Assignment 2: Modeling assignment: Using meta learning schemes with a strong and a weak learner for classification.

Shaun Pritchard

Florida Atlantic University

CAP 6778

October -03-2021

M. Khoshgoftaar

**Assignment 2: Modeling assignment: Using meta learning schemes with a strong and a weak learner for classification.**

Using the [*Lymphoma95x4026.arff*](https://canvas.fau.edu/courses/110070/files/23565112?wrap=1) microarray gene dataset while implementing the same methods used in Assignment 1. The experiment aimed to determine which model provides the best (optimal cost ratio) by applying 10-fold cross-validation to the data set while using the default settings for the meta learners (bagging, boosting), the learner (J48, decision stump), and varying the type II cost ratio from N {1, 1.5, 2 ,2.5, 3, 3.5, 4, 4.5, 5...}.

The results are then compared to the assignment I part 4 result implementations utilizing cost sensitive classifiers with varying ratios on J48 leaner for different 4 variations of ensemble meta learners and classifiers each varying the cost ratio N {1, 1.5, 2 ,2.5, 3, 3.5, 4, 4.5, 5...}.

The assignment I results were then compared to 4 different variations of ensemble meta learners and classifiers with the iteration of the meta learners (bagging, boosting) set to 25. The original meta learner interaction is normally set to a default of 10. This study implemented a total of 72 instances of analysis variations and 81 instances all together with assignment I classification analysis as follows.

For this project the experiments impacted the following learners and classification methods to perform the analysis in comparison with Assignment I experiments which were implemented with cost sensitive boosting and only J48.

* Cost sensitive classifier combined with bagging and J48
* Cost sensitive classifier combined with bagging and Decision Stump
* Cost sensitive classifier combined with boosting (AdaBoostM1) and J48
* Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump

Then the ensemble experiments were altercated by changing the met-learner increments to 25 from the default of 10. The following Table:0-1 shows the best performing classifica learners with best cost ratio. The result shows that cost sensitive boosting with decision stumps had the least amount of balanced Type II errors overall with a cost ratio of 1:1.5. The best experiment proved to be the cost sensitive classifier with AdaboostM1 and decision stump leaner with a cost ratio of 1:1, ROC 94%. Other learners performed very well in comparison to the assignment I learners with only J48, cost ratio of 1:1 ~ 1:1.5, and 9 type II errors.

Experiments with default 10-iterations:

| (10-iterations) | Cost Ratio | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I Error rate % | Type II Error rate % |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| j48 | 1:4.5 | 9 | 4 | 0.863 | 0.162 | 0.887 | 13.60% | 0.877 | 17.40% | 12.50% |
| DS | 1:5 | 19 | 3 | 0.768 | 0.163 | 0.841 | 23.10% | 0.866 | 13.00% | 26.40% |
| Ada+J48 | 1:1 | 5 | 7 | 0.874 | 0.247 | 0.871 | 12.60% | 0.83 | 30.40% | 6.90% |
| Ada+DS | 1:1 ~1:1.5 | 6 | 8 | 0.853 | 0.284 | 0.849 | 14.70% | 0.87 | 34.80% | 8.30% |

Experiments with 25-iterations:

| (25-iterations) | Cost Ratio | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I Error rate % | Type II Error rate % |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| j48 | 1:2.5 | 4 | 7 | 0.884 | 0.244 | 0.881 | 11.50% | 0.886 | 30.40% | 5.60% |
| DS | 1:3 | 18 | 6 | 0.747 | 0.258 | 0.8 | 25.20% | 0.842 | 26.10% | 25.00% |
| Ada+J48 | 1:1.5 | 4 | 8 | 0.874 | 0.277 | 0.869 | 12.60% | 0.909 | 43.50% | 4.20% |
| Ada+DS | 1:1 | 5 | 5 | 0.895 | 0.182 | 0.895 | 10.50% | 0.94 | 21.70% | 6.90% |

Assignment I J48 only:

|  | Cost Ratio | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I Error rate % | Type II Error rate % |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| J48 Only | 1:1 ~1.5 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 | 87.00% | 53.80% |

1. **Cost sensitive classifier combined with bagging and J48 compared to cost classifier with j48**

The following results in Table:1-1 show the best result for assignment II at cost ratio variance set to 4.5 with ROC of 0.877. Although at a 1:1 ratio we see less Type I and Type II errors with the lowest classification rate at 11.50%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 22%. Assignment II parameters show more significance for classification than the comparison.

Table:1-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 cost classifier ,bagging, j48** | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 3 | 8 | 0.884 | 0.274 | 0.881 | 11.50% | 0.917 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 4 | 10 | 0.853 | 0.343 | 0.853 | 14.70% | 0.876 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 5 | 10 | 0.842 | 0.346 | 0.834 | 15.70% | 0.875 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 4 | 9 | 0.863 | 0.31 | 0.858 | 13.60% | 0.883 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 7 | 9 | 0.832 | 0.32 | 0.827 | 16.80% | 0.866 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 7 | 9 | 0.832 | 0.32 | 0.827 | 16.80% | 0.866 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 8 | 8 | 0.832 | 0.291 | 0.832 | 16.80% | 0.848 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 9 | 4 | 0.863 | 0.162 | 0.887 | 13.60% | 0.877 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24.20% | 0.635 |
| 1 | 5 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.828 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24.20% | 0.635 |

Chart:1-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with bagging and J48 learner. Type II error decreases as CSC variation is increased showing better performance.

Chart:1-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with bagging on the J48 learner against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging technique.

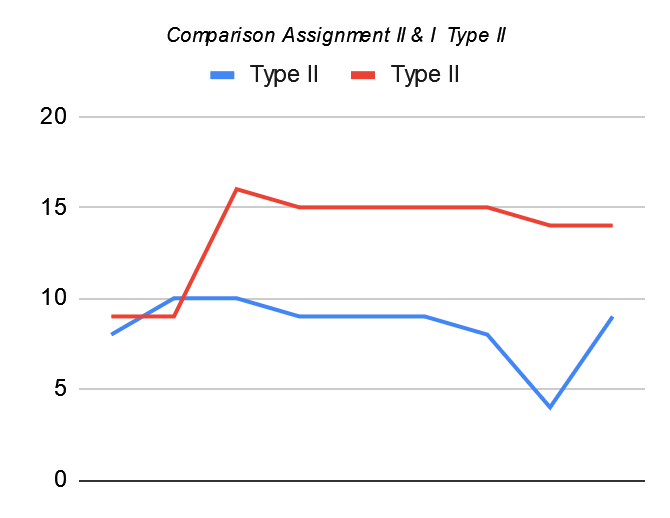
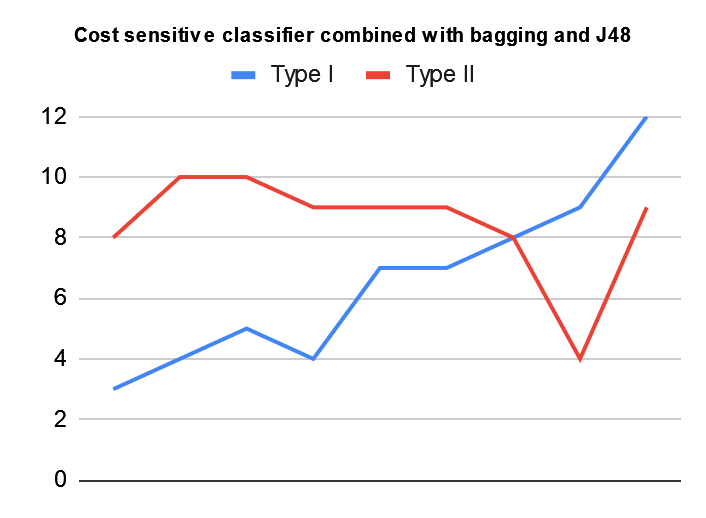


Chart:1-1 Chart:1-2

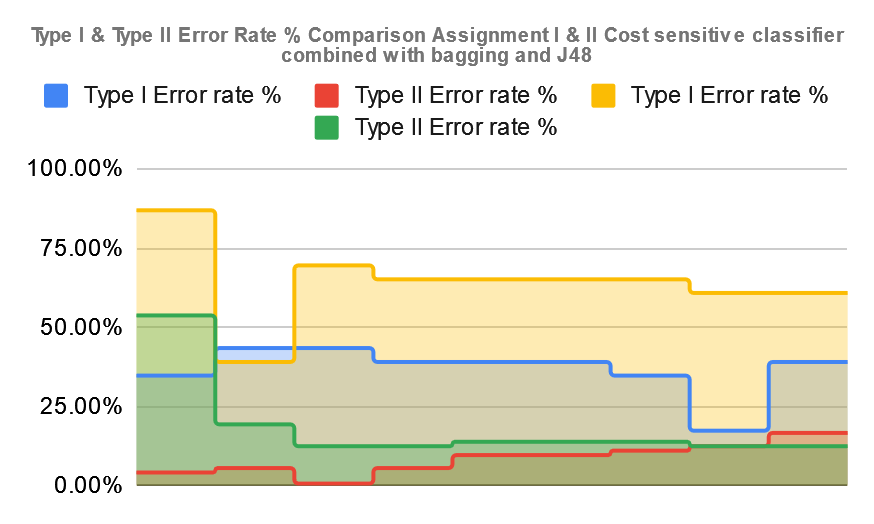
Type I & II Error Rate comparison with assignment study I

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 cost classifier ,bagging, j48** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 3 | 8 | 34.80% | 4.20% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 4 | 10 | 43.50% | 5.60% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 5 | 10 | 43.50% | 0.69% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 4 | 9 | 39.10% | 5.60% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 7 | 9 | 39.10% | 9.70% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 7 | 9 | 39.10% | 9.70% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 8 | 8 | 34.80% | 11.10% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 9 | 4 | 17.40% | 12.50% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 12 | 9 | 39.10% | 16.70% | 9 | 14 | 60.90% | 12.50% |

Table:1-2

In the following comparison Table:1-2, we use a cost ratio of 1:2 ~1:4 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:1-3 below we see the optimal models are produced using ensembles bagging with J48 learners modele perform significantly better. With the bagging and J48 base learners, the data sets performed as significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

*Note: blue and red represent the experiments Type I & Type II errors from cost sensitive classifiers with bagging on the J48 whereas green and yellow represent the Type I & Type II errors from cost sensitive classifiers with only J48 (For all Type I & II error charts for each experiment).*



1. **Cost sensitive classifier combined with bagging and Decision Stump**

The following results in Table:2-1 show the best result for assignment II Type II error at 1:5 ratio at cost ratio variance set to 5. Though the Type II errors are significantly higher at 19. The results show an ROC of 0.866. The lowest misclassification rate is at the 1:1 cost ratio variance of 1 Although at a 1:1 ratio we see less Type I and moreType II errors with the lowest misclassification rate at 16.50%. Overall, the best balanced cost ratio would be found 1:1 ~ 1:2 for bagging with a decision stump.

The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the Decision stump with misclassification rate of 22%. Assignment II parameters show more significance for classification than the comparison.

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2**  **Cost sensitive classifier combined with bagging and Decision Stump** | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 3 | 13 | 0.832 | 0.438 | 0.824 | 16.80% | 0.817 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 4 | 13 | 0.821 | 0.442 | 0.809 | 17.80% | 0.85 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 10 | 9 | 0.8 | 0.33 | 0.803 | 20% | 0.849 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 13 | 10 | 0.758 | 0.373 | 0.769 | 24% | 0.807 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 16 | 8 | 0.747 | 0.317 | 0.78 | 25.60% | 0.852 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 17 | 6 | 0.758 | 0.255 | 0.804 | 24.20% | 0.851 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 19 | 5 | 0.747 | 0.229 | 0.81 | 25.20% | 0.859 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 18 | 6 | 0.747 | 0.258 | 0.8 | 25.2%% | 0.859 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 19 | 3 | 0.768 | 0.163 | 0.841 | 23.10% | 0.866 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:2-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with bagging and Decision Stump learner. Type II error decreases as CSC variation is increased showing better performance.

Chart:2-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with bagging on the Decision Stump learner against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging technique.

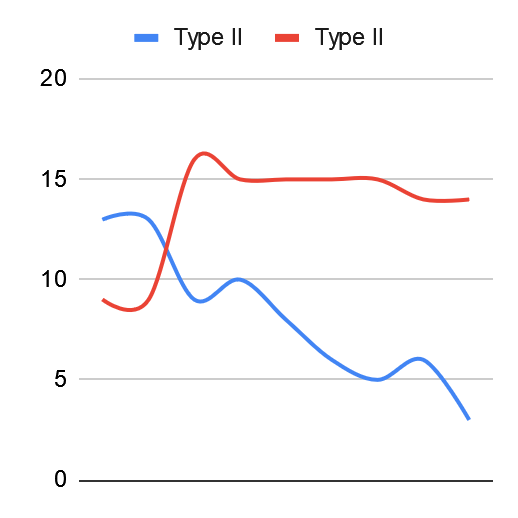
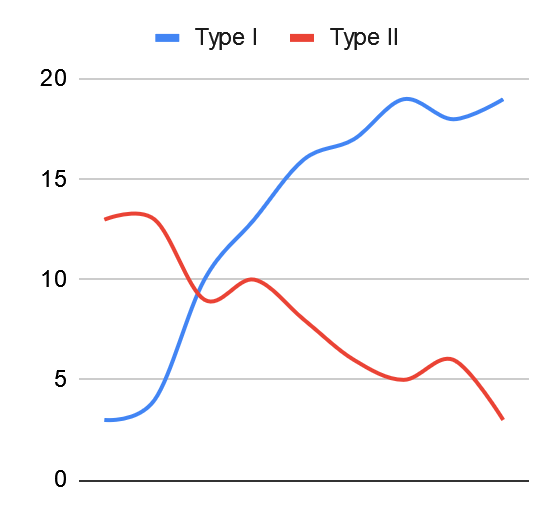


Chart:2-1 Chart:2-2

Type I & II Error Rate comparison with assignment study

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 -Cost sensitive classifier combined with bagging and Decision Stump** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 3 | 13 | 56.50% | 4.20% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 4 | 13 | 56.50% | 5.60% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 10 | 9 | 39.10% | 13.90% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 13 | 10 | 43.50% | 18.10% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 16 | 8 | 34.80% | 22.20% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 17 | 6 | 26.10% | 23.60% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 19 | 5 | 21.70% | 26.40% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 18 | 6 | 26.10% | 25.00% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 19 | 3 | 13.00% | 26.40% | 9 | 14 | 60.90% | 12.50% |

Table:1-2

In the following comparison Table:1-2, we use a cost ratio of 1:1 ~ 1:1.5 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:2-3 below we see the optimal models are produced using ensembles bagging with Decision Stump learners model perform significantly better. With the bagging and J48 base learners, the data sets performed as significantly better than to Decision Stump alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

Chart:2-3 shows significant comparison between assignment I and assignment II experiments with on J48 both Type I and Type II errors were significantly higher.

