APPLIED MACHINE LEARNING

PA2 REPORT

RAMPRASAD BOMMAGANTY AND HARISH ANNAVAJJALA

Question 1: BAGGING

****

Depth : 3 and Bags = 5

Incorrectly classified= 532

Correctly classified= 1593

Accuracy for a depth of 3 is 74.9647058824 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1561 False Positive: 532

False Negative: 0 True Positive: 32

****

Depth : 3 and Bags = 10

Incorrectly classified= 532

Correctly classified= 1593

Accuracy for a depth of 3 is 74.9647058824 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1561 False Positive: 532

False Negative: 0 True Positive: 32

****

Depth : 5 and Bags: 5

Incorrectly classified= 532

Correctly classified= 1593

Accuracy for a depth of 5 is 74.9647058824 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1561 False Positive: 532

False Negative: 0 True Positive: 32

****

Depth: 5 and Bags: 10

Incorrectly classified= 532

Correctly classified= 1593

Accuracy for a depth of 5 is 74.9647058824 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1561 False Positive: 532

False Negative: 0 True Positive: 32

****

QUESTION 2: BOOSTING

****

Depth: 1 Bags: 5

Incorrectly classified= 496

Correctly classified= 1629

Accuracy for a depth of 1 is 76.6588235294 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1597 False Positive: 496

False Negative: 0 True Positive: 32

****

Depth : 1 Bags : 10

Incorrectly classified= 514

Correctly classified= 1611

Accuracy for a depth of 1 is 75.8117647059 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1579 False Positive: 514

False Negative: 0 True Positive: 32

****

Depth : 2 Bags: 5

Incorrectly classified= 532

Correctly classified= 1593

Accuracy for a depth of 2 is 74.9647058824 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1561 False Positive: 532

False Negative: 0 True Positive: 32

****

Depth :2 Bags : 10

Incorrectly classified= 532

Correctly classified= 1593

Accuracy for a depth of 2 is 74.9647058824 %

The confusion matrix is as follows:

Predicted

Actual

True Negative: 1561 False Positive: 532

False Negative: 0 True Positive: 32

****

Note: Question 3 has been covered in Questions 1 and 2.

****

Question 4:

BAGGING and BOOSTING IN WEKA:

Running bagging in WEKA for bags : 5

OUTPUT:

=== Summary ===

Correctly Classified Instances	1593	74.9647 %
Incorrectly Classified Instances	532	25.0353 %
Kappa statistic	0.0812	
Mean absolute error	0.2504	
Root mean squared error	0.5004	
Relative absolute error	45.0609 %	
Root relative squared error	90.0298 %	
Total Number of Instances	2125	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.746	0.000	1.000	0.746	0.854	0.206	0.873	0.996	0
	1.000	0.254	0.057	1.000	0.107	0.206	0.873	0.057	1
Weighted Avg.	0.750	0.004	0.986	0.750	0.843	0.206	0.873	0.982	

=== Confusion Matrix ===

a	b	<-- classified as
1561	532	a = 0
0	32	b = 1

****

Running bagging in WEKA for bags : 10

OUTPUT:

=== Summary ===

Correctly Classified Instances	1593	74.9647 %
Incorrectly Classified Instances	532	25.0353 %
Kappa statistic	0.0812	
Mean absolute error	0.2504	
Root mean squared error	0.5004	
Relative absolute error	45.0609 %	
Root relative squared error	90.0298 %	
Total Number of Instances	2125	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.746	0.000	1.000	0.746	0.854	0.206	0.873	0.996	0
	1.000	0.254	0.057	1.000	0.107	0.206	0.873	0.057	1
Weighted Avg.	0.750	0.004	0.986	0.750	0.843	0.206	0.873	0.982	

=== Confusion Matrix ===

a	b	<-- classified as
1561	532	a = 0
0	32	b = 1

****

Boosting with iterations : 5

OUTPUT:

=== Summary ===

Correctly Classified Instances	1629	76.6588 %
Incorrectly Classified Instances	496	23.3412 %
Kappa statistic	0.0884	
Mean absolute error	0.2363	
Root mean squared error	0.4799	
Relative absolute error	42.5389 %	
Root relative squared error	86.3464 %	
Total Number of Instances	2125	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.763	0.000	1.000	0.763	0.866	0.215	0.966	0.999	0
	1.000	0.237	0.061	1.000	0.114	0.215	0.966	0.182	1
Weighted Avg.	0.767	0.004	0.986	0.767	0.854	0.215	0.966	0.987	

=== Confusion Matrix ===

a	b	<-- classified as
1597	496	a = 0
0	32	b = 1

****

Boosting with iterations : 10

OUTPUT:

=== Summary ===

Correctly Classified Instances	1611	75.8118 %
Incorrectly Classified Instances	514	24.1882 %
Kappa statistic	0.0847	
Mean absolute error	0.2452	
Root mean squared error	0.4889	
Relative absolute error	44.1395 %	
Root relative squared error	87.9716 %	
Total Number of Instances	2125	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.754	0.000	1.000	0.754	0.860	0.210	0.994	1.000	0
	1.000	0.246	0.059	1.000	0.111	0.210	0.994	0.571	1
Weighted Avg.	0.758	0.004	0.986	0.758	0.849	0.210	0.994	0.993	

=== Confusion Matrix ===

a	b	<-- classified as
1579	514	a = 0

0   32 | b = 1

****