Import Dataset

In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
data = pd.read_csv(r'data/breast-cancer.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	Age	BMI	Glucose	Insulin	HOMA	Leptin	Adiponectin	Resistin	MCP.1
0	48	23.500000	70	2.707	0.467409	8.8071	9.702400	7.99585	417.114
1	83	20.690495	92	3.115	0.706897	8.8438	5.429285	4.06405	468.786
2	82	23.124670	91	4.498	1.009651	17.9393	22.432040	9.27715	554.697
3	68	21.367521	77	3.226	0.612725	9.8827	7.169560	12.76600	928.220
4	86	21.111111	92	3.549	0.805386	6.6994	4.819240	10.57635	773.920

Pre-processing

Change Classification values (from two, one to one, zero each)

In [4]:
```
data['Classification'] = data['Classification'] - 1
```

In [5]:
```
data = data.reindex(np.random.permutation(data.index))
```

In [6]:
```
X = data.drop(['Classification'], axis = 1)
```

In [7]:
```
X = (X - X.mean()) / X.std()
```

In [8]:
```
y = data['Classification']
```

In [9]:
```
X.head()
```

Out[9]:

	Age	BMI	Glucose	Insulin	HOMA	Leptin	Adiponectin	Resis
86	-0.577289	0.108142	-0.345973	-0.742179	-0.585141	-0.577720	0.006055	0.1117
90	-1.073790	0.648273	1.341029	3.166731	2.891878	0.230587	-0.587415	0.2283
40	1.160463	0.325947	-0.656737	-0.460488	-0.437760	0.101490	-0.410752	-0.5392
39	1.222526	1.594741	-0.967500	-0.608982	-0.540199	-0.251718	-0.300339	0.2046
79	-1.011727	-0.181140	-0.035210	1.194000	0.707508	0.943051	0.484265	1.0577

In [10]:
```
y.head()
```

Out[10]:
```
86    1
90    1
40    0
39    0
79    1
Name: Classification, dtype: int64
```

Classifiers

In [11]:
```
from sklearn.neural_network import MLPClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier, AdaBoostClassifier
from sklearn.naive_bayes import GaussianNB
```

Classifier List

1. OneR
2. Nearest Neighbors
3. Linear SVM (Kernel: Linear)
4. RBF SVM (Kernel: RBF)
5. Decision Tree
6. Random Forest
7. Neural Network
8. AdaBoost (Based on the Decision Tree classifier)
9. Gaussian Naive Bayes

In [12]:
```
names = ["OneR", "Nearest Neighbors", "Linear SVM", "RBF SVM",
         "Decision Tree", "Random Forest", "Neural Net", "AdaBoost", "Naive Bayes"]
```

In [13]:
```
classifiers = [
    DecisionTreeClassifier(max_depth=1),
    KNeighborsClassifier(3),
    SVC(kernel="linear", C=0.025),
    SVC(kernel="rbf", C=0.025),
    DecisionTreeClassifier(max_depth=8),
    RandomForestClassifier(max_depth=8, n_estimators=50),
    MLPClassifier(alpha=1, hidden_layer_sizes = (50, 30)),
    AdaBoostClassifier(n_estimators = 100),
    GaussianNB()]
```

In [14]:
```
from sklearn.model_selection import cross_validate
from sklearn.model_selection import cross_val_predict
from sklearn.metrics import precision_recall_fscore_support
```

In [15]:
```
evaluations = []
```

ZeroR

In [16]:
```
max_index = data.groupby(['Classification'])['Classification'].count().idxmax()
```

In [17]:
```
evaluations.append(['ZeroR', data.groupby(['Classification'])['Classification'].coun
t()[max_index] / len(y)])
```

Using Classifier List above

In [18]:
```
for name, clf in zip(names, classifiers):
    y_pred = cross_val_predict(clf, X, y, cv = 10)
    precision_recall_f1 = precision_recall_fscore_support(y, y_pred, average='weight
ed')
    scores = cross_validate(clf, X, y, cv = 10)
    evaluations.append([name, scores['test_score'].mean(), precision_recall_f1[:3]])
```

```
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/metrics/classification.py:11
35: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to
0.0 in labels with no predicted samples.
  'precision', 'predicted', average, warn_for)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
```

```
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
/anaconda/envs/py35/lib/python3.5/site-packages/sklearn/neural_network/multilayer_pe
rceptron.py:564: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200)
reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```

In [19]:

```
evaluations.sort(key = lambda evaluations: evaluations[1])
evaluations.reverse()
```

In [20]:
```
for evaluation in evaluations:
    print("-----" * 5)
    print(evaluation[0])
    print("Accuracy: ", evaluation[1])
    if len(evaluation) > 2: print("Precision, Recall, F1", evaluation[2])
    print("-----" * 5, end = "₩n₩n")
```

```
------------------------------
Neural Net
Accuracy:  0.7469114219114219
Precision, Recall, F1 (0.7520119225037258, 0.75, 0.7505046728971962)
------------------------------

------------------------------
Linear SVM
Accuracy:  0.7203962703962704
Precision, Recall, F1 (0.7253944712231698, 0.7241379310344828, 0.7245510218834876)
------------------------------

------------------------------
Random Forest
Accuracy:  0.7058857808857809
Precision, Recall, F1 (0.7145469818282348, 0.7155172413793104, 0.7136807536081583)
------------------------------

------------------------------
Decision Tree
Accuracy:  0.6953962703962705
Precision, Recall, F1 (0.697789566755084, 0.6982758620689655, 0.6979806551905803)
------------------------------

------------------------------
OneR
Accuracy:  0.6891025641025641
Precision, Recall, F1 (0.6909941086804269, 0.6896551724137931, 0.6901198996189237)
------------------------------

------------------------------
Nearest Neighbors
Accuracy:  0.6820512820512821
Precision, Recall, F1 (0.6905390821727736, 0.6810344827586207, 0.6816749281342077)
------------------------------

------------------------------
AdaBoost
Accuracy:  0.6683566433566434
Precision, Recall, F1 (0.6777645659928656, 0.6724137931034483, 0.6732925638522022)
------------------------------

------------------------------
Naive Bayes
Accuracy:  0.6192890442890443
Precision, Recall, F1 (0.6937602627257801, 0.6206896551724138, 0.6020960108181204)
------------------------------

------------------------------
ZeroR
Accuracy:  0.5517241379310345
------------------------------

------------------------------
RBF SVM
Accuracy:  0.5516317016317016
Precision, Recall, F1 (0.30439952437574314, 0.5517241379310345, 0.39233716475095787)
```

_____