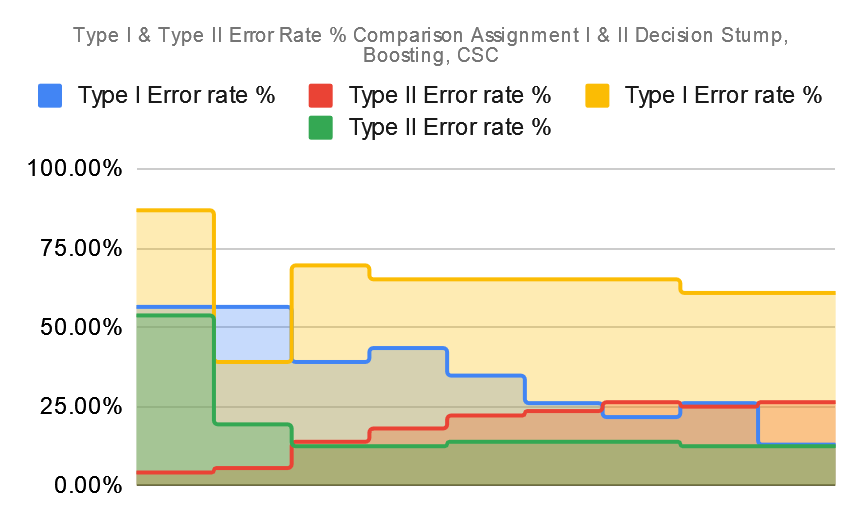


Chart:2-3

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and J48**

The following results in Table:3-1 show the best result for assignment II at cost ratio variance set to 1 with ROC of 0.83. Although at a 1:1 ratio we see less Type I and Type II errors with the lowest misclassification rate at 12.60%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 22.1%. Assignment II parameters show nearly double the significance for classification than the comparison performance by nearly .

Table:3-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 Cost sensitive classifier combined with boosting (AdaBoostM1) and J48** | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 5 | 7 | 0.874 | 0.247 | 0.871 | 12.60% | 0.83 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 4 | 9 | 0.863 | 0.31 | 0.858 | 13.60% | 0.889 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 5 | 12 | 0.821 | 0.412 | 0.809 | 17.80% | 0.74 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 8 | 13 | 0.779 | 0.455 | 0.764 | 22.10% | 0.702 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 6 | 11 | 0.821 | 0.383 | 0.811 | 17.80% | 0.667 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 6 | 11 | 0.821 | 0.383 | 0.811 | 17.80% | 0.727 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 8 | 12 | 0.789 | 0.422 | 0.788 | 21% | 0.762 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 7 | 13 | 0.789 | 0.452 | 0.774 | 21% | 0.711 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 5 | 14 | 0.842 | 0.346 | 0.834 | 16% | 0.818 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:3-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with AdBoostM1 and J48 learner. Type II error decreases as CSC variation is increased showing overall better performance.

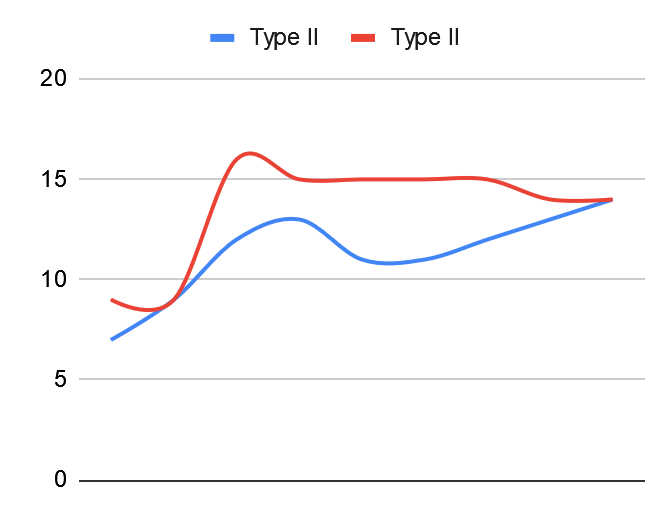
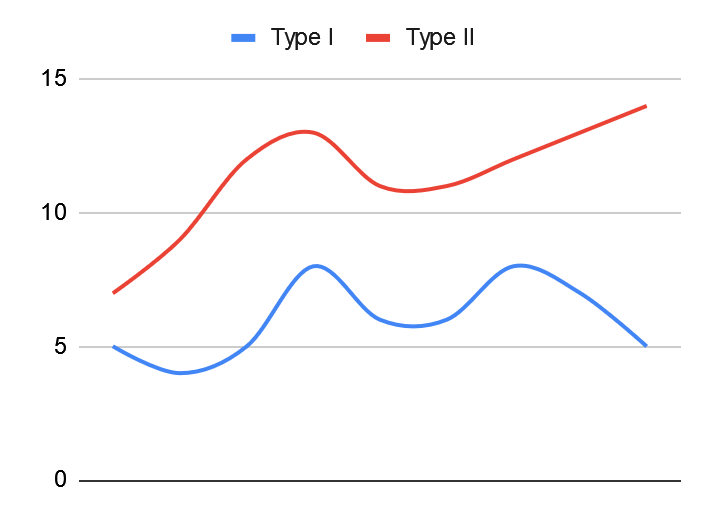
Chart:3-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with AdBoostM1 on the J48 learner against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging technique. Ultimately they converge as the ratio increase.

Chart:3-1 Chart:3-2

Type I & II Error Rate comparison with assignment study I

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 - Cost sensitive classifier combined with boosting (AdaBoostM1) and J48** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 5 | 7 | 30.40% | 6.90% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 4 | 9 | 39.10% | 5.60% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 5 | 12 | 52.20% | 6.90% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 8 | 13 | 56.50% | 11.10% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 6 | 11 | 47.80% | 8.30% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 6 | 11 | 47.80% | 8.30% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 8 | 12 | 52.20% | 11.10% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 7 | 13 | 56.50% | 9.70% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 5 | 14 | 60.90% | 6.90% | 9 | 14 | 60.90% | 12.50% |

Table:3-2

In the following comparison Table:3-2, we use a cost ratio of 1:1.5 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:3-3 below we see the optimal models are produced using ensembles AdBoostM1 with the J48 learners model performing significantly better. With the bagging and J48 base learners, the data sets performed as significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

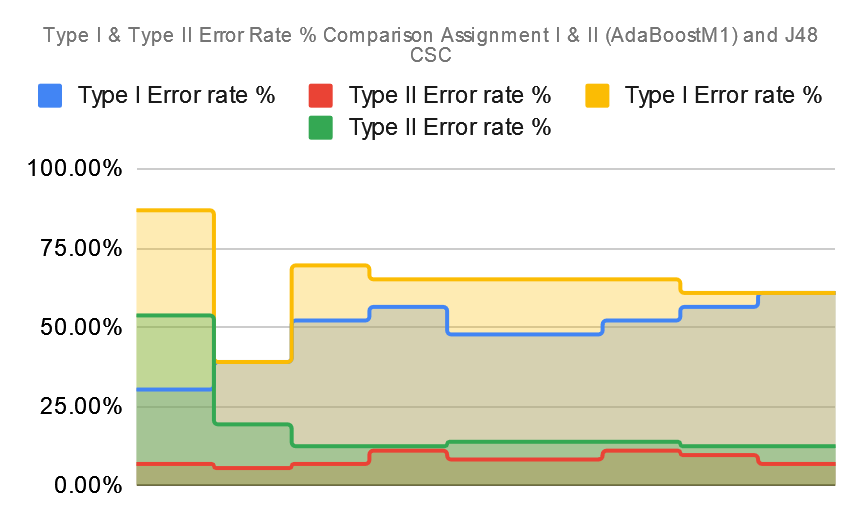


Chart:3-3

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump**

The following results in Table:4-1 show the best result for assignment II at cost ratio variance set to 1 with ROC of 0.87. Although at a 1:1 ~ 1:1.5 ratio we see less Type I and Type II errors with the lowest classification rate at 14.70%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 22%. Assignment II parameters show more significance for classification than the comparison.

Table:4-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump** | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 6 | 8 | 0.853 | 0.284 | 0.849 | 14.70% | 0.87 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 6 | 8 | 0.853 | 0.284 | 0.849 | 14.70% | 0.87 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 8 | 11 | 0.8 | 0.389 | 0.792 | 20 | 0.772 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 7 | 10 | 0.821 | 0.353 | 0.814 | 17.80% | 0.803 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 9 | 11 | 0.789 | 0.393 | 0.784 | 21% | 0.825 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 11 | 10 | 0.779 | 0.367 | 0.782 | 22.10% | 0.81 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 9 | 9 | 0.811 | 0.327 | 0.811 | 18.90% | 0.804 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 10 | 9 | 0.8 | 0.33 | 0.803 | 20% | 0.775 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 12 | 10 | 0.768 | 0.37 | 0.776 | 23.10% | 0.771 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:4-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with AdaBoostM1 and Decision Stump learner. Type II error increases as convergence of the cost ratio increases. Ultimately cost ratio 1:1 ~ 1:1.5 shows best balanced model

Chart:4-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with AdaBoostM1 and Decision Stump learner against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging with J48 alone technique.

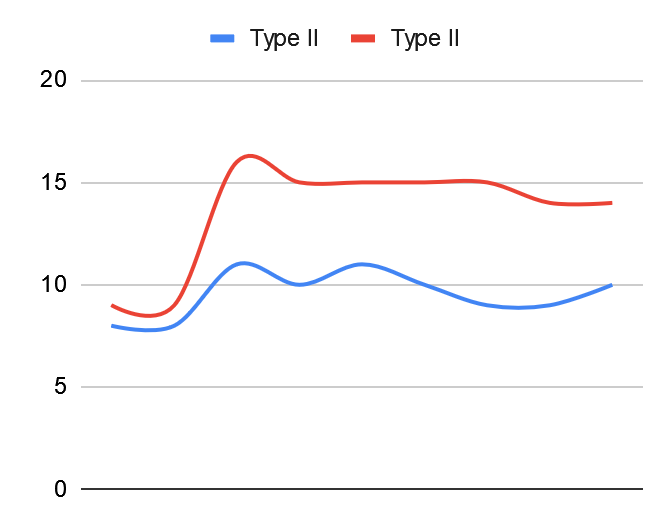
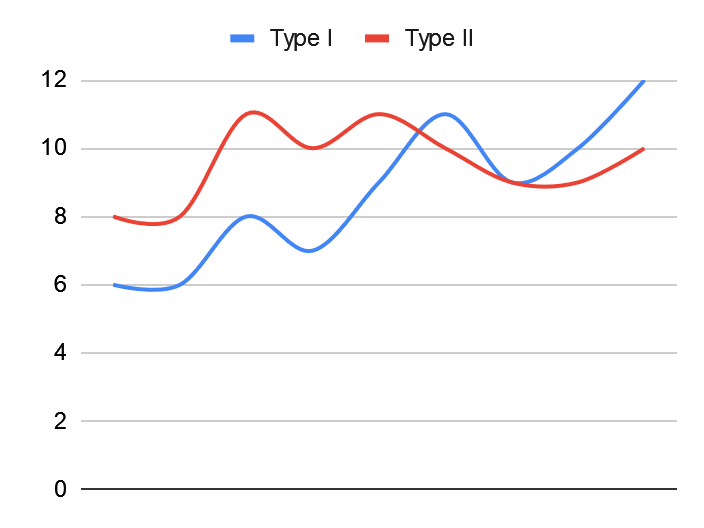
****

Chart:4-1 Chart:4-2

Type I & II Error Rate comparison with assignment study I

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 - Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 6 | 8 | 34.80% | 8.30% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 6 | 8 | 34.80% | 8.30% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 8 | 11 | 47.80% | 11.10% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 7 | 10 | 43.50% | 9.70% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 9 | 11 | 47.80% | 12.50% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 11 | 10 | 43.50% | 15.30% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 9 | 9 | 39.10% | 12.50% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 10 | 9 | 39.10% | 13.90% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 12 | 10 | 43.50% | 16.70% | 9 | 14 | 60.90% | 12.50% |

Table:4-2

In the following comparison Table:4-2, we use a cost ratio of 1:1 ~1:1.5 which proved to be more balanced but only slightly better than the models used with bagging and J48 only. Based on the Type I & Type II error rate percentages in Chart:4-3 below we see the optimal models are produced using ensembles AdaBoost with Decision Stump learners model performing significantly better. With the bagging and J48 base learners, the data sets performed as significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

