.SVC()
svm_under.fit(x_train_u, y_train_u)
y_pred_train_u = svm_under.predict(x_train_u)
y_pred_test_u = svm_under.predict(x_test_u)
```

```
[19]: #confusion matrix

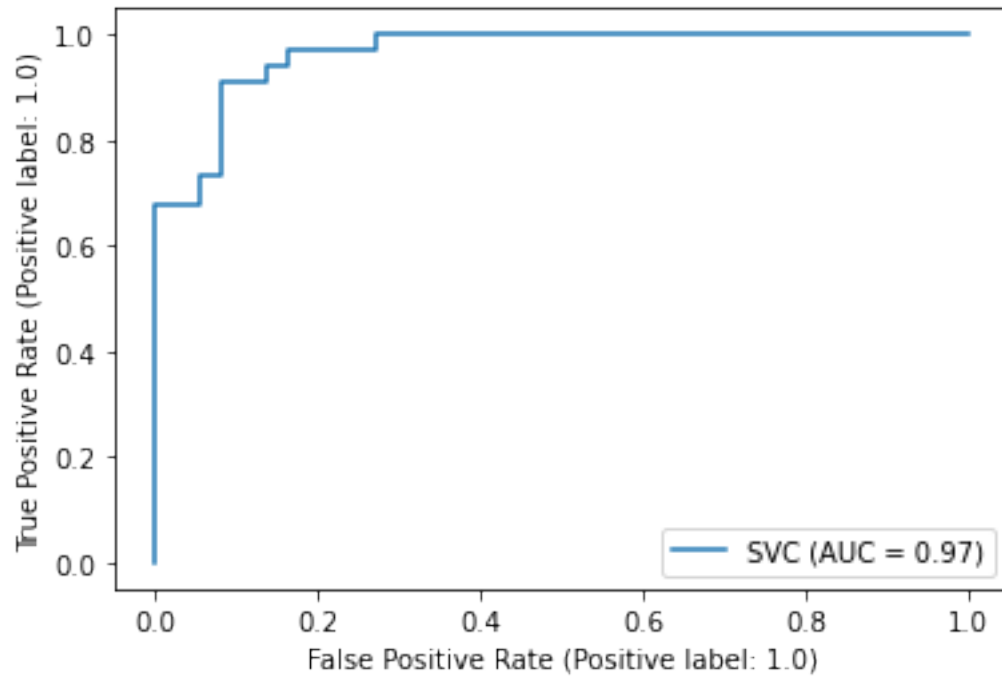
cm = confusion_matrix(y_test_u, y_pred_test_u)

sns.heatmap(cm, annot=True)
print(accuracy_score(y_test_u, y_pred_test_u))
```

0.8591549295774648



```
[20]: metrics.plot_roc_curve(svm_under, x_test_u, y_test_u)
plt.show()
```



```
[21]: from sklearn.metrics import classification_report
print(classification_report(y_test_u, y_pred_test_u))
```

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0.0 | 0.83 | 0.92 | 0.87 | 37 |
| 1.0 | 0.90 | 0.79 | 0.84 | 34 |
| accuracy | | | 0.86 | 71 |
| macro avg | 0.86 | 0.86 | 0.86 | 71 |
| weighted avg | 0.86 | 0.86 | 0.86 | 71 |

Oversampling SMOTE

```
[22]: #oversampling
class_1_over = class_1.sample(class_count_0, replace=True)
class_1_over
```

