

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

```
In [2]: dataset = pd.read_csv('G:\Dattathreya\PROJECT\DATA SETS\Human Stress.csv')
```

```
In [3]: dataset.head(2)
```

```
Out[3]:
```

| | SR | RR | T | LM | BO | REM | SR.1 | HR | SL |
|---|-------|--------|--------|-------|--------|-------|-------|-------|----|
| 0 | 93.80 | 25.680 | 91.840 | 16.60 | 89.840 | 99.60 | 1.840 | 74.20 | 3 |
| 1 | 91.64 | 25.104 | 91.552 | 15.88 | 89.552 | 98.88 | 1.552 | 72.76 | 3 |

```
In [4]: dataset.columns=[['snoring range','respiration rate','body temperature','limb movement rate',
                           'blood oxygen levels','eye movement',
                           'number of hours of sleep', 'heart rate','Stress Levels']]
```

```
In [5]: dataset.head()
```

```
Out[5]:
```

| | snoring range | respiration rate | body temperature | limb movement rate | blood oxygen levels | eye movement | number of hours of sleep | heart rate | Stress Levels |
|---|---------------|------------------|------------------|--------------------|---------------------|--------------|--------------------------|------------|---------------|
| 0 | 93.80 | 25.680 | 91.840 | 16.600 | 89.840 | 99.60 | 1.840 | 74.20 | 3 |
| 1 | 91.64 | 25.104 | 91.552 | 15.880 | 89.552 | 98.88 | 1.552 | 72.76 | 3 |
| 2 | 60.00 | 20.000 | 96.000 | 10.000 | 95.000 | 85.00 | 7.000 | 60.00 | 1 |
| 3 | 85.76 | 23.536 | 90.768 | 13.920 | 88.768 | 96.92 | 0.768 | 68.84 | 3 |
| 4 | 48.12 | 17.248 | 97.872 | 6.496 | 96.248 | 72.48 | 8.248 | 53.12 | 0 |

```
In [6]: corr = dataset.corr()
```

```
In [7]: corr['Stress Levels']
```

Out[7]:

| Stress Levels | |
|--------------------------|-----------|
| snoring range | 0.975322 |
| respiration rate | 0.963516 |
| body temperature | -0.962354 |
| limb movement rate | 0.971071 |
| blood oxygen levels | -0.961092 |
| eye movement | 0.951988 |
| number of hours of sleep | -0.973036 |
| heart rate | 0.963516 |
| Stress Levels | 1.000000 |

```
In [8]: from sklearn.model_selection import train_test_split
```

```
In [9]: x = dataset.iloc[:, :-1]  
x
```

Out[9]:

| | snoring range | respiration rate | body temperature | limb movement rate | blood oxygen levels | eye movement | number of hours of sleep | heart rate |
|-----|------------------|---------------------|---------------------|-----------------------|------------------------|-----------------|-----------------------------|------------|
| 0 | 93.800 | 25.680 | 91.840 | 16.600 | 89.840 | 99.60 | 1.840 | 74.20 |
| 1 | 91.640 | 25.104 | 91.552 | 15.880 | 89.552 | 98.88 | 1.552 | 72.76 |
| 2 | 60.000 | 20.000 | 96.000 | 10.000 | 95.000 | 85.00 | 7.000 | 60.00 |
| 3 | 85.760 | 23.536 | 90.768 | 13.920 | 88.768 | 96.92 | 0.768 | 68.84 |
| 4 | 48.120 | 17.248 | 97.872 | 6.496 | 96.248 | 72.48 | 8.248 | 53.12 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 625 | 69.600 | 20.960 | 92.960 | 10.960 | 90.960 | 89.80 | 3.440 | 62.40 |
| 626 | 48.440 | 17.376 | 98.064 | 6.752 | 96.376 | 73.76 | 8.376 | 53.44 |
| 627 | 97.504 | 27.504 | 86.880 | 17.752 | 84.256 | 101.88 | 0.000 | 78.76 |
| 628 | 58.640 | 19.728 | 95.728 | 9.728 | 94.592 | 84.32 | 6.728 | 59.32 |
| 629 | 73.920 | 21.392 | 93.392 | 11.392 | 91.392 | 91.96 | 4.088 | 63.48 |

630 rows × 8 columns

```
In [10]: y =dataset.iloc[:,-1]
y
```

```
Out[10]: 0      3
1      3
2      1
3      3
4      0
..
625    2
626    0
627    4
628    1
629    2
Name: (Stress Levels,), Length: 630, dtype: int64
```

```
In [11]: xtrain,xtest,ytrain,ytest = train_test_split(x,y,test_size = 0.20)
```

Feature Selection 1

```
In [12]: from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import mutual_info_classif
```

```
In [13]: kbest = SelectKBest(mutual_info_classif,k=4)
selectfeature = kbest.fit(xtrain,ytrain)
```