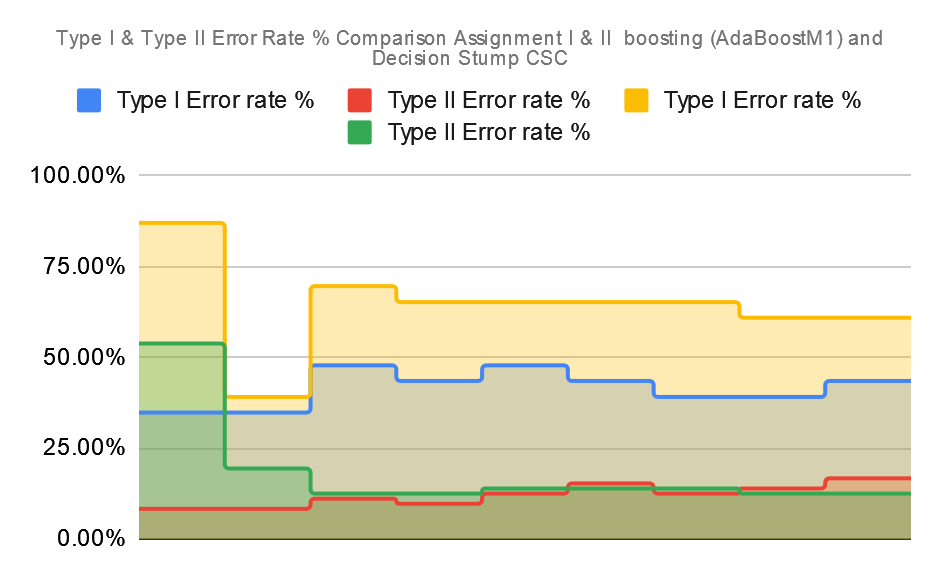
****

Chart:4-3

1. **Cost sensitive classifier combined with bagging and J48 plus meta learner iteration set to 25.**

The following results in Table:5-1 show the best result for assignment II at cost ratio 1:2.5 variance set to 2.5 with ROC of 0.886. Although at a 1:1 ratio we see less Type I and Type II errors with the lowest classification rate at 11.50%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 22%. Assignment II parameters show more significance for classification than the comparison.

Table:5-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 cost classifier ,bagging, j48 (25-iterations)** | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 3 | 10 | 0.863 | 0.34 | 0.859 | 13.60% | 0.92 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 3 | 9 | 0.874 | 0.307 | 0.87 | 12.60% | 0.911 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 5 | 11 | 0.832 | 0.379 | 0.822 | 16.80% | 0.928 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 4 | 7 | 0.884 | 0.244 | 0.881 | 11.50% | 0.886 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 8 | 8 | 0.832 | 0.291 | 0.832 | 16.80% | 0.878 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 8 | 8 | 0.832 | 0.291 | 0.832 | 16.80% | 0.878 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 12 | 12 | 0.747 | 0.436 | 0.747 | 25.20% | 0.87 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 10 | 8 | 0.811 | 0.297 | 0.817 | 18.90% | 0.873 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 11 | 10 | 0.779 | 0.367 | 0.782 | 22.10% | 0.863 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:5-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with bagging and J48 learner with 25-iterations. Type II error decreases as CSC variation is increased showing better performance.

Chart:5-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with bagging on the J48 learners with 25-iterations against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging technique with 25-iterations.

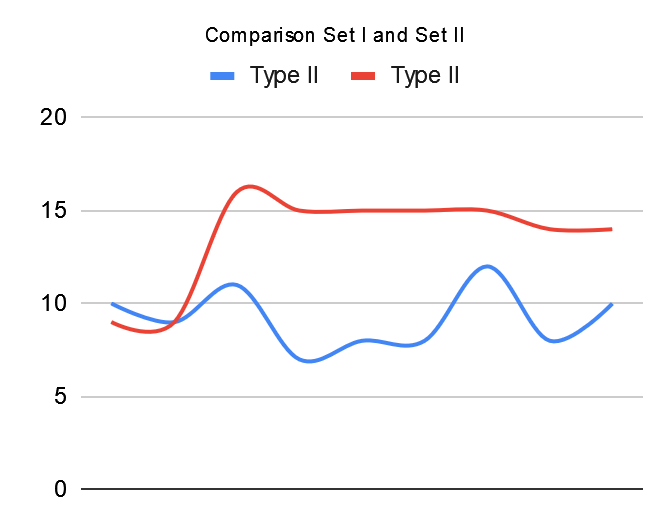
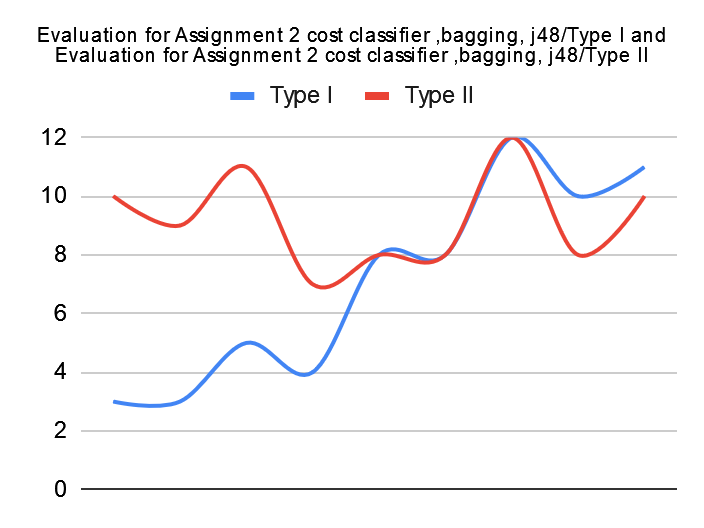


Chart:5-1 Chart:5-2

Type I & II Error Rate comparison with assignment study I

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 cost classifier ,bagging, j48 (25 Iterations)** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 3 | 10 | 43.50% | 4.20% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 3 | 9 | 39.10% | 4.20% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 5 | 11 | 47.80% | 6.90% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 4 | 7 | 30.40% | 5.60% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 8 | 8 | 34.80% | 11.10% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 8 | 8 | 34.80% | 11.10% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 12 | 12 | 52.20% | 16.70% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 10 | 8 | 34.80% | 13.90% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 11 | 10 | 43.50% | 15.30% | 9 | 14 | 60.90% | 12.50% |

Table:5-2

In the following comparison Table:5-2, we use a cost ratio of 1:2.5 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:5-3 below we see the optimal models are produced using ensembles bagging with the J48 learners with 25-iterations model performing significantly better. With the bagging and J48 base learners with 25-iterations, the data sets performed significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

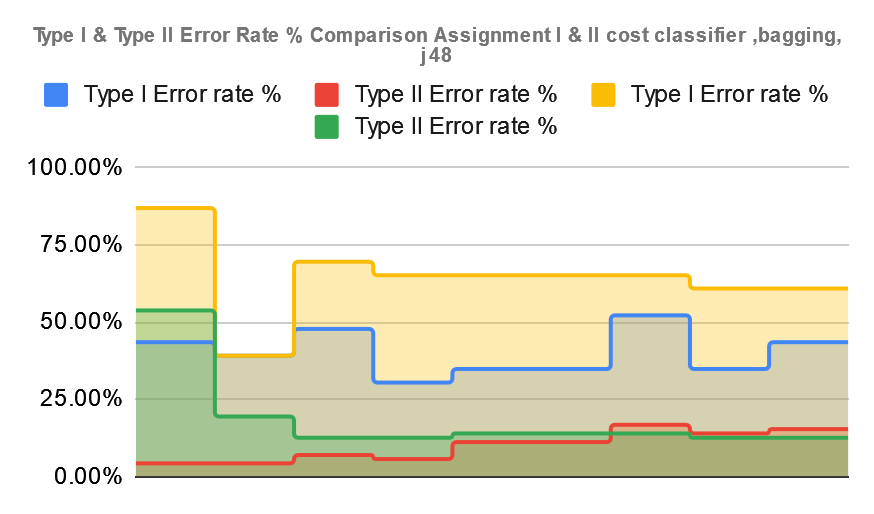


Chart:5-3

1. **Cost sensitive classifier combined with bagging and Decision Stump meta learner iteration set to 25.**

The following results in Table:6-1 show the best result for assignment II at cost ratio 1:3 variance set to 3 with ROC of 0.842. At a 1:3 ratio we see more balanced Type I and Type II errors with the lowest classification error rate at 25%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 25.20%. Assignment II parameters show more significance for classification than the comparison.

Table:1-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 Cost sensitive classifier combined with bagging and Decision Stump (25-iterations)** | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 2 | 15 | 0.821 | 0.501 | 0.818 | 17.80% | 0.825 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 2 | 12 | 0.853 | 0.402 | 0.852 | 14.70% | 0.867 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 8 | 9 | 0.821 | 0.323 | 0.819 | 17.80% | 0.839 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 15 | 9 | 0.747 | 0.347 | 0.771 | 25.20% | 0.828 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 18 | 6 | 0.747 | 0.258 | 0.8 | 25.20% | 0.842 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 22 | 7 | 0.659 | 0.305 | 0.767 | 30.52% | 0.838 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 19 | 7 | 0.726 | 0.295 | 0.78 | 27.30% | 0.85 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 23 | 6 | 0.695 | 0.275 | 0.778 | 30.50% | 0.838 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 23 | 6 | 0.695 | 0.275 | 0.778 | 30.50% | 0.838 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:6-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with bagging and Decisionlearner with 25-iterations. Type II error decreases as CSC variation is increased showing better performance.

Chart:6-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with bagging on the J48 learners with 25-iterations against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging technique with 25-iterations.

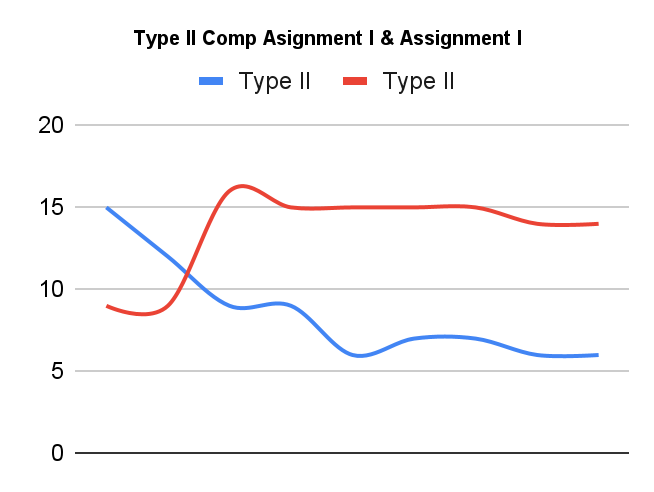
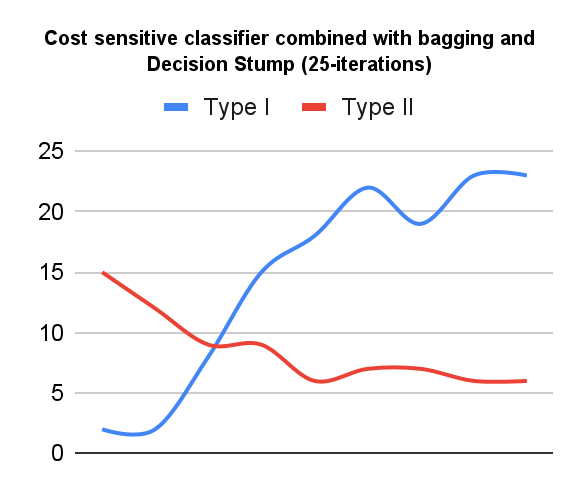
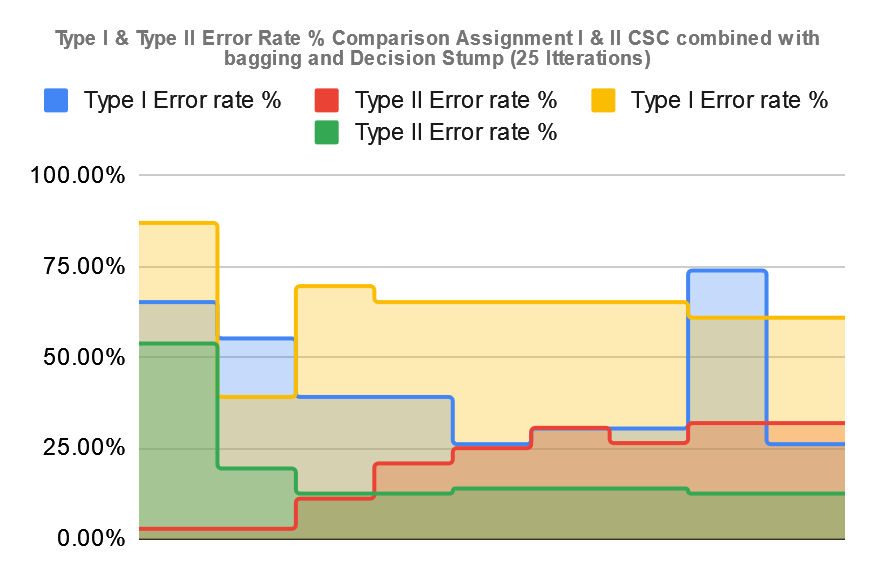


Chart:6-1 Chart:6-2

Type I & II Error Rate comparison with assignment study I

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 - Cost sensitive classifier combined with bagging and Decision Stump (25 Iterations)** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 2 | 15 | 65.20% | 2.80% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 2 | 12 | 55.20% | 2.80% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 8 | 9 | 39.10% | 11.10% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 15 | 9 | 39.10% | 20.80% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 18 | 6 | 26.10% | 25.00% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 22 | 7 | 30.40% | 30.60% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 19 | 7 | 30.40% | 26.40% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 23 | 6 | 73.90% | 31.90% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 23 | 6 | 26.10% | 31.90% | 9 | 14 | 60.90% | 12.50% |

In the following comparison Table:6-2, we use a cost ratio of 1:3 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:6-3 below we see the optimal models are produced using ensembles bagging with the J48 learners with 25-iterations model performing significantly better. With the bagging and J48 base learners with 25-iterations, the data sets performed significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.



