Evaluating the use of Different Measure Functions in the Predictive Quality of ABC-Miner

Khalid M. Salama
School of Computing
University of Kent
Canterbury, Kent CT2 7NZ
kms39@kent.ac.uk

Alex A. Freitas
School of Computing
University of Kent
Canterbury, Kent CT2 7NZ
A.A.Freitas@kent.ac.uk

ABSTRACT

Learning classifiers from datasets is a central problem in data mining and machine learning research. ABC-Miner is an Ant-based Bayesian Classification algorithm that employs the Ant Colony Optimization (ACO) meta-heuristics to learn the structure of Bayesian Augmented Naïve-Bayes (BAN) Classifiers. One of the most important aspects of the ACO algorithm is the choice of the quality measure used to evaluate a candidate solution to update pheromone. In this paper, we explore the use of various classification quality measures for evaluating the BAN classifiers constructed by the ants. The aim is to discover how the use of different evaluation measures affects the quality of the output classifier in terms of predictive accuracy. In our experiments, we use 4 different classification measures on 15 benchmark datasets.

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*Heuristic methods*

General Terms

Algorithms

Keywords

Ant Colony Optimization, Data Mining, Classification, Bayesian Network Classifiers

1. INTRODUCTION

Classification is one of the widely studied data mining tasks, in which the aim is to learn a model used to predict the class of unlabelled cases [8]. Bayesian network (BN) classifiers are used to predict the class of a case by computing the class with the highest posterior probability given the case's predictor attribute values, and learning effective BN classifiers – in terms of predictive accuracy – is our focus in this work.

GECCO'13 Companion, July 6–10, 2013, Amsterdam, The Netherlands.

ABC-Miner [7], recently introduced in the literature, is a classification algorithm that learns the structure of a Bayesian Augmented Naïve-Bayes (BAN) network using Ant Colony Optimization (ACO) – a global-search meta-heuristics for solving combinatorial optimization problems [2]. The Ant-based Bayesian Classification algorithm showed predictive effectiveness compared to other Bayesian classification algorithms, namely: Naïve-Bayes, TAN and GBN [7]. Moreover, experiments also showed that the use of *accuracy* – a classification quality measure – as a quality evaluation measure during the algorithm's training phase is more effective than the use of conventional Bayesian scoring functions.

The motivation behind this work is based on the previous conclusion; since ABC-Miner showed classification effectiveness, we work on extending the algorithm. In addition, one of the most important aspects of the ACO algorithm is the choice of the quality measure used to evaluate a candidate solution to update pheromone. In this paper, we explore the use of various classification quality measures for evaluating the BN classifiers constructed by the ants in the ABC-Miner algorithm. The aim of this investigation is to discover how the use of different evaluation measures affects the quality of the output classifier in terms of predictive accuracy. In our experiments, we explore the use of 4 different classification measures on 15 UCI Machine Learning repository [1] benchmark datasets.

2. THE ABC-Miner ALGORITHM

ABC-Miner is an ACO algorithm that learns a BN classifier by searching for the best possible Structure of a Bayesian network Augmented Naive Bayes (BAN) having at most k-dependencies (parents) at each variable node [7]. The construction graph consists of all the edges of the form $X \rightarrow Y$ where $X \neq Y$ and X, Y belong to the set of predictor attributes in the dataset. These edges represent probabilistic attribute dependencies in a BN classifier.

In essence, at each iteration, each ant incrementally constructs a candidate solution (i.e., a BN classifier). Then the quality of each candidate BN classifier is measured. The best solution produced in the colony at the current iteration undergoes local search, and then the BN classifier resulting from that local search is used to update the pheromone in the construction graph path corresponding to that classifier. After that, the system compares the quality of the current iteration's best solution $Q(tbest)$ with the quality of the global best solution $Q(gbest)$, in order to keep track of the best solution found along the entire search so far. This is repeated until the algorithm converges, or the predefined maximum number of iterations is reached.