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.

```
warnings.warn(
```

```
In [14]: xtrain.columns[selectfeature.get_support()]
```

```
Out[14]: MultiIndex([(      'snoring range',),
                    (      'respiration rate',),
                    ('limb movement rate',),
                    (      'heart rate',)],
                  )
```

```
In [15]: x1 = dataset[['snoring range','respiration rate','blood oxygen levels','heart rate']]
```

```
In [16]: x1.head(1)
```

```
Out[16]:
```

| | snoring range | respiration rate | blood oxygen levels | heart rate |
|---|---------------|------------------|---------------------|------------|
| 0 | 93.8 | 25.68 | 89.84 | 74.2 |

```
In [17]: x1train,x1test,y1train,y1test = train_test_split(x1,y,test_size = 0.20)
```

```
In [18]: x1train.head(1)
```

```
Out[18]:
```

| | snoring range | respiration rate | blood oxygen levels | heart rate |
|-----|---------------|------------------|---------------------|------------|
| 117 | 76.0 | 21.6 | 91.6 | 64.0 |

```
In [19]: from sklearn.ensemble import AdaBoostClassifier
model1 = AdaBoostClassifier()
model1.fit(x1train,y1train)
```

```
C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
```

```
warnings.warn(
```

```
Out[19]: AdaBoostClassifier()
```

```
In [20]: print('train score',model1.score(x1train,y1train))
print('test score',model1.score(x1test,y1test))
```

```
train score 0.8214285714285714
```

```
test score 0.7063492063492064
```

```
C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
```

```
warnings.warn(
```

```
C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
```

```
warnings.warn(
```

Feature Selection 2

```
In [21]: from sklearn.feature_selection import SelectPercentile
from sklearn.feature_selection import mutual_info_classif
```

```
In [22]: sp = SelectPercentile(mutual_info_classif,percentile = 2)
selectfeature2 = sp.fit(xtrain,ytrain)
```

```
C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
```

```
warnings.warn(
```

```
In [23]: xtrain.columns[selectfeature2.get_support()]
```

```
Out[23]: MultiIndex([['heart rate',)],
)
```

Train model

```
In [24]: xtrain.head(2)
```

| Out[24]: | snoring range | respiration rate | body temperature | limb movement rate | blood oxygen levels | eye movement | number of hours of sleep | heart rate |
|------------|------------------|---------------------|---------------------|-----------------------|------------------------|-----------------|-----------------------------|------------|
| 215 | 87.44 | 23.984 | 90.992 | 14.480 | 88.992 | 97.48 | 0.992 | 69.96 |
| 147 | 45.08 | 16.032 | 96.048 | 4.064 | 95.032 | 60.32 | 7.032 | 50.08 |

```
In [25]: model2 = AdaBoostClassifier()
model2.fit(xtrain,ytrain)
```

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
warnings.warn(

```
Out[25]: AdaBoostClassifier()
```

```
In [26]: print('train score',model2.score(xtrain,ytrain))
print('test score',model2.score(xtest,ytest))
```

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
warnings.warn(

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
warnings.warn(

```
train score 0.6031746031746031
test score 0.5714285714285714
```