1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and J48 meta learner iteration set to 25.**

The following results in Table:7-1 show the best result for assignment II at cost ratio 1:1 ~ 1:1.5 variance set to 1.5 with ROC of 0.909 which is significantly higher than all the experiments this far. Although at a 1:1 ~ 1:1.5 ratio we still see less Type I and Type II errors with the lowest average misclassification rate at 12.60%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 22%. Assignment II parameters show more significance for classification than the comparison.

Table:7-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2** - Cost sensitive classifier combined with boosting (AdaBoostM1) and J48 - (25 iterations) | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 6 | 8 | 0.853 | 0.284 | 0.849 | 14.70% | 0.852 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 4 | 8 | 0.874 | 0.277 | 0.869 | 12.60% | 0.909 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 3 | 10 | 0.863 | 0.34 | 0.859 | 13.60% | 0.797 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 6 | 12 | 0.811 | 0.416 | 0.798 | 18.90% | 0.73 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 6 | 11 | 0.821 | 0.383 | 0.811 | 17.80% | 0.715 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 3 | 11 | 0.853 | 0.373 | 0.847 | 14.70% | 0.795 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 7 | 8 | 0.842 | 0.287 | 0.84 | 15.70% | 0.793 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 6 | 13 | 0.8 | 0.449 | 0.784 | 20% | 0.755 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 5 | 12 | 0.821 | 0.412 | 0.809 | 17.80% | 0.746 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:7-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with AdaBoostM1 and J48 learner with 25-iterations. Type II error decreases as CSC variation is increased showing better performance.

Chart:7-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC with AdaBoostM1 on the J48 learners with 25-iterations against the CSC with only the J48 as resulted in assignment I.Type II errors were significantly less using the bagging technique with 25-iterations.

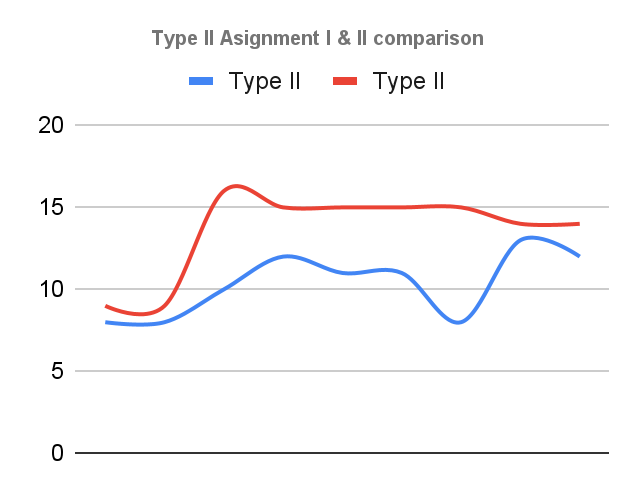
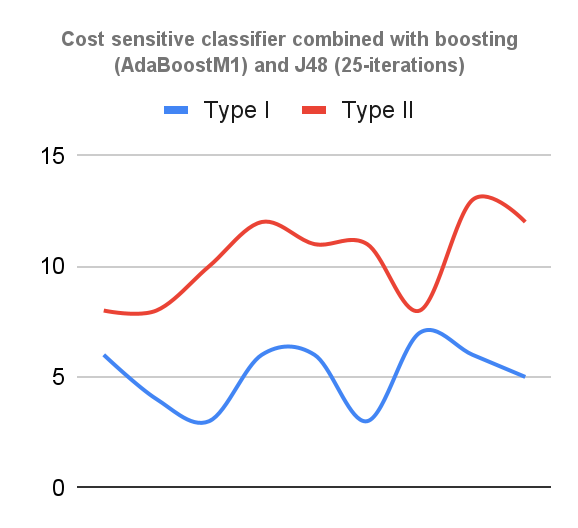


Chart:7-1 Chart:7-2

Type I & II Error Rate comparison with assignment study I

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2 cost classifier ,bagging, j48 25-iterations** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 6 | 8 | 34.80% | 8.30% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 4 | 8 | 43.50% | 4.20% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 3 | 10 | 34.80% | 5.60% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 6 | 12 | 52.20% | 8.30% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 6 | 11 | 47.80% | 8.30% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 3 | 11 | 47.80% | 4.20% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 7 | 8 | 34.80% | 9.70% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 6 | 13 | 56.50% | 8.30% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 5 | 12 | 52.20% | 6.90% | 9 | 14 | 60.90% | 12.50% |

In the following comparison Table:7-2, we use a cost ratio of 1:1.5 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:7-3 below we see the optimal models are produced using ensembles AdaBoostM1 with the J48 learners with 25-iterations model performing significantly better. With the AdaBoostM1 and J48 base learners with 25-iterations, the data sets performed significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

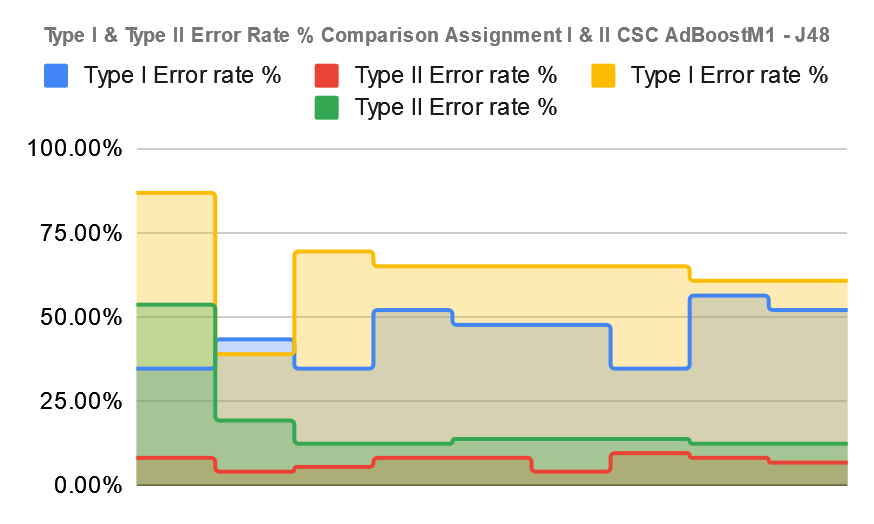


Chart:7-3

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump meta learner iteration set to 25.**

The following results in Table:8-1 show the best result for assignment II at cost ratio 1:1 variance set to 1 penalty with overall best ROC of 94%. Although at a 1:1 ratio we see less Type I and Type II errors with the lowest average misclassification rate at 10.50%. The comparison shows the cost classifier variance of 1 for assignment I with cost sensitive classifier and only the J48 with misclassification rate of 22%. Assignment II parameters show more significance for classification than the comparison.

Table:1-1

| **Cost Sensitive Classifier Adj.** | | **Evaluation for Assignment 2** - Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump - (25 iterations) | | | | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC | Type I | Type II | TP Rate | FP Rate | Precision | Misclassification Rate % | ROC |
| 1 | 1 | 5 | 5 | 0.895 | 0.182 | 0.895 | 10.50% | 0.94 | 12 | 9 | 0.779 | 0.337 | 0.789 | 22.10% | 0.732 |
| 1 | 1.5 | 3 | 11 | 0.853 | 0.373 | 0.847 | 14.70% | 0.876 | 14 | 9 | 0.758 | 0.344 | 0.758 | 24.20% | 0.681 |
| 1 | 2 | 2 | 9 | 0.884 | 0.303 | 0.883 | 11.50% | 0.855 | 9 | 16 | 0.737 | 0.557 | 0.71 | 26.30% | 0.591 |
| 1 | 2.5 | 2 | 10 | 0.874 | 0.336 | 0.873 | 12.60% | 0.886 | 9 | 15 | 0.747 | 0.525 | 0.726 | 25.20% | 0.616 |
| 1 | 3 | 7 | 11 | 0.811 | 0.386 | 0.801 | 18.90% | 0.882 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 3.5 | 9 | 10 | 0.8 | 0.36 | 0.797 | 20% | 0.889 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4 | 6 | 9 | 0.842 | 0.317 | 0.836 | 15.70% | 0.882 | 10 | 15 | 0.737 | 0.528 | 0.718 | 26.30% | 0.61 |
| 1 | 4.5 | 4 | 8 | 0.874 | 0.277 | 0.869 | 12.60% | 0.854 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |
| 1 | 5 | 6 | 11 | 0.821 | 0.383 | 0.811 | 17.80% | 0.838 | 9 | 14 | 0.758 | 0.492 | 0.741 | 24% | 0.635 |

Chart:8-1 shows the comparison of type I and type II errors evaluated for cost sensitive classifier(CSC) with AdaBoostM1 and Decision Stump learner with 25-iterations. Type II error decreases as CSC variation is increased showing better performance.

Chart:8-2 Shows comparison of both assignments I & II Type II errors. The blue-line indicates The CSC withAdaBoostM1 and Decision Stump with 25-iterations against the CSC with only the J48 as resulted in assignment I. Type II errors were significantly less using the bagging technique with 25-iterations.

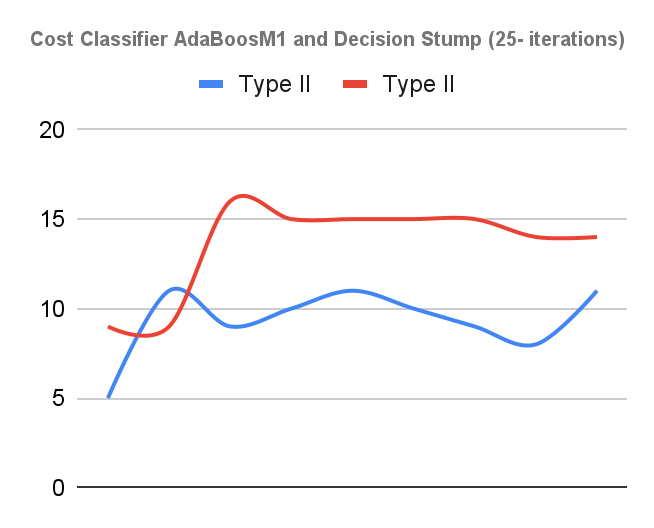
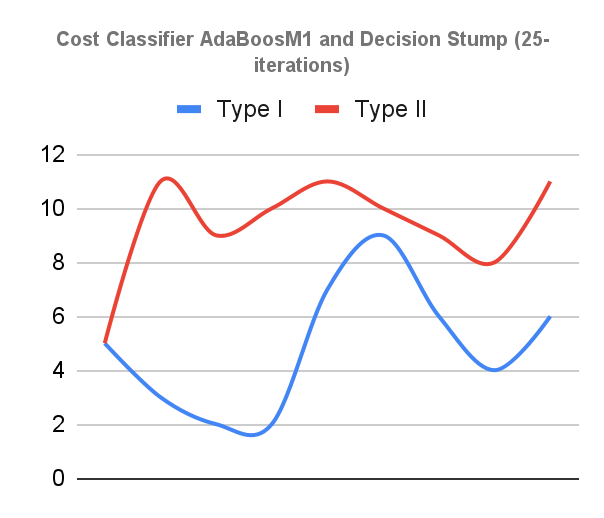


Chart:8-1 Chart:8-2

Type I & II Error Rate comparison with assignment study I

Table:8-2

|  |  | **Evaluation for Assignment 2 - Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump**  **- (25 Iterations)** | | | | **Evaluation for Assignment 1 - Cost Classifier with j48** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cost of Type I Error | Cost of Type II Error | Type I | Type II | Type I Error rate % | Type II Error rate % | Type I | Type II | Type I Error rate % | Type II Error rate % |
| 1 | 1 | 5 | 5 | 21.70% | 6.90% | 12 | 9 | 87.00% | 53.80% |
| 1 | 1.5 | 3 | 11 | 42.80% | 4.20% | 14 | 9 | 39.10% | 19.40% |
| 1 | 2 | 2 | 9 | 43.50% | 2.80% | 9 | 16 | 69.60% | 12.50% |
| 1 | 2.5 | 2 | 10 | 47.80% | 9.70% | 9 | 15 | 65.20% | 12.50% |
| 1 | 3 | 7 | 11 | 43.50% | 12.50% | 10 | 15 | 65.20% | 13.90% |
| 1 | 3.5 | 9 | 10 | 39.10% | 8.30% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4 | 6 | 9 | 34.80% | 5.60% | 10 | 15 | 65.20% | 13.90% |
| 1 | 4.5 | 4 | 8 | 47.80% | 8.30% | 9 | 14 | 60.90% | 12.50% |
| 1 | 5 | 6 | 11 | 43.50% | 15.30% | 9 | 14 | 60.90% | 12.50% |

In the following comparison Table:8-2, we use a cost ratio of 1:1 which proved to be more balanced. Based on the Type I & Type II error rate percentages in Chart:2-3 below we see the optimal models are produced using ensembles AdaBoostM1 and Decision Stump learners with 25-iterations models performing significantly better. With the AdaBoostM1 and Decision Stump base learners with 25-iterations, the data sets performed significantly better than to J48 alone. Type I errors are higher, but type II errors and total misclassification rates are significantly lower.

