

Python: 3.8.8 (default, Feb 24 2021, 15:54:32) [MSC v.1928 64 bit (AMD64)]

scipy: 1.7.1

numpy: 1.19.2

pandas: 1.1.3

sklearn: 0.24.2

Hello World!

	sepal-length	sepal-width	petal-length	petal-width	class
0	5.4	3.7	1.5	0.2	Iris-setosa
1	4.8	3.4	1.6	0.2	Iris-setosa
2	4.8	3.0	1.4	0.1	Iris-setosa
3	4.3	3.0	1.1	0.1	Iris-setosa
4	5.8	4.0	1.2	0.2	Iris-setosa
..
115	6.7	3.0	5.2	2.3	Iris-virginica
116	6.3	2.5	5.0	1.9	Iris-virginica
117	6.5	3.0	5.2	2.0	Iris-virginica
118	6.2	3.4	5.4	2.3	Iris-virginica
119	5.9	3.0	5.1	1.8	Iris-virginica

[120 rows x 5 columns]

	class	counts
0	Iris-setosa	40
1	Iris-versicolor	30
2	Iris-virginica	50

Part-1

Number of samples in fold1:60

Number of samples in fold2:60

Number of mislabeled points out of a total 120 points : 4

Accuracy of the Neural network model: 0.966667

Confusion Matrix of Neural network:

```
[[40 0 0]
```

```
[ 0 26 4]
```

```
[ 0 0 50]]
```

Class balanced accuracy of the Neural Network Model model: 0.9308641975308642

Balanced accuracy of the Neural Network Model model: 0.9682539682539683

Sklearn balanced accuracy of the Neural Network Model model: 0.9555555555555556

Part-2

#####

Random oversampling

#####

Random oversampled dataset shape Counter({'Iris-setosa': 50, 'Iris-versicolor': 50, 'Iris-virginica': 50})

Number of samples in fold1:75

Number of samples in fold2:75

Number of mislabeled points out of a total 150 points : 5

Accuracy of the Neural network model with Random oversampling: 0.966667

Confusion Matrix of Neural network with Random oversampling:

```
[[50 0 0]
```

```
[ 0 47 3]
```

```
[ 0 2 48]]
```

#####

SMOTE oversampling

#####

SMOTE oversampled dataset shape Counter({'Iris-setosa': 50, 'Iris-versicolor': 50, 'Iris-virginica': 50})

Number of samples in fold1:75

Number of samples in fold2:75

Number of mislabeled points out of a total 150 points : 2

Accuracy of the Neural network model with SMOTE oversampling: 0.986667

Confusion Matrix of Neural network with SMOTE oversampling:

```
[[50 0 0]
```

```
[ 0 48 2]
```

```
[ 0 0 50]]
```

#####

ADASYN oversampling

#####

ADASYN oversampled dataset shape Counter({'Iris-versicolor': 51, 'Iris-virginica': 50, 'Iris-setosa': 40})

Number of samples in fold1:70

Number of samples in fold2:71

Number of mislabeled points out of a total 141 points : 7

Accuracy of the Neural network model with ADASYN oversampling: 0.950355

Confusion Matrix of Neural network with ADASYN oversampling:

```
[[40 0 0]
```

```
[ 0 47 4]
```

```
[ 0 3 47]]
```

Part-3

#####

Random undersampling

#####

Random undersampled dataset shape Counter({'Iris-setosa': 30, 'Iris-versicolor': 30, 'Iris-virginica': 30})

Number of samples in fold1:45

Number of samples in fold2:45

Number of mislabeled points out of a total 90 points : 2

Accuracy of the Neural network model with Random undersampling: 0.977778

Confusion Matrix of Neural network with Random undersampling:

```
[[30 0 0]
```

```
[ 0 28 2]
```

```
[ 0 0 30]]
```

#####

Cluster undersampling

#####

Cluster undersampled dataset shape Counter({'Iris-setosa': 30, 'Iris-versicolor': 30, 'Iris-virginica': 30})

Number of samples in fold1:45

Number of samples in fold2:45

Number of mislabeled points out of a total 90 points : 2

Accuracy of the Neural network model with Cluster undersampling: 0.977778

Confusion Matrix of Neural network with Cluster undersampling:

```
[[30 0 0]
```

```
[ 0 28 2]
```

```
[ 0 0 30]]
```

#####

Tomek Links undersampling

#####

Tomek Links undersampled dataset shape Counter({'Iris-virginica': 49, 'Iris-setosa': 40, 'Iris-versicolor': 30})

Number of samples in fold1:59

Number of samples in fold2:60

Number of mislabeled points out of a total 119 points : 4

Accuracy of the Neural network model with Tomek Links undersampling: 0.966387

Confusion Matrix of Neural network with Tomek Links undersampling:

```
[[40 0 0]
```

```
[ 0 26 4]
```

```
[ 0 0 49]]
```