I.Dataset Url:

http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.data

II.Preprocessing :

There were no NA or the empty values in the data.So, it was not required to remove them.

1. Normalized all the values between 0 and 1

Pseudocode:

```
numFolds = 10
classifiers = {Decision Trees, Perceptron, Neural Net, Deep Learning, SVM, naïve Bayes, Logistic Regression, k-Nearest Neighbors, Bagging, Random Forests, AdaBoost, Gradient Boosting} // list of n classifiers with best parameters
split the data into 10 folds d[1…10]
for i in 1 to 10
// create training dataset by combining all folds except d[i]
train = {d[1] + d[2] + … + d[i-1] + d[i+1] + … + d[10]}
// create test dataset using d[i]
test = d[i]
for c in classifiers classifiers
// create a model of type c using train
model <- createModel(c, train)
// find accuracy of model of type c on test
for classifier c: accuracy[i] <- findAccuracy(model, test)
for classifier c: other_parameter[i] <- findEvaluation(model, test)
next c
next i
print average accuracies
print precisions
```

Evaluation Metrics used:

1. Accuracy
2. Precision

III.Finding best classifier parameters:

Experiment#	Classifier	CrossValidation fold	Parameter1	Parameter2	Parameter3	Parameter4	Average Accuracy of 10 folds
1	Decision Trees	10	maxDepth:30	Cp: 0.01	Minbucket: 10	Minsplit: 3	78.5
	Decision Trees	10	maxDepth:30	Cp: 0.01	Minbucket: 1	Minsplit: 3	84.5
	Decision Trees	10	maxDepth:30	Cp: 0.001	Minbucket: 1	Minsplit: 3	94.5
2	Perceptron	10	Threshold: 500				55.92
	Perceptron	10	Threshold:100				87.3
	Perceptron	10	Threshold:10				98.6
3	Neural Net	10	#of Hidden layers:3	Nodes: (3,3,3)	Threshold:10		65.02
	Neural Net	10	#of Hidden layers:4	Nodes:(4,4,4,4)	Threshold:10		62.9
	Neural Net	10	#of Hidden layers:3	Nodes :(4,4,4)	Threshold:1		98.1
4	Deep Learning	10	#of Hidden layers:7	Nodes : 10 each	Threshold:10		65.0
	Deep Learning	10	#of Hidden layers:9	Nodes :10 each	Threshold:0.1		93.9
	Deep Learning	10	#of Hidden layers:12	Nodes :10 each	Threshold:0.1		98.24
5	SVM	10	Kernel: radial				88.17
	SVM	10	Kernel:				65.02

			polynomial				
	SVM	10	Kernel: sigmoid				66.34
6	naïve Bayes	10					96.3
7	Logistic Regression	10	maxit=10	trace= true			96.75
	Logistic Regression	10					
8	k-Nearest Neighbors	10	K = 2				55.5
	k-Nearest Neighbors	10	k=50				63.6
	k-Nearest Neighbors	10	k=300				65
9	Bagging	10	Nb bag: 0.1	coob=true			96.0
	Bagging	10	Nb bag: 100	coob=true			96.3
	Bagging	10	Nb bag: 100	coob=false			95.9
10	Random Forests	10	Importance: true	Proximity: true	Ntree: 1		94
	Random Forests	10	Importance: false	Proximity: false	Ntree: 100		96
	Random Forests	10	Importance: true	Proximity: true	Ntree: 500		97.3
11	AdaBoost	10	Iter: 20	Nu: 1			96.7
	AdaBoost	10	Iter: 2	Nu: 1			94
	AdaBoost	10	Iter: 20	Nu: 2			94.2
12	Gradient Boosting	10	Minobsinnode: 1	Trees: 1	Distribution: gaussian		74.3
	Gradient Boosting	10	Minobsinnode: 1	Trees: 1	Distribution: bernoulli		84.6

	Gradient Boosting	10	Minobsinnode: 1	Trees: 100	Distribution: bernoulli		71.3

# of instances in dataset: 699
# of attributes in dataset: 10
10 fold cross-validation performed

IV.Results with best parameters:

Classifier	**Best Parameters**	**Accuracy**	**Precision**
Decision Trees	maxDepth:30 Cp: 0.001 Minbucket: 1 Minsplit: 3	94.5	89.5
Perceptron	Threshold:10	98.6	97.8
Neural Net	#of Hidden layers:3 Nodes :(4,4,4) Threshold:1	98.1	98.2
Deep Learning	#of Hidden layers:12 Nodes :10 each Threshold:0.1	98.24	97.9
SVM	Kernel: radial	88.17	169.2
naïve Bayes		96.3	96.5
Logistic Regression	maxit=10 trace= true	96.75	95.6
k-Nearest Neighbors	k=300	65	169.6
Bagging	Nb bag: 100 coob=true	96.3	179.12
Random Forests	Importance: true Proximity: true Ntree: 500	97.3	96
AdaBoost	Iter: 20 Nu: 1	96.7	96.9
Gradient Boosting	Minobsinnode:1 Trees: 1 Distribution: bernoulli	84.6	73.8

Analysis:

1. The algorithm working majorly depends on the data we select.
2. Perceptron and Neural Net gave the best results on the data where SVM, Gradient boost were not effective on this data.
3. Accuracy gave the best measure of the correctness. Though we were able to identify the precision of the classifier, it didn't give a clear picture of the correctness.
4. On contrary to theory, the number of nodes in the neural net or in deep learning did not change the accuracy too much as the data is scaled between 0&1 and weights were also very low.