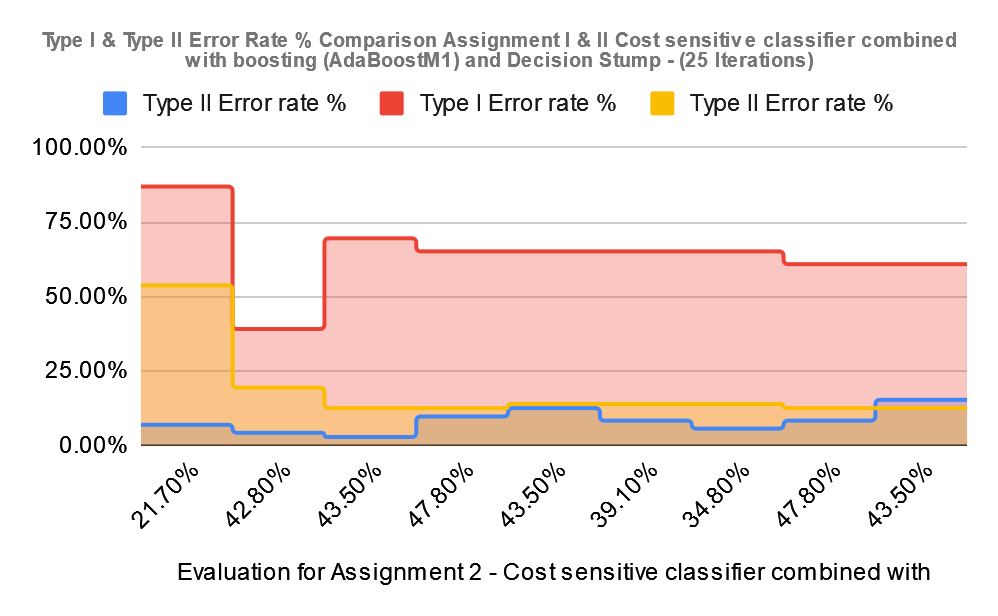


Chart:8-3

Overall the AdaBoostM1 and Decision Stump with 25 iterations performed the overall best and produced the most balanced classification models.

**Appendices**

Appendix of weka data analysis based on assignment I and assignment II results and instances.

1. Part 4 results Assignment 1 cost sensitive classifier combined with J48
   1. Cost set to 1
   2. Cost set to 2
   3. Cost set to 2.5
   4. Cost set to 3
   5. Cost set to 3.5
   6. Cost set to 4
   7. Cost set to 4.5
   8. Cost set to 5
2. Cost sensitive classifier combined with bagging and J48
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
3. Cost sensitive classifier combined with bagging and Decision Stump
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
4. Cost sensitive classifier combined with boosting (AdaBoostM1) and J48
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
5. Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
6. Cost sensitive classifier combined with bagging and J48 plus meta learner iteration set to 25.
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
7. Cost sensitive classifier combined with bagging and Decision Stump plus meta learner iteration set to 25
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
8. Cost sensitive classifier combined with boosting (AdaBoostM1) and J48 plus meta learner iteration set to 25
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
9. Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump plus meta learner iteration set to 25
   1. Cost set to 0.5
   2. Cost set to 1
   3. Cost set to 2
   4. Cost set to 2.5
10. **Part 4 results Assignment 1 cost sensitive classifier combined with J48** 
    1. **Cost set to 1**

Classifier Model

J48 pruned tree

------------------

GENE3941X <= 0.94

| GENE1125X <= 1.01

| | GENE1567X <= 2.62

| | | GENE1391X <= 0.57: nonACL (66.9)

| | | GENE1391X > 0.57

| | | | GENE3732X <= -0.74: nonACL (3.0)

| | | | GENE3732X > -0.74: ACL (2.0)

| | GENE1567X > 2.62

| | | GENE2996X <= 0: nonACL (2.1)

| | | GENE2996X > 0: ACL (6.0)

| GENE1125X > 1.01: ACL (5.0)

GENE3941X > 0.94: ACL (10.0)

Number of Leaves : 7

Size of the tree : 13

Cost Matrix

0 1

1 0

Time taken to build model: 0.13 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 74 77.8947 %

Incorrectly Classified Instances 21 22.1053 %

Kappa statistic 0.4232

Mean absolute error 0.2178

Root mean squared error 0.4496

Relative absolute error 58.8069 %

Root relative squared error 104.8305 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.167 0.538 0.609 0.571 0.425 0.736 0.562 ACL

0.833 0.391 0.870 0.833 0.851 0.425 0.731 0.853 nonACL

Weighted Avg. 0.779 0.337 0.789 0.779 0.783 0.425 0.732 0.782

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

12 60 | b = nonACL

* 1. **Cost set to 1.5**

Classifier Model

J48 pruned tree

------------------

GENE3941X <= 0.94

| GENE1567X <= 2.62

| | GENE1131X <= 0.7: nonACL (61.84/1.34)

| | GENE1131X > 0.7: ACL (5.8/1.78)

| GENE1567X > 2.62: ACL (13.98/1.94)

GENE3941X > 0.94: ACL (13.38)

Number of Leaves : 4

Size of the tree : 7

Cost Matrix

0 1.5

1 0

Time taken to build model: 0.16 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 72 75.7895 %

Incorrectly Classified Instances 23 24.2105 %

Kappa statistic 0.3857

Mean absolute error 0.2499

Root mean squared error 0.4763

Relative absolute error 67.4837 %

Root relative squared error 111.0661 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.194 0.500 0.609 0.549 0.389 0.679 0.486 ACL

0.806 0.391 0.866 0.806 0.835 0.389 0.682 0.832 nonACL

Weighted Avg. 0.758 0.344 0.777 0.758 0.765 0.389 0.681 0.748

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

14 58 | b = nonACL

* 1. **Cost set to 2**

Classifier Model

J48 pruned tree

------------------

GENE3941X <= 0.94

| GENE1567X <= 2.62

| | GENE1126X <= 1.27

| | | GENE212X <= 0.63: nonACL (55.27)

| | | GENE212X > 0.63: ACL (4.14/0.92)

| | GENE1126X > 1.27: ACL (3.22)

| GENE1567X > 2.62: ACL (16.27/1.78)

GENE3941X > 0.94: ACL (16.1)

Number of Leaves : 5

Size of the tree : 9

Cost Matrix

0 2

1 0

Time taken to build model: 0.16 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 70 73.6842 %

Incorrectly Classified Instances 25 26.3158 %

Kappa statistic 0.2001

Mean absolute error 0.2582

Root mean squared error 0.4954

Relative absolute error 69.7095 %

Root relative squared error 115.5121 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.304 0.125 0.438 0.304 0.359 0.205 0.591 0.317 ACL

0.875 0.696 0.797 0.875 0.834 0.205 0.591 0.795 nonACL

Weighted Avg. 0.737 0.557 0.710 0.737 0.719 0.205 0.591 0.680

=== Confusion Matrix ===

a b <-- classified as

7 16 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 2.5**

Classifier Model

J48 pruned tree

------------------

GENE1610X <= -0.77: nonACL (30.08)

GENE1610X > -0.77

| GENE3781X <= -0.77: nonACL (11.74)

| GENE3781X > -0.77

| | GENE3332X <= 2.16

| | | GENE1879X <= 0.36: ACL (43.65/1.47)

| | | GENE1879X > 0.36: nonACL (2.84)

| | GENE3332X > 2.16: nonACL (6.69)

Number of Leaves : 5

Size of the tree : 9

Cost Matrix

0 2.5

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.2445

Mean absolute error 0.2482

Root mean squared error 0.4934

Relative absolute error 67.0348 %

Root relative squared error 115.0524 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.348 0.125 0.471 0.348 0.400 0.249 0.616 0.347 ACL

0.875 0.652 0.808 0.875 0.840 0.249 0.616 0.804 nonACL

Weighted Avg. 0.747 0.525 0.726 0.747 0.733 0.249 0.616 0.694

=== Confusion Matrix ===

a b <-- classified as

8 15 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 3**

Classifier Model

J48 pruned tree

------------------

GENE1610X <= -0.77: nonACL (27.62)

GENE1610X > -0.77

| GENE3781X <= -0.77: nonACL (10.78)

| GENE3781X > -0.77

| | GENE3332X <= 2.16

| | | GENE1879X <= 0.36: ACL (47.84/1.35)

| | | GENE1879X > 0.36: nonACL (2.62)

| | GENE3332X > 2.16: nonACL (6.14)

Number of Leaves : 5

Size of the tree : 9

Cost Matrix

0 3

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 70 73.6842 %

Incorrectly Classified Instances 25 26.3158 %

Kappa statistic 0.2256

Mean absolute error 0.2593

Root mean squared error 0.505

Relative absolute error 70.0206 %

Root relative squared error 117.7553 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.348 0.139 0.444 0.348 0.390 0.228 0.609 0.332 ACL

0.861 0.652 0.805 0.861 0.832 0.228 0.609 0.802 nonACL

Weighted Avg. 0.737 0.528 0.718 0.737 0.725 0.228 0.609 0.688

=== Confusion Matrix ===

a b <-- classified as

8 15 | a = ACL

10 62 | b = nonACL

* 1. **Cost set to 3.5**

Classifier Model

J48 pruned tree

------------------

GENE1610X <= -0.77: nonACL (25.54)

GENE1610X > -0.77

| GENE3781X <= -0.77: nonACL (9.97)

| GENE3781X > -0.77

| | GENE3328X <= 0.87

| | | GENE1173X <= 1.15: ACL (48.59/0.62)

| | | GENE1173X > 1.15

| | | | GENE2543X <= -0.7: ACL (2.18)

| | | | GENE2543X > -0.7: nonACL (3.74)

| | GENE3328X > 0.87: nonACL (4.98)

Number of Leaves : 6

Size of the tree : 11

Cost Matrix

0 3.5

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 70 73.6842 %

Incorrectly Classified Instances 25 26.3158 %

Kappa statistic 0.2256

Mean absolute error 0.2597

Root mean squared error 0.5026

Relative absolute error 70.1321 %

Root relative squared error 117.195 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.348 0.139 0.444 0.348 0.390 0.228 0.610 0.336 ACL

0.861 0.652 0.805 0.861 0.832 0.228 0.610 0.803 nonACL

Weighted Avg. 0.737 0.528 0.718 0.737 0.725 0.228 0.610 0.690

=== Confusion Matrix ===

a b <-- classified as

8 15 | a = ACL

10 62 | b = nonACL

* 1. **Cost set to 4**

Classifier Model

J48 pruned tree

------------------

GENE1610X <= -0.77: nonACL (23.75)

GENE1610X > -0.77

| GENE3781X <= -0.77: nonACL (9.27)

| GENE3781X > -0.77

| | GENE3328X <= 0.87

| | | GENE1173X <= 1.15: ACL (51.55/0.58)

| | | GENE1173X > 1.15

| | | | GENE2543X <= -0.7: ACL (2.32)

| | | | GENE2543X > -0.7: nonACL (3.48)

| | GENE3328X > 0.87: nonACL (4.63)

Number of Leaves : 6

Size of the tree : 11

Cost Matrix

0 4

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 70 73.6842 %

Incorrectly Classified Instances 25 26.3158 %

Kappa statistic 0.2256

Mean absolute error 0.26

Root mean squared error 0.5036

Relative absolute error 70.2213 %

Root relative squared error 117.4372 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.348 0.139 0.444 0.348 0.390 0.228 0.610 0.339 ACL

0.861 0.652 0.805 0.861 0.832 0.228 0.610 0.803 nonACL

Weighted Avg. 0.737 0.528 0.718 0.737 0.725 0.228 0.610 0.690

=== Confusion Matrix ===

a b <-- classified as

8 15 | a = ACL

10 62 | b = nonACL

* 1. **Cost set to 4.5**

Classifier Model

J48 pruned tree

------------------

GENE1610X <= -0.77: nonACL (22.19)

GENE1610X > -0.77

| GENE3781X <= -0.77: nonACL (8.66)

| GENE3781X > -0.77

| | GENE3328X <= 0.87

| | | GENE1173X <= 1.15: ACL (54.13/0.54)

| | | GENE1173X > 1.15

| | | | GENE2543X <= -0.7: ACL (2.44)

| | | | GENE2543X > -0.7: nonACL (3.25)

| | GENE3328X > 0.87: nonACL (4.33)

Number of Leaves : 6

Size of the tree : 11

Cost Matrix

0 4.5

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 72 75.7895 %

Incorrectly Classified Instances 23 24.2105 %

Kappa statistic 0.2876

Mean absolute error 0.2414

Root mean squared error 0.4867

Relative absolute error 65.1847 %

Root relative squared error 113.4836 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.391 0.125 0.500 0.391 0.439 0.291 0.638 0.362 ACL

0.875 0.609 0.818 0.875 0.846 0.291 0.638 0.815 nonACL

Weighted Avg. 0.758 0.492 0.741 0.758 0.747 0.291 0.638 0.705

=== Confusion Matrix ===

a b <-- classified as

9 14 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 5**

Classifier Model

J48 pruned tree

------------------

GENE1610X <= -0.77: nonACL (20.83)

GENE1610X > -0.77

| GENE3781X <= -0.77: nonACL (8.13)

| GENE3781X > -0.77

| | GENE3328X <= 0.87

| | | GENE1173X <= 1.15: ACL (56.39/0.51)

| | | GENE1173X > 1.15

| | | | GENE2543X <= -0.7: ACL (2.54)

| | | | GENE2543X > -0.7: nonACL (3.05)

| | GENE3328X > 0.87: nonACL (4.06)