feature importance again train model

```
In [27]: imp = model2.feature_importances_
imp = pd.DataFrame(imp)

feature = xtrain.columns
feature = pd.DataFrame(feature)

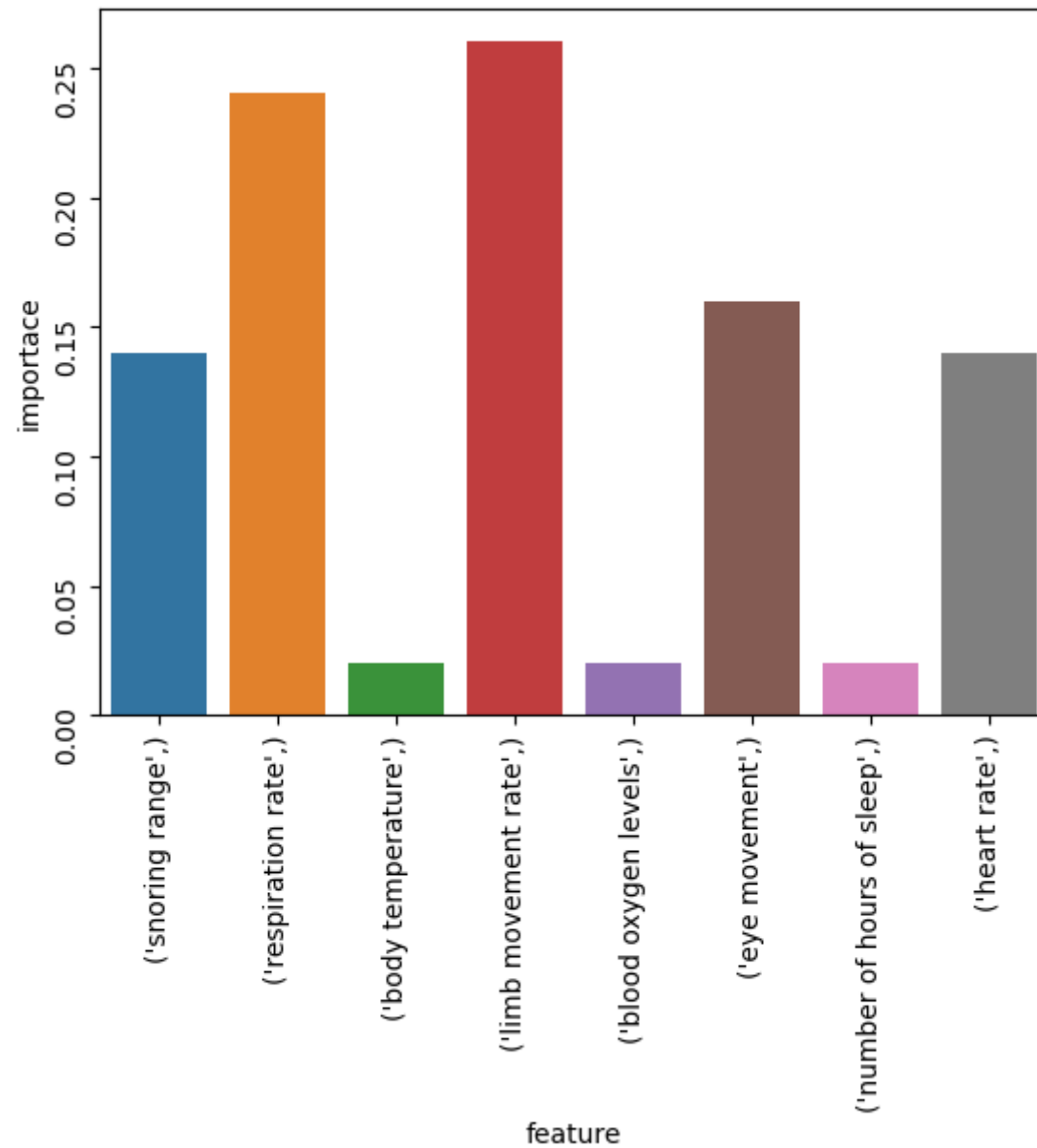
importfeature = pd.concat([imp,feature],axis=1)
importfeature.columns=['importace','feature']

importfeature
```

Out[27]:

| | importace | feature |
|---|-----------|-----------------------------|
| 0 | 0.14 | (snoring range,) |
| 1 | 0.24 | (respiration rate,) |
| 2 | 0.02 | (body temperature,) |
| 3 | 0.26 | (limb movement rate,) |
| 4 | 0.02 | (blood oxygen levels,) |
| 5 | 0.16 | (eye movement,) |
| 6 | 0.02 | (number of hours of sleep,) |
| 7 | 0.14 | (heart rate,) |

```
In [28]: sns.barplot(x= importfeature['feature'],y = importfeature['importace'])  
plt.tick_params(rotation =90)
```



```
In [29]: x2 = dataset[['body temperature','blood oxygen levels','number of hours of sleep','respiration rate']]
```

```
In [30]: x2.head(2)
```


Out[30]:

| | body temperature | blood oxygen levels | number of hours of sleep | respiration rate |
|--|------------------|---------------------|--------------------------|------------------|
|--|------------------|---------------------|--------------------------|------------------|

| | | | | |
|---|--------|--------|-------|--------|
| 0 | 91.840 | 89.840 | 1.840 | 25.680 |
|---|--------|--------|-------|--------|

| | | | | |
|---|--------|--------|-------|--------|
| 1 | 91.552 | 89.552 | 1.552 | 25.104 |
|---|--------|--------|-------|--------|

In [31]: `x2train,x2test,y2train,y2test = train_test_split(x2,y,test_size = 0.20)`

In [32]: `x2train.head(2)`

Out[32]:

| | body temperature | blood oxygen levels | number of hours of sleep | respiration rate |
|--|------------------|---------------------|--------------------------|------------------|
|--|------------------|---------------------|--------------------------|------------------|

| | | | | |
|-----|--------|--------|-------|--------|
| 323 | 85.080 | 82.096 | 0.000 | 26.064 |
|-----|--------|--------|-------|--------|

| | | | | |
|-----|--------|--------|-------|--------|
| 173 | 95.632 | 94.448 | 6.632 | 19.632 |
|-----|--------|--------|-------|--------|

In [33]: `model3 = AdaBoostClassifier()
model3.fit(x2train,y2train)`

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
warnings.warn(

Out[33]: `AdaBoostClassifier()`

In [34]: `print('train score',model3.score(x2train,y2train))
print('test score',model3.score(x2test,y2test))`

train score 0.6130952380952381

test score 0.5396825396825397

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
warnings.warn(

C:\Users\HIMAVANTH\anaconda3\lib\site-packages\sklearn\utils\validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple']. An error will be raised in 1.2.
warnings.warn(