IRIS DATASET

Applying Decision Stump on Iris Data set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.trees.DecisionStump -t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances	*100*	*66.6667 %*
Incorrectly Classified Instances	*50*	*33.3333 %*

Applying Bagging on Decision Stump:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.trees.DecisionStump-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances	*118*	*78.6667 %*
Incorrectly Classified Instances	*32*	*21.3333 %*

Applying Boosting on Decision Stump:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.trees.DecisionStump-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances	*142*	*94.6667 %*
Incorrectly Classified Instances	*8*	*5.3333 %*

Applying Naïve Bayes on Iris Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 144 96 %

Incorrectly Classified Instances 6 4 %

Applying Bagging on Naïve Bayes

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 143 95.3333 %

Incorrectly Classified Instances 7 4.6667 %

Applying Boosting on Naïve Bayes

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 141 94 %

Incorrectly Classified Instances 9 6 %

Applying Multi Layer Perceptron on Iris Data set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 100 -V 0 -S 0 -E 20 -H 10-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 144 96 %

Incorrectly Classified Instances 6 4 %

Applying Bagging on Multi Layer Perceptron:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.functions.MultilayerPerceptron -- -L 0.3 -M 0.2 -N 100 -V 0 -S 0 -E 20 -H 10-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 144 96 %

Incorrectly Classified Instances 6 4 %

Applying Boosting on Multi Layer Perceptron:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.functions.MultilayerPerceptron -- -L 0.3 -M 0.2 -N 100 -V 0 -S 0 -E 20 -H 10-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 144 96 %

Incorrectly Classified Instances 6 4 %

Analysis of Iris Data Set using Instance Based Learning (K-Nearest):

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 143 95.3333 %

Incorrectly Classified Instances 7 4.6667 %

Analysis of Bagging on K-Nearest:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""-t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 144 96 %

Incorrectly Classified Instances 6 4 %

Analysis of Boosting on K-Nearest:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.lazy.IBk -- -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""t " \home\workspace\Boosting_Bagging\iris.arff"

Correctly Classified Instances 143 95.3333 %

Incorrectly Classified Instances 7 4.6667 %

ADULT DATA SET

Analysis of Naïve Bayes on Adult Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27165 83.428 %

Incorrectly Classified Instances 5396 16.572 %

Analysis of Bagging on Naïve Bayes:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27123 83.299 %

Incorrectly Classified Instances 5438 16.701 %

Analysis of Boosting on Naïve Bayes:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27165 83.428 %

Incorrectly Classified Instances 5396 16.572 %

Analysis of Logistic Regression on Adult Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.functions.Logistic -R 1.0E-8 -M 10-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27660 84.9483 %

Incorrectly Classified Instances 4901 15.0517 %

Analysis of Bagging on Logistic Regression:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M 10-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27657 84.939 %

Incorrectly Classified Instances 4904 15.061 %

Analysis of Boosting on Logistic Regression:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M 10-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27660 84.9483 %

Incorrectly Classified Instances 4901 15.0517 %

Analysis of Multi Layer Perceptron on Adult Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 100 -V 0 -S 0 -E 20 -H 10-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27109 83.2574 %

Incorrectly Classified Instances 5452 16.7426 %

Analysis of Bagging on Adult Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.functions.MultilayerPerceptron -- -L 0.3 -M 0.2 -N 100 -V 0 -S 0 -E 20 -H 10-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27271 83.7565 %

Incorrectly Classified Instances 5290 16.2435 %

Analysis of Boosting on Adult Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.functions.MultilayerPerceptron -- -L 0.3 -M 0.2 -N 100 -V 0 -S 0 -E 20 -H 10-t " \home\workspace\Boosting_Bagging\adult.arff"

Correctly Classified Instances 27127 83.3116 %

Incorrectly Classified Instances 5434 16.6884 %

CREDIT APPROVAL DATA SET

Analysis of Naïve Bayes on Credit Approval Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.bayes.NaiveBayes -t " \home\workspace\Boosting_Bagging\credit_approval.arff"

Correctly Classified Instances 536 77.6812 %

Incorrectly Classified Instances 154 22.3188 %

Analysis of Bagging on Naïve Bayes:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.bayes.NaiveBayes -t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 534 77.3913 %

Incorrectly Classified Instances 156 22.6087 %

Analysis of Boosting on Naïve Bayes:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.bayes.NaiveBayes-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 562 81.4493 %

Incorrectly Classified Instances 128 18.5507 %

Analysis of J48 on Credit Approval Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.trees.J48 -C 0.25 -M 2 -A-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 594 86.087 %

Incorrectly Classified Instances 96 13.913 %

Analysis of Bagging on J48:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.trees.J48 -- -C 0.25 -M 2 -A-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 601 87.1014 %

Incorrectly Classified Instances 89 12.8986 %

Analysis of Boosting on J48:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.trees.J48 -- -C 0.25 -M 2 -A-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 595 86.2319 %

Incorrectly Classified Instances 95 13.7681 %

Analysis of K-Nearest on Credit Approval Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.lazy.IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 560 81.1594 %

Incorrectly Classified Instances 130 18.8406 %

Analysis of Boosting on Credit Approval Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.lazy.IBk -- -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 560 81.1594 %