Number of Leaves : 6

Size of the tree : 11

Cost Matrix

0 5

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 72 75.7895 %

Incorrectly Classified Instances 23 24.2105 %

Kappa statistic 0.2876

Mean absolute error 0.2415

Root mean squared error 0.4872

Relative absolute error 65.2018 %

Root relative squared error 113.6003 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.391 0.125 0.500 0.391 0.439 0.291 0.638 0.362 ACL

0.875 0.609 0.818 0.875 0.846 0.291 0.634 0.813 nonACL

Weighted Avg. 0.758 0.492 0.741 0.758 0.747 0.291 0.635 0.703

=== Confusion Matrix ===

a b <-- classified as

9 14 | a = ACL

9 63 | b = nonACL

1. **Cost sensitive classifier combined with bagging and J48**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 0.8 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 84 88.4211 %

Incorrectly Classified Instances 11 11.5789 %

Kappa statistic 0.6593

Mean absolute error 0.2399

Root mean squared error 0.3133

Relative absolute error 64.7884 %

Root relative squared error 73.0507 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.042 0.833 0.652 0.732 0.667 0.917 0.842 ACL

0.958 0.348 0.896 0.958 0.926 0.667 0.917 0.960 nonACL

Weighted Avg. 0.884 0.274 0.881 0.884 0.879 0.667 0.917 0.932

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 0.74 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5593

Mean absolute error 0.2376

Root mean squared error 0.332

Relative absolute error 64.1583 %

Root relative squared error 77.421 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.056 0.765 0.565 0.650 0.570 0.876 0.773 ACL

0.944 0.435 0.872 0.944 0.907 0.570 0.876 0.926 nonACL

Weighted Avg. 0.853 0.343 0.846 0.853 0.845 0.570 0.876 0.889

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 0.7 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 80 84.2105 %

Incorrectly Classified Instances 15 15.7895 %

Kappa statistic 0.5354

Mean absolute error 0.2583

Root mean squared error 0.3365

Relative absolute error 69.7525 %

Root relative squared error 78.4697 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.069 0.722 0.565 0.634 0.542 0.875 0.782 ACL

0.931 0.435 0.870 0.931 0.899 0.542 0.875 0.937 nonACL

Weighted Avg. 0.842 0.346 0.834 0.842 0.835 0.542 0.875 0.899

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

5 67 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 0.61 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 82 86.3158 %

Incorrectly Classified Instances 13 13.6842 %

Kappa statistic 0.5973

Mean absolute error 0.2467

Root mean squared error 0.3277

Relative absolute error 66.6135 %

Root relative squared error 76.4094 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.056 0.778 0.609 0.683 0.605 0.883 0.803 ACL

0.944 0.391 0.883 0.944 0.913 0.605 0.883 0.953 nonACL

Weighted Avg. 0.863 0.310 0.858 0.863 0.857 0.605 0.883 0.917

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 0.63 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.5271

Mean absolute error 0.2692

Root mean squared error 0.3528

Relative absolute error 72.6953 %

Root relative squared error 82.2739 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.097 0.667 0.609 0.636 0.528 0.866 0.700 ACL

0.903 0.391 0.878 0.903 0.890 0.528 0.866 0.923 nonACL

Weighted Avg. 0.832 0.320 0.827 0.832 0.829 0.528 0.866 0.869

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

7 65 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 0.62 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.5271

Mean absolute error 0.2831

Root mean squared error 0.3647

Relative absolute error 76.4552 %

Root relative squared error 85.0469 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.097 0.667 0.609 0.636 0.528 0.852 0.716 ACL

0.903 0.391 0.878 0.903 0.890 0.528 0.852 0.944 nonACL

Weighted Avg. 0.832 0.320 0.827 0.832 0.829 0.528 0.852 0.889

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

7 65 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 0.63 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.5411

Mean absolute error 0.2881

Root mean squared error 0.3742

Relative absolute error 77.8024 %

Root relative squared error 87.2478 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.111 0.652 0.652 0.652 0.541 0.848 0.646 ACL

0.889 0.348 0.889 0.889 0.889 0.541 0.848 0.945 nonACL

Weighted Avg. 0.832 0.291 0.832 0.832 0.832 0.541 0.848 0.873

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 0.61 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 82 86.3158 %

Incorrectly Classified Instances 13 13.6842 %

Kappa statistic 0.6528

Mean absolute error 0.2934

Root mean squared error 0.3645

Relative absolute error 79.2142 %

Root relative squared error 85.0012 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.826 0.125 0.679 0.826 0.745 0.659 0.877 0.709 ACL

0.875 0.174 0.940 0.875 0.906 0.659 0.877 0.926 nonACL

Weighted Avg. 0.863 0.162 0.877 0.863 0.867 0.659 0.877 0.873

=== Confusion Matrix ===

a b <-- classified as

19 4 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 5

1 0

Time taken to build model: 0.6 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 74 77.8947 %

Incorrectly Classified Instances 21 22.1053 %

Kappa statistic 0.4232

Mean absolute error 0.3049

Root mean squared error 0.3852

Relative absolute error 82.3353 %

Root relative squared error 89.8248 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.167 0.538 0.609 0.571 0.425 0.828 0.634 ACL

0.833 0.391 0.870 0.833 0.851 0.425 0.828 0.938 nonACL

Weighted Avg. 0.779 0.337 0.789 0.779 0.783 0.425 0.828 0.864

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

12 60 | b = nonACL

1. **Cost sensitive classifier combined with bagging and Decision Stump**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 0.3 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.4614

Mean absolute error 0.2737

Root mean squared error 0.3617

Relative absolute error 73.9169 %

Root relative squared error 84.3486 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.435 0.042 0.769 0.435 0.556 0.490 0.817 0.712 ACL

0.958 0.565 0.841 0.958 0.896 0.490 0.817 0.926 nonACL

Weighted Avg. 0.832 0.438 0.824 0.832 0.814 0.490 0.817 0.874

=== Confusion Matrix ===

a b <-- classified as

10 13 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 0.31 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.4375

Mean absolute error 0.2886

Root mean squared error 0.3596

Relative absolute error 77.9282 %

Root relative squared error 83.8506 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.435 0.056 0.714 0.435 0.541 0.458 0.850 0.690 ACL

0.944 0.565 0.840 0.944 0.889 0.458 0.850 0.946 nonACL

Weighted Avg. 0.821 0.442 0.809 0.821 0.805 0.458 0.850 0.884

=== Confusion Matrix ===

a b <-- classified as

10 13 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 0.3 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 76 80 %

Incorrectly Classified Instances 19 20 %

Kappa statistic 0.463

Mean absolute error 0.3162

Root mean squared error 0.3704

Relative absolute error 85.3793 %

Root relative squared error 86.3664 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.139 0.583 0.609 0.596 0.463 0.849 0.738 ACL

0.861 0.391 0.873 0.861 0.867 0.463 0.849 0.940 nonACL

Weighted Avg. 0.800 0.330 0.803 0.800 0.801 0.463 0.849 0.891

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

10 62 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 0.29 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 72 75.7895 %

Incorrectly Classified Instances 23 24.2105 %

Kappa statistic 0.3683

Mean absolute error 0.3415

Root mean squared error 0.4036

Relative absolute error 92.2027 %

Root relative squared error 94.1048 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.181 0.500 0.565 0.531 0.370 0.807 0.557 ACL

0.819 0.435 0.855 0.819 0.837 0.370 0.807 0.930 nonACL

Weighted Avg. 0.758 0.373 0.769 0.758 0.763 0.370 0.807 0.840

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

13 59 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 0.3 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.3844

Mean absolute error 0.3376

Root mean squared error 0.3945

Relative absolute error 91.1523 %

Root relative squared error 91.9956 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.222 0.484 0.652 0.556 0.393 0.852 0.697 ACL

0.778 0.348 0.875 0.778 0.824 0.393 0.852 0.945 nonACL

Weighted Avg. 0.747 0.317 0.780 0.747 0.759 0.393 0.852 0.885

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

16 56 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 0.29 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 72 75.7895 %

Incorrectly Classified Instances 23 24.2105 %

Kappa statistic 0.4326

Mean absolute error 0.3317

Root mean squared error 0.3915

Relative absolute error 89.5706 %

Root relative squared error 91.2943 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.739 0.236 0.500 0.739 0.596 0.449 0.851 0.637 ACL

0.764 0.261 0.902 0.764 0.827 0.449 0.851 0.952 nonACL

Weighted Avg. 0.758 0.255 0.804 0.758 0.771 0.449 0.851 0.876

=== Confusion Matrix ===

a b <-- classified as

17 6 | a = ACL

17 55 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 0.29 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.4297

Mean absolute error 0.3376

Root mean squared error 0.3927

Relative absolute error 91.1716 %

Root relative squared error 91.5747 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.783 0.264 0.486 0.783 0.600 0.456 0.859 0.605 ACL

0.736 0.217 0.914 0.736 0.815 0.456 0.859 0.955 nonACL

Weighted Avg. 0.747 0.229 0.810 0.747 0.763 0.456 0.859 0.870

=== Confusion Matrix ===

a b <-- classified as

18 5 | a = ACL

19 53 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 0.29 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.4154

Mean absolute error 0.3356

Root mean squared error 0.3945

Relative absolute error 90.62 %

Root relative squared error 91.9842 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.739 0.250 0.486 0.739 0.586 0.434 0.859 0.647 ACL

0.750 0.261 0.900 0.750 0.818 0.434 0.859 0.954 nonACL

Weighted Avg. 0.747 0.258 0.800 0.747 0.762 0.434 0.859 0.879

=== Confusion Matrix ===

a b <-- classified as

17 6 | a = ACL

18 54 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 5

1 0

Time taken to build model: 0.28 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 73 76.8421 %

Incorrectly Classified Instances 22 23.1579 %

Kappa statistic 0.4897

Mean absolute error 0.3529

Root mean squared error 0.4073

Relative absolute error 95.3042 %

Root relative squared error 94.9758 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.870 0.264 0.513 0.870 0.645 0.527 0.866 0.665 ACL

0.736 0.130 0.946 0.736 0.828 0.527 0.866 0.955 nonACL

Weighted Avg. 0.768 0.163 0.841 0.768 0.784 0.527 0.866 0.885

=== Confusion Matrix ===

a b <-- classified as

20 3 | a = ACL

19 53 | b = nonACL

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and J48**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 83 87.3684 %

Incorrectly Classified Instances 12 12.6316 %

Kappa statistic 0.6453

Mean absolute error 0.1286

Root mean squared error 0.3554

Relative absolute error 34.7182 %

Root relative squared error 82.8727 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.696 0.069 0.762 0.696 0.727 0.646 0.829 0.676 ACL

0.931 0.304 0.905 0.931 0.918 0.646 0.830 0.903 nonACL

Weighted Avg. 0.874 0.247 0.871 0.874 0.872 0.646 0.830 0.848

=== Confusion Matrix ===

a b <-- classified as

16 7 | a = ACL

5 67 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 1.45 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 82 86.3158 %

Incorrectly Classified Instances 13 13.6842 %

Kappa statistic 0.5973

Mean absolute error 0.1442

Root mean squared error 0.3715

Relative absolute error 38.9344 %

Root relative squared error 86.6184 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.056 0.778 0.609 0.683 0.605 0.886 0.806 ACL

0.944 0.391 0.883 0.944 0.913 0.605 0.891 0.946 nonACL

Weighted Avg. 0.863 0.310 0.858 0.863 0.857 0.605 0.889 0.912

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 1.24 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.456

Mean absolute error 0.1777

Root mean squared error 0.404

Relative absolute error 47.9947 %

Root relative squared error 94.1999 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.478 0.069 0.688 0.478 0.564 0.468 0.733 0.537 ACL

0.931 0.522 0.848 0.931 0.887 0.468 0.742 0.854 nonACL

Weighted Avg. 0.821 0.412 0.809 0.821 0.809 0.468 0.740 0.777

=== Confusion Matrix ===

a b <-- classified as

11 12 | a = ACL

5 67 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 1.15 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 74 77.8947 %

Incorrectly Classified Instances 21 22.1053 %

Kappa statistic 0.3495

Mean absolute error 0.2197

Root mean squared error 0.4585

Relative absolute error 59.3189 %

Root relative squared error 106.9038 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.435 0.111 0.556 0.435 0.488 0.354 0.689 0.462 ACL

0.889 0.565 0.831 0.889 0.859 0.354 0.706 0.847 nonACL

Weighted Avg. 0.779 0.455 0.764 0.779 0.769 0.354 0.702 0.754

=== Confusion Matrix ===

a b <-- classified as

10 13 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 1.4 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.4734

Mean absolute error 0.1896

Root mean squared error 0.4256

Relative absolute error 51.1992 %

Root relative squared error 99.2339 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.083 0.667 0.522 0.585 0.479 0.671 0.442 ACL

0.917 0.478 0.857 0.917 0.886 0.479 0.665 0.820 nonACL

Weighted Avg. 0.821 0.383 0.811 0.821 0.813 0.479 0.667 0.728

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 1.23 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.4734

Mean absolute error 0.1844

Root mean squared error 0.4164

Relative absolute error 49.7936 %

Root relative squared error 97.1069 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.083 0.667 0.522 0.585 0.479 0.718 0.549 ACL

0.917 0.478 0.857 0.917 0.886 0.479 0.730 0.851 nonACL

Weighted Avg. 0.821 0.383 0.811 0.821 0.813 0.479 0.727 0.778

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 1.1 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 75 78.9474 %

Incorrectly Classified Instances 20 21.0526 %

Kappa statistic 0.3902

Mean absolute error 0.2058

Root mean squared error 0.4455

Relative absolute error 55.5598 %

Root relative squared error 103.8722 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.478 0.111 0.579 0.478 0.524 0.393 0.747 0.555 ACL

