A Structured Approach to Ensemble Learning for Alzheimer's Disease Prediction

Matthew Seeley
BYU CS Department
Provo, UT 84602-6576
seeleymatthew@hotmail.com

Mark Clement
BYU CS Department
Provo, UT 84602-6576

Christophe Giraud-Carrier
BYU CS Department
Provo, UT 84602-6576

Quinn Snell
BYU CS Department
Provo, UT 84602-6576

Paul Bodily
BYU CS Department
Provo, UT 84602-6576

Stanley Fujimoto
BYU CS Department
Provo, UT 84602-6576

ABSTRACT

This research employs an exhaustive search of different attribute selection algorithms in order to provide a more structured approach to learning design for prediction of Alzheimer's clinical dementia rating (CDR).

Categories and Subject Descriptors

I.5.1 [**Pattern Recognition**]: Models

General Terms

Design, Experimentation, Performance

Keywords

Ensemble Machine Learning, Alzheimer's Disease, CDR

1. INTRODUCTION

Alzheimer's disease causes progressive and fatal degeneration of the brain. It leads to dementia, a condition in which a loss of memory and mental abilities is severe enough to affect a person's daily life [3]. It is a commonly occurring disease amongst the elderly; in 2006, there were more than 26 million cases of Alzheimer's disease worldwide [3]. Moreover, because of the global increase in average lifespan, it is estimated that Alzheimer's disease may affect over 115 million people worldwide by the year 2050 [4]. A recent study suggests these predictions may underestimate the true prevalence of the disease [1]. If that is the case, current estimates that healthcare costs associated with Alzheimer's disease in the US in 2013 exceeded 200 billion USD [4] may be conservative.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is an ongoing project that aims to identify biomarkers for Alzheimer's disease [2] so that such tests can be developed. The ADNI data set currently has information about more than 1,000 participants. For each participant, over 100 attributes are recorded; some of these attributes include age, volumes of various parts of the brain, levels of specific proteins in cerebrospinal fluid, and biomarkers from the Myriad Genetics Rules Based Medicine panel, and alleles for certain genes. (A recently published study suggests that the levels of certain lipids found in the blood would also be helpful attributes [4], but the levels of these lipids in the ADNI participants are not yet available.)

The ADNI data includes the clinical dementia rating (CDR) of 821 participants. In the ADNI data, participants are divided into three possible categories defined in the CDR scale: normal (no cognitive impairment), mildly cognitively impaired, and moderately cognitively impaired (i.e., requiring some help for basic personal care). While the most recent version of the CDR scale classifies patients into five different categories rather than three [5], the CDR scale used in the ADNI data can still serve as a useful output class for predictive models.

2. METHODS

Attribute selection is an important element in creating an effective machine learning ensemble. Even though it is acknowledged as important, approaches to attribute selection are often ad-hoc or incomplete. This research employed an exhaustive search of different attribute selection algorithms in order to provide a more structured approach to learning design.

Several different algorithms were used for attribute selection. ClassifierSubsetEval, CFSSubsetEval, and ConsistencySubsetEval were paired with the GreedyStepwise, BestFirst, and SubsetSizeForwardSelection search methods. The lists of attributes produced by each evaluator/method pair were then compared. Not all of these pairings selected the same set of attributes. However, there were 15 attributes that ranked in the top 20 in sets produced by at least three of the ten pairings. These 15 were noted. Before we completed the next stage of our experiments, a set of 18 additional attributes we had not included in the original 108 became available. These new attributes comprised biologically significant SNP haplotype data. These SNP data were

BCB'14, September 20–23, 2014, Newport Beach, CA, USA.
ACM 978-1-4503-2894-4/14/09.
http://dx.doi.org/10.1145/2649387.2660797.

Model	Accuracy	Specificity	Sensitivity
LogitBoost (numIterations=13) - DecisionStump	79.29%	50.40%	90.50%
Bagging - REPTree (numFolds=3)	78.08%	33.50%	95.40%
Decorate - BayesNet	76.00%	57.40%	83.20%
DTNB	77.83%	35.20%	94.40%
LogitBoost - DecisionTable	79.29%	51.30%	90.20%
Ridor	76.49%	25.70%	96.30%
MultiBoostAB - MultiLayerPerceptron (learningRate=0.1, hiddenLayers=32)	73.45%	40.90%	86.10%
CostSensitiveClassifier([0,5.0;1.0,0], minimizeExpectedCost=TRUE)	61.63%	80.00%	54.50%
NaiveBayes	70.04%	64.30%	72.30%
Decorate - LogitBoost - DecisionTable	79.29%	46.50%	92.00%
Vote (majority vote) - {logitBoost - DecisionStump, Decorate – DTNB}	81.49%	48.70%	94.20%

