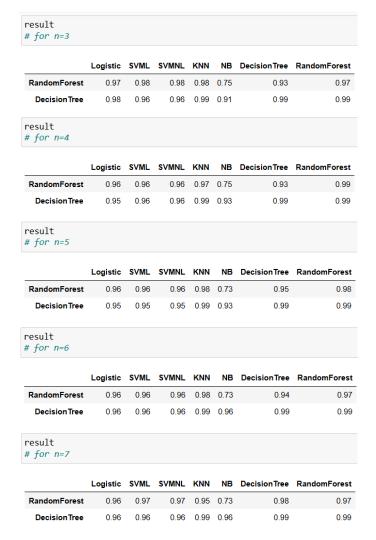# BACKWARD SELECTION / BACKWARD ELIMINATION

Backward Selection is a feature selection method which is used to identify the important features from a huge dataset that improves the model's performance.

Backward Selection algorithm is provided by mlxtend as a function named "SequentialFeatureSelection" from feature_selection.

Backward Selection working follows by initially training the models using all the available features. Next, it removes the feature with the least impact on the model, which is measured and chosen using score metrics. The model is re-evaluated and the model's performance is evaluated after each time a feature is removed. This feature elimination takes place iteratively until fixed number of features are reached or further removal of feature has no positive impact on the model performance.

```
result
# for n=3
```

	Logistic	SVML	SVMNL	KNN	NB	DecisionTree	RandomForest
RandomForest	0.97	0.98	0.98	0.98	0.75	0.93	0.97
DecisionTree	0.98	0.96	0.96	0.99	0.91	0.99	0.99

```
result
# for n=4
```

	Logistic	SVML	SVMNL	KNN	NB	DecisionTree	RandomForest
RandomForest	0.96	0.96	0.96	0.97	0.75	0.93	0.99
DecisionTree	0.95	0.96	0.96	0.99	0.93	0.99	0.99

```
result
# for n=5
```

	Logistic	SVML	SVMNL	KNN	NB	DecisionTree	RandomForest
RandomForest	0.96	0.96	0.96	0.98	0.73	0.95	0.98
DecisionTree	0.95	0.95	0.95	0.99	0.93	0.99	0.99

```
result
# for n=6
```

	Logistic	SVML	SVMNL	KNN	NB	DecisionTree	RandomForest
RandomForest	0.96	0.96	0.96	0.98	0.73	0.94	0.97
DecisionTree	0.96	0.96	0.96	0.99	0.96	0.99	0.99

```
result
# for n=7
```

	Logistic	SVML	SVMNL	KNN	NB	DecisionTree	RandomForest
RandomForest	0.96	0.97	0.97	0.95	0.73	0.98	0.97
DecisionTree	0.96	0.96	0.96	0.99	0.96	0.99	0.99

Conclusion:

From the above displayed tables, the accuracy score results are varying as the number of features increased. For n = 5, the results are stable. Hence, we can finalize the number of features required for the prediction of CKD using backward Selection is n=5.