0.889 0.522 0.842 0.889 0.865 0.393 0.767 0.879 nonACL

Weighted Avg. 0.789 0.422 0.778 0.789 0.782 0.393 0.762 0.801

=== Confusion Matrix ===

a b <-- classified as

11 12 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 1.15 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 75 78.9474 %

Incorrectly Classified Instances 20 21.0526 %

Kappa statistic 0.3704

Mean absolute error 0.2072

Root mean squared error 0.4482

Relative absolute error 55.9404 %

Root relative squared error 104.5123 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.435 0.097 0.588 0.435 0.500 0.377 0.704 0.544 ACL

0.903 0.565 0.833 0.903 0.867 0.377 0.713 0.837 nonACL

Weighted Avg. 0.789 0.452 0.774 0.789 0.778 0.377 0.711 0.766

=== Confusion Matrix ===

a b <-- classified as

10 13 | a = ACL

7 65 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 5

1 0

Time taken to build model: 1.16 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 76 80 %

Incorrectly Classified Instances 19 20 %

Kappa statistic 0.3713

Mean absolute error 0.1987

Root mean squared error 0.4276

Relative absolute error 53.6606 %

Root relative squared error 99.713 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.391 0.069 0.643 0.391 0.486 0.389 0.697 0.535 ACL

0.931 0.609 0.827 0.931 0.876 0.389 0.711 0.840 nonACL

Weighted Avg. 0.800 0.478 0.783 0.800 0.782 0.389 0.708 0.766

=== Confusion Matrix ===

a b <-- classified as

9 14 | a = ACL

5 67 | b = nonACL

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump.**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 0.53 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5862

Mean absolute error 0.143

Root mean squared error 0.3335

Relative absolute error 38.628 %

Root relative squared error 77.7726 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.083 0.714 0.652 0.682 0.587 0.870 0.780 ACL

0.917 0.348 0.892 0.917 0.904 0.587 0.870 0.946 nonACL

Weighted Avg. 0.853 0.284 0.849 0.853 0.850 0.587 0.870 0.906

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 0.57 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5862

Mean absolute error 0.1635

Root mean squared error 0.3693

Relative absolute error 44.1632 %

Root relative squared error 86.1191 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.083 0.714 0.652 0.682 0.587 0.809 0.693 ACL

0.917 0.348 0.892 0.917 0.904 0.587 0.809 0.902 nonACL

Weighted Avg. 0.853 0.284 0.849 0.853 0.850 0.587 0.809 0.852

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 0.56 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 76 80 %

Incorrectly Classified Instances 19 20 %

Kappa statistic 0.4297

Mean absolute error 0.2006

Root mean squared error 0.4243

Relative absolute error 54.1804 %

Root relative squared error 98.9511 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.111 0.600 0.522 0.558 0.431 0.772 0.665 ACL

0.889 0.478 0.853 0.889 0.871 0.431 0.772 0.890 nonACL

Weighted Avg. 0.800 0.389 0.792 0.800 0.795 0.431 0.772 0.836

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 0.55 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.4897

Mean absolute error 0.1851

Root mean squared error 0.3948

Relative absolute error 49.9873 %

Root relative squared error 92.0498 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.097 0.650 0.565 0.605 0.492 0.803 0.681 ACL

0.903 0.435 0.867 0.903 0.884 0.492 0.803 0.904 nonACL

Weighted Avg. 0.821 0.353 0.814 0.821 0.817 0.492 0.803 0.850

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

7 65 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 0.54 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 75 78.9474 %

Incorrectly Classified Instances 20 21.0526 %

Kappa statistic 0.4088

Mean absolute error 0.2102

Root mean squared error 0.4218

Relative absolute error 56.7721 %

Root relative squared error 98.3679 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.125 0.571 0.522 0.545 0.410 0.825 0.628 ACL

0.875 0.478 0.851 0.875 0.863 0.410 0.825 0.938 nonACL

Weighted Avg. 0.789 0.393 0.784 0.789 0.786 0.410 0.825 0.863

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 0.55 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 74 77.8947 %

Incorrectly Classified Instances 21 22.1053 %

Kappa statistic 0.4064

Mean absolute error 0.2156

Root mean squared error 0.4281

Relative absolute error 58.2236 %

Root relative squared error 99.8315 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.153 0.542 0.565 0.553 0.407 0.810 0.601 ACL

0.847 0.435 0.859 0.847 0.853 0.407 0.810 0.928 nonACL

Weighted Avg. 0.779 0.367 0.782 0.779 0.781 0.407 0.810 0.849

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

11 61 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 0.55 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 77 81.0526 %

Incorrectly Classified Instances 18 18.9474 %

Kappa statistic 0.4837

Mean absolute error 0.1949

Root mean squared error 0.4102

Relative absolute error 52.6163 %

Root relative squared error 95.6484 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.125 0.609 0.609 0.609 0.484 0.804 0.599 ACL

0.875 0.391 0.875 0.875 0.875 0.484 0.804 0.912 nonACL

Weighted Avg. 0.811 0.327 0.811 0.811 0.811 0.484 0.804 0.836

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 0.87 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 76 80 %

Incorrectly Classified Instances 19 20 %

Kappa statistic 0.463

Mean absolute error 0.2139

Root mean squared error 0.4342

Relative absolute error 57.7679 %

Root relative squared error 101.2403 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.139 0.583 0.609 0.596 0.463 0.775 0.588 ACL

0.861 0.391 0.873 0.861 0.867 0.463 0.775 0.884 nonACL

Weighted Avg. 0.800 0.330 0.803 0.800 0.801 0.463 0.775 0.812

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

10 62 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 5

1 0

Time taken to build model: 0.56 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 73 76.8421 %

Incorrectly Classified Instances 22 23.1579 %

Kappa statistic 0.3871

Mean absolute error 0.2374

Root mean squared error 0.462

Relative absolute error 64.0961 %

Root relative squared error 107.7361 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.167 0.520 0.565 0.542 0.388 0.771 0.567 ACL

0.833 0.435 0.857 0.833 0.845 0.388 0.771 0.892 nonACL

Weighted Avg. 0.768 0.370 0.776 0.768 0.772 0.388 0.771 0.813

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

12 60 | b = nonACL

1. **Cost sensitive classifier combined with bagging and J48 plus meta learner iteration set to 25.**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 3.25 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 82 86.3158 %

Incorrectly Classified Instances 13 13.6842 %

Kappa statistic 0.584

Mean absolute error 0.2442

Root mean squared error 0.317

Relative absolute error 65.9371 %

Root relative squared error 73.9305 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.042 0.813 0.565 0.667 0.599 0.920 0.834 ACL

0.958 0.435 0.873 0.958 0.914 0.599 0.920 0.968 nonACL

Weighted Avg. 0.863 0.340 0.859 0.863 0.854 0.599 0.920 0.935

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 2.52 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 83 87.3684 %

Incorrectly Classified Instances 12 12.6316 %

Kappa statistic 0.6223

Mean absolute error 0.247

Root mean squared error 0.3203

Relative absolute error 66.7043 %

Root relative squared error 74.695 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.042 0.824 0.609 0.700 0.634 0.911 0.814 ACL

0.958 0.391 0.885 0.958 0.920 0.634 0.911 0.966 nonACL

Weighted Avg. 0.874 0.307 0.870 0.874 0.867 0.634 0.911 0.929

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 2.32 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.4964

Mean absolute error 0.2557

Root mean squared error 0.3193

Relative absolute error 69.0532 %

Root relative squared error 74.4614 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.069 0.706 0.522 0.600 0.505 0.928 0.850 ACL

0.931 0.478 0.859 0.931 0.893 0.505 0.928 0.971 nonACL

Weighted Avg. 0.832 0.379 0.822 0.832 0.822 0.505 0.928 0.942

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

5 67 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 2.17 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 84 88.4211 %

Incorrectly Classified Instances 11 11.5789 %

Kappa statistic 0.6698

Mean absolute error 0.2671

Root mean squared error 0.3375

Relative absolute error 72.1303 %

Root relative squared error 78.6991 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.696 0.056 0.800 0.696 0.744 0.673 0.886 0.780 ACL

0.944 0.304 0.907 0.944 0.925 0.673 0.886 0.957 nonACL

Weighted Avg. 0.884 0.244 0.881 0.884 0.881 0.673 0.886 0.914

=== Confusion Matrix ===

a b <-- classified as

16 7 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 2.1 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.5411

Mean absolute error 0.2707

Root mean squared error 0.3455

Relative absolute error 73.1107 %

Root relative squared error 80.5675 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.111 0.652 0.652 0.652 0.541 0.878 0.745 ACL

0.889 0.348 0.889 0.889 0.889 0.541 0.878 0.949 nonACL

Weighted Avg. 0.832 0.291 0.832 0.832 0.832 0.541 0.878 0.900

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 2.05 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 79 83.1579 %

Incorrectly Classified Instances 16 16.8421 %

Kappa statistic 0.5411

Mean absolute error 0.2739

Root mean squared error 0.3395

Relative absolute error 73.9629 %

Root relative squared error 79.1645 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.111 0.652 0.652 0.652 0.541 0.899 0.807 ACL

0.889 0.348 0.889 0.889 0.889 0.541 0.899 0.963 nonACL

Weighted Avg. 0.832 0.291 0.832 0.832 0.832 0.541 0.899 0.925

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 2.11 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.3116

Mean absolute error 0.2825

Root mean squared error 0.353

Relative absolute error 76.2807 %

Root relative squared error 82.3159 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.478 0.167 0.478 0.478 0.478 0.312 0.870 0.731 ACL

0.833 0.522 0.833 0.833 0.833 0.312 0.870 0.956 nonACL

Weighted Avg. 0.747 0.436 0.747 0.747 0.747 0.312 0.870 0.902

=== Confusion Matrix ===

a b <-- classified as

11 12 | a = ACL

12 60 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 2.02 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 77 81.0526 %

Incorrectly Classified Instances 18 18.9474 %

Kappa statistic 0.4985

Mean absolute error 0.3001

Root mean squared error 0.3602

Relative absolute error 81.0261 %

Root relative squared error 83.9862 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.139 0.600 0.652 0.625 0.499 0.873 0.759 ACL

0.861 0.348 0.886 0.861 0.873 0.499 0.873 0.952 nonACL

Weighted Avg. 0.811 0.297 0.817 0.811 0.813 0.499 0.873 0.905

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

10 62 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 1

1 0

Time taken to build model: 0.96 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.3963

Mean absolute error 0.2837

Root mean squared error 0.3615

Relative absolute error 76.6151 %

Root relative squared error 84.2905 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.348 0.028 0.800 0.348 0.485 0.447 0.825 0.716 ACL

0.972 0.652 0.824 0.972 0.892 0.447 0.825 0.929 nonACL

Weighted Avg. 0.821 0.501 0.818 0.821 0.793 0.447 0.825 0.877

=== Confusion Matrix ===

a b <-- classified as

8 15 | a = ACL

2 70 | b = nonACL

1. **Cost sensitive classifier combined with bagging and Decision Stump plus meta learner iteration set to 25**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 0.96 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.3963

Mean absolute error 0.2837

Root mean squared error 0.3615

Relative absolute error 76.6151 %

Root relative squared error 84.2905 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.348 0.028 0.800 0.348 0.485 0.447 0.825 0.716 ACL

0.972 0.652 0.824 0.972 0.892 0.447 0.825 0.929 nonACL

Weighted Avg. 0.821 0.501 0.818 0.821 0.793 0.447 0.825 0.877

=== Confusion Matrix ===

a b <-- classified as

8 15 | a = ACL

2 70 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 1.02 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5287

Mean absolute error 0.2947

Root mean squared error 0.3516

Relative absolute error 79.5759 %

Root relative squared error 81.9946 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.478 0.028 0.846 0.478 0.611 0.561 0.867 0.779 ACL

0.972 0.522 0.854 0.972 0.909 0.561 0.867 0.947 nonACL

Weighted Avg. 0.853 0.402 0.852 0.853 0.837 0.561 0.867 0.906

=== Confusion Matrix ===

a b <-- classified as

11 12 | a = ACL

2 70 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 1 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.5051

Mean absolute error 0.32

Root mean squared error 0.3695

Relative absolute error 86.4148 %

Root relative squared error 86.1571 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.111 0.636 0.609 0.622 0.505 0.839 0.738 ACL

0.889 0.391 0.877 0.889 0.883 0.505 0.839 0.936 nonACL

Weighted Avg. 0.821 0.323 0.819 0.821 0.820 0.505 0.839 0.888

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

8 64 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 0.76 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.3677

Mean absolute error 0.3405

Root mean squared error 0.3936

Relative absolute error 91.9488 %

Root relative squared error 91.7766 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.208 0.483 0.609 0.538 0.372 0.828 0.698 ACL

0.792 0.391 0.864 0.792 0.826 0.372 0.828 0.936 nonACL

Weighted Avg. 0.747 0.347 0.771 0.747 0.756 0.372 0.828 0.878

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

15 57 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 0.96 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 71 74.7368 %

Incorrectly Classified Instances 24 25.2632 %

Kappa statistic 0.4154

Mean absolute error 0.3459

Root mean squared error 0.3945

Relative absolute error 93.403 %

Root relative squared error 91.9885 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.739 0.250 0.486 0.739 0.586 0.434 0.842 0.721 ACL

0.750 0.261 0.900 0.750 0.818 0.434 0.842 0.942 nonACL

Weighted Avg. 0.747 0.258 0.800 0.747 0.762 0.434 0.842 0.889

=== Confusion Matrix ===

a b <-- classified as

17 6 | a = ACL

18 54 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 0.76 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 66 69.4737 %