Figure 1: Models with the most notable results

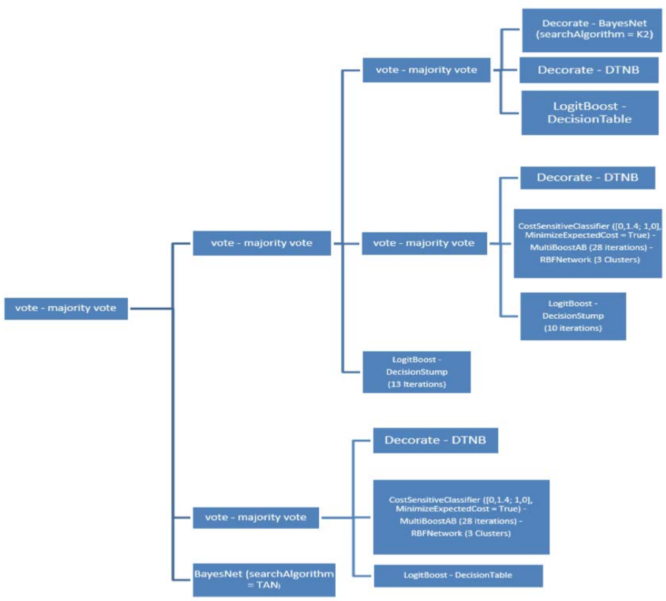

Figure 2: Best model (82.095% accuracy)

added to the set of 15 selected in the initial screening and we proceeded with a second screening on the new intermediate set of 33 attributes.

In all, Over 100 different models were generated with fifty-three models used as a base set from which to build majority-voting ensembles. These models were chosen because they either (1) had an accuracy of greater than 76% or (2) had an unusually high specificity, since most of the models with accuracy had relatively low specificity.

We used a simple program to do an exhaustive search of all possible combinations of three models selected from the 53 models chosen without replacement. The ensemble with the highest accuracy was then constructed and added to the group of models. Several meta techniques proved to be helpful. When paired with rule-based classifiers, Decorate and Logitboost produced all the non-ensemble models that achieved 79+% accuracy shown in Fig. 1. CostSensitiveClassifier produced the best RandomForest model and, combined with MultiBoostAB, the best RBFNetwork model. Bagging helped produce the best REPTree model (Fig. 2).

Our basic program showed that it is possible to search through a large number of potential ensemble combinations quickly without having to build the models in Weka. In a matter of minutes, we were able to search through $\binom{53}{3}$ = 23,426 majority-vote three-model ensembles; building all of these models manually in Weka likely would have taken weeks. Furthermore, our exhaustive search revealed an encouraging fact: there were many different ensembles that achieved an overall prediction accuracy of 80% or better. Time constraints made it impractical to build every ensemble with 80+% prediction accuracy in Weka so that it could be added to the original 53 models, so we only built the best ensemble format each iteration. However, it is possible that the other ensembles might have had different confusion matrices and, therefore, might have had the potential to be combined into even better ensembles of ensembles.

3. CONCLUSIONS

A structured approach to ensemble learning can effectively examine the space of attribute selection algorithms and classification algorithms by generating all combinations of three models, and examining their output to determine an ensemble's accuracy on the test data. Some models appear to improve accuracy when they are included multiple times in the ensemble hierarchy. It is unlikely that this hierarchy would have been discovered by traditional ad-hoc methods.

4. ADDITIONAL AUTHORS

Additional authors: John Kauwe and Perry Ridge (BYU Biology Department, 401 WIDB, Provo, UT, 84602).

5. REFERENCES

[1] A. Association et al. Alzheimer's disease fact sheet. *Alzheimer's Association*, 2011.

[2] D. G. Flood, G. J. Marek, and M. Williams. Developing predictive csf biomarkers?a challenge critical to success in alzheimer's disease and neuropsychiatric translational medicine. *Biochemical pharmacology*, 81(12):1422–1434, 2011.

[3] L. C. Lu and J. H. Bludau. *Alzheimer's Disease.* Greenwood Publishing Group, 2011.

[4] J. Lundkvist, M. M. Halldin, J. Sandin, G. Nordvall, P. Forsell, S. Svensson, L. Jansson, G. Johansson, B. Winblad, and J. Ekstrand. The battle of alzheimer?s disease–the beginning of the future unleashing the potential of academic discoveries. *Frontiers in pharmacology*, 5, 2014.

[5] J. C. Morris. The clinical dementia rating (cdr): current version and scoring rules. *Neurology*, 1993.