Incorrectly Classified Instances 130 18.8406 %

Analysis of Bagging on Credit Approval Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.lazy.IBk -- -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""- t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 562 81.4493 %

Incorrectly Classified Instances 128 18.5507 %

Analysis of Logistic Regression on Credit Approval Data Set:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M 10-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 578 83.7681 %

Incorrectly Classified Instances 112 16.2319 %

Analysis of Bagging on Logistic Regression:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 30 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M 10-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 586 84.9275 %

Incorrectly Classified Instances 104 15.0725 %

Analysis of Boosting on Logistic Regression:

java -cp "%CLASSPATH%;C:\Users\Snow_Leopard\workspace\tmp\Weka\weka-3-7-12\weka.jar" weka.classifiers.meta.AdaBoostM1 -P 100000 -S 1 -I 30 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M 10-t " \home\workspace\Boosting_Bagging\ credit_approval.arff"

Correctly Classified Instances 578 83.7681 %

Incorrectly Classified Instances 112 16.2319 %

Adult Dataset

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Naive Bayes	*83.428%*	*83.299%*	*83.428%*
Logistic	*84.9483%*	*84.939%*	*84.9483%*
MultiPerceptron	*83.2574%*	*83.7565%*	*83.3116%*

Credit Approval

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Naive Bayes	*77.6812 %*	*77.3913%*	*81.4493%*
Logistic	*83.7681%*	*84.9275%*	*83.7681%*
J48	*86.087 %*	*87.1014%*	*86.2319%*
K-Nearest	*81.1594%*	*81.1594%*	*81.4493%*

Iris

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Decision Stump	*66.6667%*	*78.6667%*	*94.6667%*
Naïve Bayes	*95.3333%*	*96%*	*94%*
MultiPerceptron	*96%*	*96%*	*96%*
K-Nearest	*95.3333%*	*96%*	*95.3333%*

With 100 iterations (Credit Approval):

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Naïve Bayes	*77.6812%*	*77.5342%*	*81.4493%*
K-Nearest	*81.1594%*	*81.0145%*	*81.1594%*

Adult Dataset (Iteration = 100)

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Naive Bayes	*83.428%*	*83.299%*	*83.428%*
Logistic	*84.9483%*	*84.939%*	*84.9483%*
MultiPerceptron	*83.2574%*	*83.7565%*	*83.3116%*

With 100 iterations (Iris):

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Decision Stump	*66.6667%*	*86%*	*93.3333%*
Naïve Bayes	*95.3333%*	*96%*	*94%*
MultiPerceptron	*96%*	*96%*	*96%*
K-Nearest	*95.3333%*	*96%*	*95.3333%*

With 150 iterations (Iris Data Set)

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Decision Stump	*66.6667%*	*86%*	*92.6667%*
K-Nearest	*95.3333%*	*96%*	*95.3333%*
MultiPerceptron	*96%*	*96%*	*96%*
K-Nearest	*95.3333%*	*96%*	*95.3333%*

With 150 Iterations (Adult Data Set)

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Naive Bayes	*83.428%*	*83.299%*	*83.428%*
Logistic	*84.9483%*	*84.939%*	*84.9483%*
MultiPerceptron	*83.2574%*	*83.7565%*	*83.3116%*

With 100 iterations (Credit Approval):

Base Learner	*Vanilla*	*Bagging*	*Boosting*
Naïve Bayes	*77.6812%*	*77.5362%*	*81.4493%*
K-Nearest	*81.1594%*	*85.5072%*	*84.2029%*

The algorithms which have been improved by Bagging are as under:

1. Decision Stump (Significant Improvement)
2. J-48 (Less Significant)
3. K-Nearest (Less Significant)
4. Multi Layer Perceptron (Less Significant)

The algorithms which have been improved by Boosting are as under:

1. Decision Stump (Significant Improvement)
2. Naïve Bayes (Moderate Improvement – Credit Approval)
3. J-48 (Less Significant)
4. Multi Layer Perceptron (Low Increase – Adult Data Set)
5. K-Nearest (Less Increase – Adult Data Set)

No improvement was found for the Logistic Regression Algorithm.

Bagging works by reducing the variance by averaging where as Boosting works by initializing the data weighing coefficients for the data set, evaluating the quantities and updating the coefficients back again. Boosting works by reducing the bias through weight coefficient update.

Decision Stump has high variance due to which it is effectively worked over by Bagging. Boosting also works well on decision stump as it can work both on unstable and stable classifiers. Similarly, boosting works significantly well on Naïve Bayes which is a stable algorithm and increases its accuracy for the Credit Approval Data Set.

J-48 and Decision Stump in summary have variance and are hence effectively worked upon by Bagging. They are improved by Boosting as well by reducing the bias in these algorithms. There is a very less improvement with K-Nearest and Multi Layer Perceptron indicating the results had moderate variance although Multi Layer Perceptron is known to be an unstable algorithm.

On the other side, Naïve Bayes works well with Boosting which means for Credit Approval Data Set, the algorithm (NB) threw high bias and hence improved upon by Boosting. There were slight improvements in J48, Multi Layer Perceptron and K-Nearest indicating moderate bias.

Logistic Regression has remained unbiased for both bagging and boosting and did not show any improvement.