Solution 4

a) **Comparison of accuracy and time-taken of my implementation Vs Scikitlearn's built-in random forest classifier**

Number of Trees: 5 (in both classifiers)

My implementation

```
Test accuracy:  67.82608695652175 %
CPU times: user 1min 8s, sys: 1.47 s, total: 1min 10s
Wall time: 1min 9s
```

Vs

Sklearn RandomForestClassifier

```
Test Accuracy:  94.02310666792418 %
CPU times: user 1.2 ms, sys: 0 ns, total: 1.2 ms
Wall time: 934 µs
```

b) The sensitivity of Random Forests to the parameter m (the number of features used for best split)

Number of Trees: 5
m = SQRT(features)

```
m: 7
CPU times: user 1min 8s, sys: 1.47 s, total: 1min 10s
Wall time: 1min 9s
==============================
OOB Error:  21.732157177225343 %
Test Error:  32.17391304347826 %
Test accuracy:  67.82608695652175 %
==============================
```

Number of Trees: 5
m = All Features considered

```
m: 58
CPU times: user 10min 16s, sys: 9.57 s, total: 10min 25s
Wall time: 10min 22s
==============================
OOB Error:  19.518072289156628 %
Test Error:  35.79710144927537 %
Test accuracy:  64.20289855072463 %
==============================
```

c) Plot the OOB (out-of-bag) error and the test error against for m = (7, 58)