```
[22]: baseline value accelerations fetal_movement uterine_contractions \
645          123.0          0.000          0.000          0.000
1766          120.0          0.000          0.000          0.005
697           131.0          0.001          0.369          0.003
2028          129.0          0.000          0.001          0.005
1960          133.0          0.001          0.000          0.009
...           ...           ...           ...           ...
```

| | | | | |
|------|-------|-------|-------|-------|
| 1768 | 120.0 | 0.000 | 0.000 | 0.005 |
| 54 | 148.0 | 0.000 | 0.003 | 0.000 |
| 473 | 150.0 | 0.000 | 0.003 | 0.000 |
| 1768 | 120.0 | 0.000 | 0.000 | 0.005 |
| 773 | 130.0 | 0.000 | 0.000 | 0.000 |

| | light_decelerations | severe_decelerations | prolongued_decelerations | \ |
|------|---------------------|----------------------|--------------------------|---|
| 645 | 0.000 | 0.0 | 0.000 | |
| 1766 | 0.006 | 0.0 | 0.001 | |
| 697 | 0.003 | 0.0 | 0.002 | |
| 2028 | 0.008 | 0.0 | 0.002 | |
| 1960 | 0.004 | 0.0 | 0.004 | |
| ... | ... | ... | ... | |
| 1768 | 0.006 | 0.0 | 0.002 | |
| 54 | 0.000 | 0.0 | 0.000 | |
| 473 | 0.000 | 0.0 | 0.000 | |
| 1768 | 0.006 | 0.0 | 0.002 | |
| 773 | 0.000 | 0.0 | 0.000 | |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 645 | 76.0 | 0.2 | |
| 1766 | 62.0 | 1.5 | |
| 697 | 29.0 | 2.0 | |
| 2028 | 65.0 | 2.9 | |
| 1960 | 63.0 | 2.1 | |
| ... | ... | ... | |
| 1768 | 64.0 | 1.1 | |
| 54 | 70.0 | 0.3 | |
| 473 | 80.0 | 0.4 | |
| 1768 | 64.0 | 1.1 | |
| 773 | 79.0 | 0.2 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|------|--|-----|---|
| 645 | 85.0 | ... | |
| 1766 | 0.0 | ... | |
| 697 | 0.0 | ... | |
| 2028 | 0.0 | ... | |
| 1960 | 0.0 | ... | |
| ... | ... | ... | |
| 1768 | 0.0 | ... | |
| 54 | 69.0 | ... | |
| 473 | 24.0 | ... | |
| 1768 | 0.0 | ... | |
| 773 | 91.0 | ... | |

| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
|-----|---------------|---------------|---------------------------|---|
| 645 | 118.0 | 129.0 | 1.0 | |

| | | | |
|------|-------|-------|------|
| 1766 | 69.0 | 154.0 | 4.0 |
| 697 | 54.0 | 165.0 | 14.0 |
| 2028 | 50.0 | 151.0 | 6.0 |
| 1960 | 52.0 | 185.0 | 4.0 |
| ... | ... | ... | ... |
| 1768 | 69.0 | 154.0 | 2.0 |
| 54 | 136.0 | 154.0 | 3.0 |
| 473 | 124.0 | 176.0 | 1.0 |
| 1768 | 69.0 | 154.0 | 2.0 |
| 773 | 127.0 | 133.0 | 1.0 |

| | histogram_number_of_zeroes | histogram_mode | histogram_mean | \ |
|------|----------------------------|----------------|----------------|---|
| 645 | 0.0 | 123.0 | 122.0 | |
| 1766 | 1.0 | 86.0 | 101.0 | |
| 697 | 0.0 | 129.0 | 112.0 | |
| 2028 | 2.0 | 105.0 | 84.0 | |
| 1960 | 0.0 | 125.0 | 94.0 | |
| ... | ... | ... | ... | |
| 1768 | 1.0 | 86.0 | 98.0 | |
| 54 | 0.0 | 150.0 | 148.0 | |
| 473 | 0.0 | 150.0 | 152.0 | |
| 1768 | 1.0 | 86.0 | 98.0 | |
| 773 | 0.0 | 130.0 | 130.0 | |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 645 | 124.0 | 0.0 | 0.0 | 1.0 |
| 1766 | 98.0 | 64.0 | -1.0 | 1.0 |
| 697 | 128.0 | 103.0 | 1.0 | 1.0 |
| 2028 | 111.0 | 12.0 | 0.0 | 1.0 |
| 1960 | 97.0 | 147.0 | -1.0 | 1.0 |
| ... | ... | ... | ... | ... |
| 1768 | 95.0 | 44.0 | -1.0 | 1.0 |
| 54 | 150.0 | 0.0 | 1.0 | 1.0 |
| 473 | 153.0 | 2.0 | 0.0 | 1.0 |
| 1768 | 95.0 | 44.0 | -1.0 | 1.0 |
| 773 | 131.0 | 0.0 | 1.0 | 1.0 |

[1950 rows x 22 columns]

