Error Sensitive Grading (ESGrading) for Model Combination

CSE 572 - Project

Surendra Singhi

April 28, 2005

Outline

Two Steps of creating an Ensemble

Model Generation - Create the ensemble

- How do we create diverse collection of classifiers?
- Re-sampling, using different learning algorithm, various other strategies.

Two Steps of creating an Ensemble

Model Generation - Create the ensemble

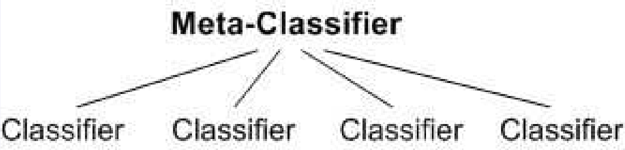
- How do we create diverse collection of classifiers?
- Re-sampling, using different learning algorithm, various other strategies.

Model Combination - Combine the predictions of base classifiers

- What is the best way of combining the predictions of the classifiers?

Meta Classification Framework

- Has classifiers at two levels.
- Base level or low level classifiers generated during model generation phase.
- Meta level classifier created during model combination phase.



Categorization of Model Combination Strategies

Voting- simple but popular

- Majority Voting
- Weighted Voting
- Thresholded Voting

Stacking - tries to find the relationship between predictions of base classifier and the actual class

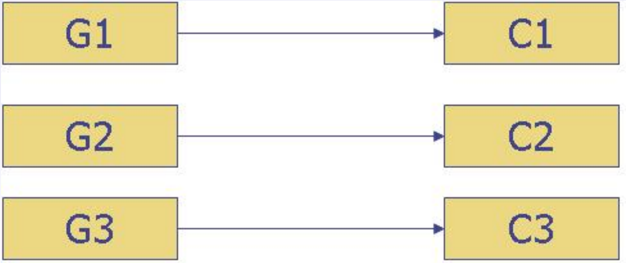
- Stacking with MLR, MDT
- StackingC

Grading/Referee Method - how much should the base classifier be trusted

- Uses another ensemble to do model combination

Grading

- For each base classifier, there is a grader classifier.
- When grader predicts base classifier to be correct, use it else ignore it.
- Combine the predictions of base classifiers which were predicted to be correct by using weighted voting.



Type A error vs. Type B error

Definition (Type A error)

Base classifier is correct but grader predicts it to be wrong.

Definition (Type B error)

Base classifier is wrong but grader predicts it to be correct.

Type A error vs. Type B error

Definition (Type A error)

Base classifier is correct but grader predicts it to be wrong.

Definition (Type B error)

Base classifier is wrong but grader predicts it to be correct.

Fact

*Both errors are bad but Type B error is **worse**.*

Basic Idea

Idea

Reduce Type B error, may be at the expense of increasing Type A error.

Basic Idea

Idea

Reduce Type B error, may be at the expense of increasing Type A error.

Why it may work?

If there are large numbers of base classifiers then being conservative about picking base classifiers and missing some good classifiers may not hurt performance much, but choosing some wrong base classifiers may degrade the accuracy.

Basic Idea

Idea

Reduce Type B error, may be at the expense of increasing Type A error.

Why it may work?

If there are large numbers of base classifiers then being conservative about picking base classifiers and missing some good classifiers may not hurt performance much, but choosing some wrong base classifiers may degrade the accuracy.

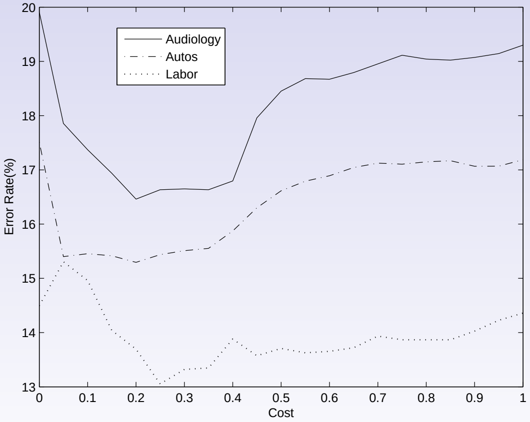
How do we do it?

Use Cost-Sensitive Learning.

Some More Details. . .

- Type A error has some cost associated with it.
- Type B error also has some cost.
- But because it is binary classification problem, a **single cost ratio** can be used to express the trade-off.
- Determine a good cost-ratio, and then train the graders using cost-sensitive learning.

Error vs. Cost-ratio



How do we determine cost-ratio?

Cross-validation

Try out different cost-ratio values using cross-validation.

Time Complexity

If c-fold cross-validation is done with t different cost-ratio values then, the complexity is $O(c * t * G)$ where G is the time complexity of the grading algorithm.

Experiments

- Used 19 datasets from UCI-ML Repository
- 10 * 10-fold cross-validation.
- Modified existing Weka classifiers to implement everything and do experiments
- Used one-tailed paired t-test, loss-tie-win at 95% significance level.

Maj. Voting	*Wt. Voting*	*Stacking MT*	*Stacking MLR*	*StackingC*	*Grading*
3/10/6	2/13/4	0/10/9	0/12/7	0/13/6	2/10/7

Conclusion & Future Work

Conclusion

- Introduced a novel approach of doing grading.
- Suggested using cross-validation to determine cost-ratio.
- Showed that this method gives better results than existing methods on benchmark datasets.

Future Work

- Research on efficient way of determine cost, maybe ROC analysis.
- Research if similar modification can be made to Stacking methods.

References

- Saso Dzeroski and Bernard Zenko.
  Is combining classifiers better than selecting the best one?
  Machine Learning, 54(3):255–273, March 2004.
- Alexander K. Seewald and J. Furnkranz.
  An evaluation of grading classifiers.
  In *Advances in Intelligent Data Analysis:Proceedings of the 4th International Symposium*. Springer-Verlag, 2001.

Thank you.