

## =====Part 1: Imbalanced Data Set=====

Confusion Matrix:

```
[[40  0  0]
 [ 0 26  4]
 [ 0  0 50]]
```

- a.Accuracy: 0.9666666666666667
- b.Class Balanced Accuracy (slide 16): 0.9308641975308642
- c.Balanced Accuracy (slide 16): 0.9805621224060941
- d.Balanced Accuracy (sklearn function): 0.9555555555555556

## =====Part 2: Oversampling=====

a. Random Oversampling:

Confusion Matrix:

```
[[50  0  0]
 [ 0 47  3]
 [ 0  1 49]]
```

Accuracy: 0.9733333333333334

b. SMOTE Oversampling:

Confusion Matrix:

```
[[50  0  0]
 [ 0 46  4]
 [ 0  0 50]]
```

Accuracy: 0.9733333333333334

c. ADASYN Oversampling:

Handling run-time error: Not any neighbours belong to the majority class. This case will induce a NaN case with a division by zero.

ADASYN is not suited for this specific dataset. Use SMOTE instead.

Using SMOTE instead:

Confusion Matrix:

```
[[50  0  0]
 [ 0 46  4]
 [ 0  1 49]]
```

Accuracy: 0.9666666666666667

## =====Part 3: Undersampling=====

a. Random Undersampling:

Confusion Matrix:

```
[[30  0  0]
 [ 0 28  2]
 [ 0  0 30]]
Accuracy: 0.9777777777777777
```

b. Cluster Undersampling:

Confusion Matrix:

```
[[30  0  0]
 [ 0 28  2]
 [ 0  0 30]]
Accuracy: 0.9777777777777777
```

c. Tomek Undersampling:

Confusion Matrix:

```
[[40  0  0]
 [ 0 26  4]
 [ 0  0 49]]
Accuracy: 0.9663865546218487
```