```
[23]: data_over = pd.concat([class_1_over, class_0], axis=0)
data_over
```

```
[23]: baseline value accelerations fetal_movement uterine_contractions \
645          123.0          0.000          0.000          0.000
1766          120.0          0.000          0.000          0.005
697          131.0          0.001          0.369          0.003
2028          129.0          0.000          0.001          0.005
```


| | | | | |
|------|-------|-------|-------|-------|
| 1960 | 133.0 | 0.001 | 0.000 | 0.009 |
| ... | ... | ... | ... | ... |
| 2121 | 140.0 | 0.000 | 0.000 | 0.007 |
| 2122 | 140.0 | 0.001 | 0.000 | 0.007 |
| 2123 | 140.0 | 0.001 | 0.000 | 0.007 |
| 2124 | 140.0 | 0.001 | 0.000 | 0.006 |
| 2125 | 142.0 | 0.002 | 0.002 | 0.008 |

| | light_decelerations | severe_decelerations | prolongued_decelerations | \ |
|------|---------------------|----------------------|--------------------------|---|
| 645 | 0.000 | 0.0 | 0.000 | |
| 1766 | 0.006 | 0.0 | 0.001 | |
| 697 | 0.003 | 0.0 | 0.002 | |
| 2028 | 0.008 | 0.0 | 0.002 | |
| 1960 | 0.004 | 0.0 | 0.004 | |
| ... | ... | ... | ... | |
| 2121 | 0.000 | 0.0 | 0.000 | |
| 2122 | 0.000 | 0.0 | 0.000 | |
| 2123 | 0.000 | 0.0 | 0.000 | |
| 2124 | 0.000 | 0.0 | 0.000 | |
| 2125 | 0.000 | 0.0 | 0.000 | |

| | abnormal_short_term_variability | mean_value_of_short_term_variability | \ |
|------|---------------------------------|--------------------------------------|---|
| 645 | 76.0 | 0.2 | |
| 1766 | 62.0 | 1.5 | |
| 697 | 29.0 | 2.0 | |
| 2028 | 65.0 | 2.9 | |
| 1960 | 63.0 | 2.1 | |
| ... | ... | ... | |
| 2121 | 79.0 | 0.2 | |
| 2122 | 78.0 | 0.4 | |
| 2123 | 79.0 | 0.4 | |
| 2124 | 78.0 | 0.4 | |
| 2125 | 74.0 | 0.4 | |

| | percentage_of_time_with_abnormal_long_term_variability | ... | \ |
|------|--|-----|---|
| 645 | 85.0 | ... | |
| 1766 | 0.0 | ... | |
| 697 | 0.0 | ... | |
| 2028 | 0.0 | ... | |
| 1960 | 0.0 | ... | |
| ... | ... | ... | |
| 2121 | 25.0 | ... | |
| 2122 | 22.0 | ... | |
| 2123 | 20.0 | ... | |
| 2124 | 27.0 | ... | |
| 2125 | 36.0 | ... | |

| | histogram_min | histogram_max | histogram_number_of_peaks | \ |
|------|---------------|---------------|---------------------------|---|
| 645 | 118.0 | 129.0 | 1.0 | |
| 1766 | 69.0 | 154.0 | 4.0 | |
| 697 | 54.0 | 165.0 | 14.0 | |
| 2028 | 50.0 | 151.0 | 6.0 | |
| 1960 | 52.0 | 185.0 | 4.0 | |
| ... | ... | ... | ... | |
| 2121 | 137.0 | 177.0 | 4.0 | |
| 2122 | 103.0 | 169.0 | 6.0 | |
| 2123 | 103.0 | 170.0 | 5.0 | |
| 2124 | 103.0 | 169.0 | 6.0 | |
| 2125 | 117.0 | 159.0 | 2.0 | |

| | histogram_number_of_zeroes | histogram_mode | histogram_mean | \ |
|------|----------------------------|----------------|----------------|---|
| 645 | 0.0 | 123.0 | 122.0 | |
| 1766 | 1.0 | 86.0 | 101.0 | |
| 697 | 0.0 | 129.0 | 112.0 | |
| 2028 | 2.0 | 105.0 | 84.0 | |
| 1960 | 0.0 | 125.0 | 94.0 | |
| ... | ... | ... | ... | |
| 2121 | 0.0 | 153.0 | 150.0 | |
| 2122 | 0.0 | 152.0 | 148.0 | |
| 2123 | 0.0 | 153.0 | 148.0 | |
| 2124 | 0.0 | 152.0 | 147.0 | |
| 2125 | 1.0 | 145.0 | 143.0 | |

| | histogram_median | histogram_variance | histogram_tendency | fetal_health |
|------|------------------|--------------------|--------------------|--------------|
| 645 | 124.0 | 0.0 | 0.0 | 1.0 |
| 1766 | 98.0 | 64.0 | -1.0 | 1.0 |
| 697 | 128.0 | 103.0 | 1.0 | 1.0 |
| 2028 | 111.0 | 12.0 | 0.0 | 1.0 |
| 1960 | 97.0 | 147.0 | -1.0 | 1.0 |
| ... | ... | ... | ... | ... |
| 2121 | 152.0 | 2.0 | 0.0 | 0.0 |
| 2122 | 151.0 | 3.0 | 1.0 | 0.0 |
| 2123 | 152.0 | 4.0 | 1.0 | 0.0 |
| 2124 | 151.0 | 4.0 | 1.0 | 0.0 |
| 2125 | 145.0 | 1.0 | 0.0 | 0.0 |

[3900 rows x 22 columns]