Incorrectly Classified Instances 29 30.5263 %

Kappa statistic 0.3192

Mean absolute error 0.3487

Root mean squared error 0.4068

Relative absolute error 94.1532 %

Root relative squared error 94.8652 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.696 0.306 0.421 0.696 0.525 0.341 0.838 0.714 ACL

0.694 0.304 0.877 0.694 0.775 0.341 0.838 0.941 nonACL

Weighted Avg. 0.695 0.305 0.767 0.695 0.715 0.341 0.838 0.886

=== Confusion Matrix ===

a b <-- classified as

16 7 | a = ACL

22 50 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 0.74 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 69 72.6316 %

Incorrectly Classified Instances 26 27.3684 %

Kappa statistic 0.3667

Mean absolute error 0.3533

Root mean squared error 0.4041

Relative absolute error 95.4033 %

Root relative squared error 94.2334 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.696 0.264 0.457 0.696 0.552 0.383 0.850 0.699 ACL

0.736 0.304 0.883 0.736 0.803 0.383 0.850 0.948 nonACL

Weighted Avg. 0.726 0.295 0.780 0.726 0.742 0.383 0.850 0.888

=== Confusion Matrix ===

a b <-- classified as

16 7 | a = ACL

19 53 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 0.73 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 66 69.4737 %

Incorrectly Classified Instances 29 30.5263 %

Kappa statistic 0.3353

Mean absolute error 0.3578

Root mean squared error 0.4126

Relative absolute error 96.6163 %

Root relative squared error 96.2123 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.739 0.319 0.425 0.739 0.540 0.364 0.838 0.676 ACL

0.681 0.261 0.891 0.681 0.772 0.364 0.838 0.944 nonACL

Weighted Avg. 0.695 0.275 0.778 0.695 0.715 0.364 0.838 0.879

=== Confusion Matrix ===

a b <-- classified as

17 6 | a = ACL

23 49 | b = nonACL

* 1. **Cost set to 5**

Time taken to build model: 0.93 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 66 69.4737 %

Incorrectly Classified Instances 29 30.5263 %

Kappa statistic 0.3353

Mean absolute error 0.3679

Root mean squared error 0.4244

Relative absolute error 99.3489 %

Root relative squared error 98.953 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.739 0.319 0.425 0.739 0.540 0.364 0.838 0.699 ACL

0.681 0.261 0.891 0.681 0.772 0.364 0.838 0.944 nonACL

Weighted Avg. 0.695 0.275 0.778 0.695 0.715 0.364 0.838 0.885

=== Confusion Matrix ===

a b <-- classified as

17 6 | a = ACL

23 49 | b = nonACL

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and J48 plus meta learner iteration set to 25.**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 0.13 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5862

Mean absolute error 0.1503

Root mean squared error 0.3794

Relative absolute error 40.5846 %

Root relative squared error 88.4727 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.083 0.714 0.652 0.682 0.587 0.832 0.691 ACL

0.917 0.348 0.892 0.917 0.904 0.587 0.859 0.926 nonACL

Weighted Avg. 0.853 0.284 0.849 0.853 0.850 0.587 0.852 0.869

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 3.54 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 83 87.3684 %

Incorrectly Classified Instances 12 12.6316 %

Kappa statistic 0.6341

Mean absolute error 0.1269

Root mean squared error 0.3402

Relative absolute error 34.2791 %

Root relative squared error 79.3359 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.056 0.789 0.652 0.714 0.639 0.895 0.823 ACL

0.944 0.348 0.895 0.944 0.919 0.639 0.914 0.958 nonACL

Weighted Avg. 0.874 0.277 0.869 0.874 0.869 0.639 0.909 0.926

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 3.01 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 82 86.3158 %

Incorrectly Classified Instances 13 13.6842 %

Kappa statistic 0.584

Mean absolute error 0.143

Root mean squared error 0.3722

Relative absolute error 38.606 %

Root relative squared error 86.794 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.042 0.813 0.565 0.667 0.599 0.759 0.599 ACL

0.958 0.435 0.873 0.958 0.914 0.599 0.809 0.901 nonACL

Weighted Avg. 0.863 0.340 0.859 0.863 0.854 0.599 0.797 0.828

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 3 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 77 81.0526 %

Incorrectly Classified Instances 18 18.9474 %

Kappa statistic 0.4334

Mean absolute error 0.1835

Root mean squared error 0.424

Relative absolute error 49.5467 %

Root relative squared error 98.8595 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.478 0.083 0.647 0.478 0.550 0.441 0.711 0.519 ACL

0.917 0.522 0.846 0.917 0.880 0.441 0.736 0.862 nonACL

Weighted Avg. 0.811 0.416 0.798 0.811 0.800 0.441 0.730 0.779

=== Confusion Matrix ===

a b <-- classified as

11 12 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 2.78 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.4734

Mean absolute error 0.1834

Root mean squared error 0.4223

Relative absolute error 49.5208 %

Root relative squared error 98.4687 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.083 0.667 0.522 0.585 0.479 0.692 0.486 ACL

0.917 0.478 0.857 0.917 0.886 0.479 0.722 0.856 nonACL

Weighted Avg. 0.821 0.383 0.811 0.821 0.813 0.479 0.715 0.766

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 3.07 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5445

Mean absolute error 0.1529

Root mean squared error 0.3842

Relative absolute error 41.2843 %

Root relative squared error 89.5844 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.042 0.800 0.522 0.632 0.564 0.766 0.664 ACL

0.958 0.478 0.863 0.958 0.908 0.564 0.805 0.897 nonACL

Weighted Avg. 0.853 0.373 0.847 0.853 0.841 0.564 0.795 0.841

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 2.82 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 80 84.2105 %

Incorrectly Classified Instances 15 15.7895 %

Kappa statistic 0.5633

Mean absolute error 0.1609

Root mean squared error 0.3964

Relative absolute error 43.4554 %

Root relative squared error 92.4227 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.097 0.682 0.652 0.667 0.564 0.771 0.641 ACL

0.903 0.348 0.890 0.903 0.897 0.564 0.800 0.895 nonACL

Weighted Avg. 0.842 0.287 0.840 0.842 0.841 0.564 0.793 0.834

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

7 65 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 2.88 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 76 80 %

Incorrectly Classified Instances 19 20 %

Kappa statistic 0.3921

Mean absolute error 0.1978

Root mean squared error 0.4423

Relative absolute error 53.3999 %

Root relative squared error 103.1264 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.435 0.083 0.625 0.435 0.513 0.402 0.725 0.615 ACL

0.917 0.565 0.835 0.917 0.874 0.402 0.765 0.875 nonACL

Weighted Avg. 0.800 0.449 0.784 0.800 0.787 0.402 0.755 0.812

=== Confusion Matrix ===

a b <-- classified as

10 13 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 5

1 0

Time taken to build model: 2.89 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.456

Mean absolute error 0.1748

Root mean squared error 0.4145

Relative absolute error 47.1946 %

Root relative squared error 96.6536 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.478 0.069 0.688 0.478 0.564 0.468 0.704 0.565 ACL

0.931 0.522 0.848 0.931 0.887 0.468 0.759 0.873 nonACL

Weighted Avg. 0.821 0.412 0.809 0.821 0.809 0.468 0.746 0.799

=== Confusion Matrix ===

a b <-- classified as

11 12 | a = ACL

5 67 | b = nonACL

1. **Cost sensitive classifier combined with boosting (AdaBoostM1) and Decision Stump plus meta learner iteration set to 25**
   1. **Cost set to 1**

Cost Matrix

0 1

1 0

Time taken to build model: 1.36 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 85 89.4737 %

Incorrectly Classified Instances 10 10.5263 %

Kappa statistic 0.7132

Mean absolute error 0.0995

Root mean squared error 0.2919

Relative absolute error 26.8809 %

Root relative squared error 68.0697 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.783 0.069 0.783 0.783 0.783 0.713 0.940 0.901 ACL

0.931 0.217 0.931 0.931 0.931 0.713 0.940 0.978 nonACL

Weighted Avg. 0.895 0.182 0.895 0.895 0.895 0.713 0.940 0.959

=== Confusion Matrix ===

a b <-- classified as

18 5 | a = ACL

5 67 | b = nonACL

* 1. **Cost set to 1.5**

Cost Matrix

0 1.5

1 0

Time taken to build model: 1.66 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 81 85.2632 %

Incorrectly Classified Instances 14 14.7368 %

Kappa statistic 0.5445

Mean absolute error 0.1449

Root mean squared error 0.3628

Relative absolute error 39.1255 %

Root relative squared error 84.5966 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.042 0.800 0.522 0.632 0.564 0.877 0.779 ACL

0.958 0.478 0.863 0.958 0.908 0.564 0.876 0.932 nonACL

Weighted Avg. 0.853 0.373 0.847 0.853 0.841 0.564 0.876 0.895

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

3 69 | b = nonACL

* 1. **Cost set to 2**

Cost Matrix

0 2

1 0

Time taken to build model: 1.37 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 84 88.4211 %

Incorrectly Classified Instances 11 11.5789 %

Kappa statistic 0.648

Mean absolute error 0.1349

Root mean squared error 0.3408

Relative absolute error 36.4165 %

Root relative squared error 79.4708 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.028 0.875 0.609 0.718 0.665 0.855 0.772 ACL

0.972 0.391 0.886 0.972 0.927 0.665 0.855 0.937 nonACL

Weighted Avg. 0.884 0.303 0.883 0.884 0.877 0.665 0.855 0.897

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

2 70 | b = nonACL

* 1. **Cost set to 2.5**

Cost Matrix

0 2.5

1 0

Time taken to build model: 1.62 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 83 87.3684 %

Incorrectly Classified Instances 12 12.6316 %

Kappa statistic 0.6096

Mean absolute error 0.1392

Root mean squared error 0.3479

Relative absolute error 37.5914 %

Root relative squared error 81.126 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.028 0.867 0.565 0.684 0.631 0.886 0.802 ACL

0.972 0.435 0.875 0.972 0.921 0.631 0.886 0.948 nonACL

Weighted Avg. 0.874 0.336 0.873 0.874 0.864 0.631 0.886 0.913

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

2 70 | b = nonACL

* 1. **Cost set to 3**

Cost Matrix

0 3

1 0

Time taken to build model: 1.39 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 77 81.0526 %

Incorrectly Classified Instances 18 18.9474 %

Kappa statistic 0.4512

Mean absolute error 0.1826

Root mean squared error 0.4086

Relative absolute error 49.3097 %

Root relative squared error 95.2693 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.097 0.632 0.522 0.571 0.455 0.882 0.763 ACL

0.903 0.478 0.855 0.903 0.878 0.455 0.882 0.959 nonACL

Weighted Avg. 0.811 0.386 0.801 0.811 0.804 0.455 0.882 0.911

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

7 65 | b = nonACL

* 1. **Cost set to 3.5**

Cost Matrix

0 3.5

1 0

Time taken to build model: 1.44 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 76 80 %

Incorrectly Classified Instances 19 20 %

Kappa statistic 0.4468

Mean absolute error 0.1879

Root mean squared error 0.4104

Relative absolute error 50.7335 %

Root relative squared error 95.7093 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.565 0.125 0.591 0.565 0.578 0.447 0.889 0.742 ACL

0.875 0.435 0.863 0.875 0.869 0.447 0.889 0.964 nonACL

Weighted Avg. 0.800 0.360 0.797 0.800 0.798 0.447 0.889 0.910

=== Confusion Matrix ===

a b <-- classified as

13 10 | a = ACL

9 63 | b = nonACL

* 1. **Cost set to 4**

Cost Matrix

0 4

1 0

Time taken to build model: 1.38 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 80 84.2105 %

Incorrectly Classified Instances 15 15.7895 %

Kappa statistic 0.5498

Mean absolute error 0.1518

Root mean squared error 0.3666

Relative absolute error 40.9966 %

Root relative squared error 85.4955 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.609 0.083 0.700 0.609 0.651 0.552 0.882 0.721 ACL

0.917 0.391 0.880 0.917 0.898 0.552 0.882 0.958 nonACL

Weighted Avg. 0.842 0.317 0.836 0.842 0.838 0.552 0.882 0.900

=== Confusion Matrix ===

a b <-- classified as

14 9 | a = ACL

6 66 | b = nonACL

* 1. **Cost set to 4.5**

Cost Matrix

0 4.5

1 0

Time taken to build model: 1.35 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 83 87.3684 %

Incorrectly Classified Instances 12 12.6316 %

Kappa statistic 0.6341

Mean absolute error 0.1417

Root mean squared error 0.36

Relative absolute error 38.2564 %

Root relative squared error 83.9359 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.652 0.056 0.789 0.652 0.714 0.639 0.854 0.678 ACL

0.944 0.348 0.895 0.944 0.919 0.639 0.854 0.944 nonACL

Weighted Avg. 0.874 0.277 0.869 0.874 0.869 0.639 0.854 0.880

=== Confusion Matrix ===

a b <-- classified as

15 8 | a = ACL

4 68 | b = nonACL

* 1. **Cost set to 5**

Cost Matrix

0 5

1 0

Time taken to build model: 1.57 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 78 82.1053 %

Incorrectly Classified Instances 17 17.8947 %

Kappa statistic 0.4734

Mean absolute error 0.1802

Root mean squared error 0.4096

Relative absolute error 48.6726 %

Root relative squared error 95.5086 %

Total Number of Instances 95

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.522 0.083 0.667 0.522 0.585 0.479 0.838 0.617 ACL

0.917 0.478 0.857 0.917 0.886 0.479 0.838 0.943 nonACL

Weighted Avg. 0.821 0.383 0.811 0.821 0.813 0.479 0.838 0.864

=== Confusion Matrix ===

a b <-- classified as

12 11 | a = ACL

6 66 | b = nonACL