```
[24]: X_over = data_over.drop('fetal_health', axis = 1)
y_over = data_over['fetal_health']
x_train_o,x_test_o,y_train_o,y_test_o = \
    ↪train_test_split(X_over,y_over,test_size=0.2, random_state=142)
svm_over = svm.SVC()
svm_over.fit(x_train_o,y_train_o)
```

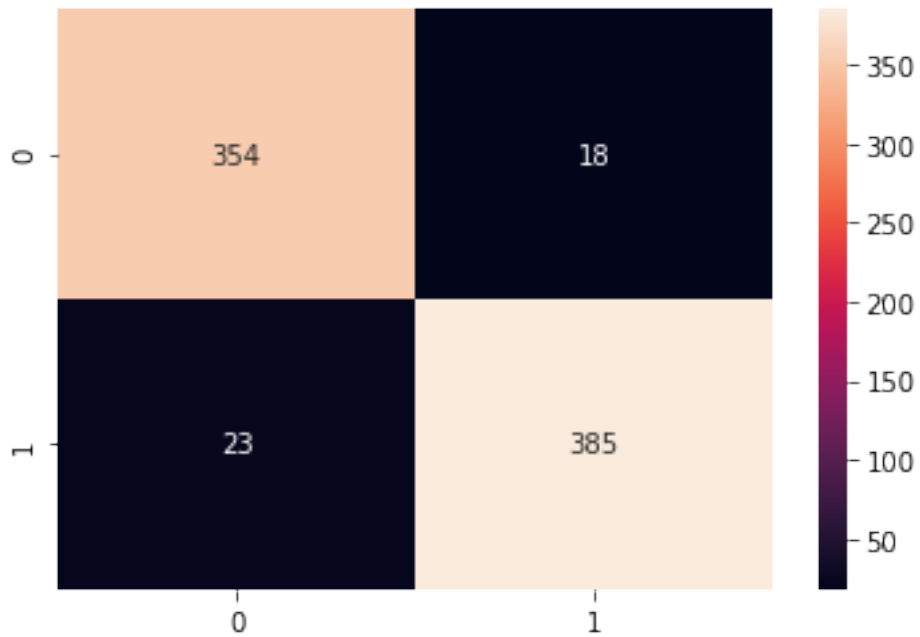
```
y_pred_train_o = svm_over.predict(x_train_o)
y_pred_test_o = svm_over.predict(x_test_o)
```

```
[25]: #confusion matrix

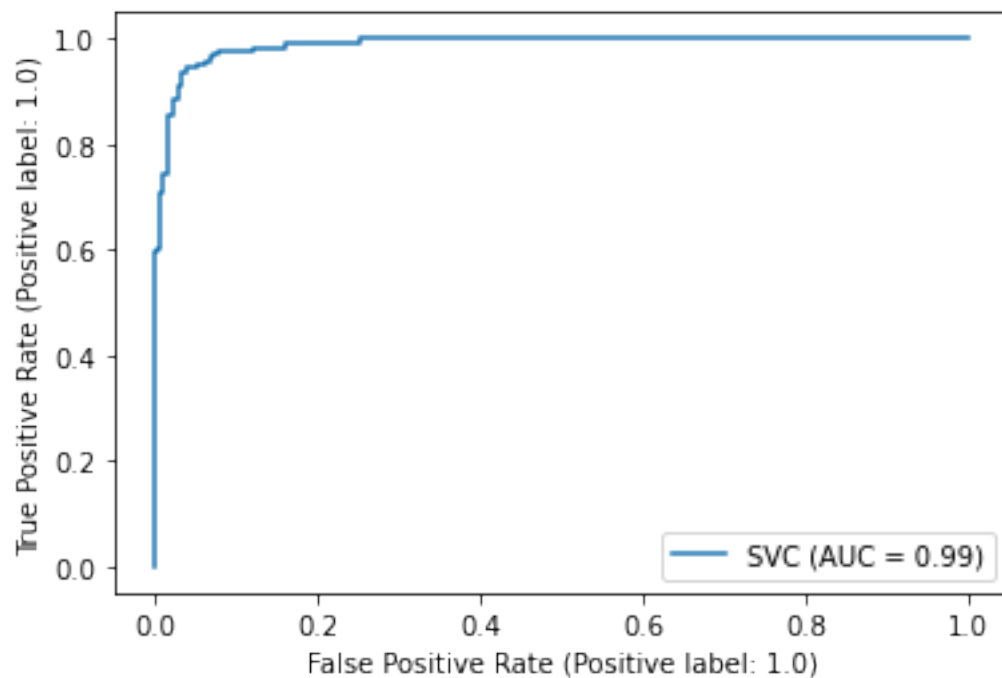
cm = confusion_matrix(y_test_o, y_pred_test_o)

sns.heatmap(cm, annot=True,annot_kws={"size": 10},fmt="d")
print(accuracy_score(y_test_o, y_pred_test_o))
```

0.9474358974358974



```
[26]: metrics.plot_roc_curve(svm_over, x_test_o, y_test_o)
plt.show()
```



```
[27]: from sklearn.metrics import classification_report
print(classification_report(y_test_o, y_pred_test_o))
```

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0.0 | 0.94 | 0.95 | 0.95 | 372 |
| 1.0 | 0.96 | 0.94 | 0.95 | 408 |
| accuracy | | | 0.95 | 780 |
| macro avg | 0.95 | 0.95 | 0.95 | 780 |
| weighted avg | 0.95 | 0.95 | 0.95 | 780 |