An ant starts by considering a very simple BN classifier structure, namely a Naïve-Bayes structure, where each variable has just one parent, namely the class variable. Then the ant incrementally builds a more complex network, in the form of a Bayesian Augmented Naïve-Bayes (BAN) structure, by adding one edge at a time to the current network structure. The selection of the edge to be added at each step is based on both the heuristic function value and the pheromone amount associated with each valid candidate edge that could be added at this step, using the probabilistic state transition formula in [7]. An edge is valid to be added to the current partial BN classifier if the inclusion of that edge in the classifier does not create a directed cycle and does not exceed the upper limit of k parents for the current node (a limit chosen by the current ant). Once an edge is added to the current partial BN classifier, all the invalid edges are eliminated from the construction graph available for that ant. This process is repeated until no valid edges are available for that ant.

When the BN structure constructed by an ant is finished, the CPT (Conditional Probability Table) is computed for each variable, producing a complete BN classifier. Then the quality of the solution is evaluated and all the edges become available for constructing further candidate solutions.

The ABC-Miner algorithm evaluates the quality of the candidate constructed BN classifier during the training phase using *accuracy* [7], a conventional predictive measure, since the goal is to build a BN only for predicting the value of a specific class attribute, unlike conventional BN learning algorithms whose scoring function does not distinguish between the predictor and the class attributes.

3. CLASSIFIER QUALITY MEASURES

To Evaluate the predictive performance of a classifier, we count the cases (validation cases in the training phase and test cases in the test phase) correctly and incorrectly predicted by the classifier. These counts are organized in a tabular structure known as a *confusion matrix*.

For binary classification problems, where the class variable has exactly two values, only one confusion matrix is computed. However, in multi-class problems, where the class variable has more than two values, several matrices are computed, one for each class value considered as the positive class, with all the other classes being grouped together to form the negative class. One common approach for calculating the overall quality from several confusion matrices is to calculate the quality on each class using a specific measure with each confusion matrix separately, and take the average of the qualities calculated across all the classes.

Various classification quality evaluation measures are formulated using these elements of the confusion matrix, with different biases and quality aspects' importance. Several works aimed to study the effectiveness of these measures, yet in different classification contexts such as classification rule induction [6, 5, 3, 4], which highlighted the importance of rule quality measure chosen to be used to guide the search. We explore the effect of these different quality evaluation measures in guiding the ACO search to construct effective BN classifiers. The measure functions used in our experiements are `Accuracy`, `F-measure`, `Sensitivity` $\times$ `Specificity` and `Jaccard Coefficient`.

4. EXPERIMENTS AND RESULTS

The performance of classification quality measures was evaluated using 15 public-domain datasets from the well-known UCI (University of California at Irvine) dataset repository [1].

According to the results, `Sensitivity` $\times$ `Specificity` achieved the highest predictive accuracy amongst all quality measures in 8 datasets, while `Jaccard` achieved the highest accuracy in 5 datasets, `accuracy` in 2 datasets and `F-measure` in 1 dataset.

`Sensitivity` $\times$ `Specificity` obtained the best overall averaging raking with a value of **2.1**, followed by `Jaccard` that obtained an overall averaging ranking with a value of **2.4**. `accuracy` and `F-measure` came in the third and the fourth places respectively, with overall average ranking values **2.6** and **3.1** respectively.

5. CONCLUSION

ABC-Miner is a recently introduced algorithm that employs the ACO meta-heuristics to construct BN classifiers. We explored the effect of using 4 different quality measures, namely `accuracy`, `F-measure`, `Sensitivity` $\times$ `Specificity` and `Jaccard` for evaluating the candidate BN classifiers constructed by the ants and updating pheromone during the training phase of ABC-Miner. The quality of the final models are evaluated in terms of predictive accuracy.

Empirical evaluation on 10 UCI datasets has shown that the performance of different quality measures varies substantially across different datasets. However, the `Sensitivity` $\times$ `Specificity` measure has obtained the best overall average predictive performance. One possible research direction is to try to combine the use of several quality measures in the same learning procedure, which is left to future work.

6. REFERENCES

[1] UCI Repository of Machine Learning Databases. Retrieved Oct 2011 from, URL:www.ics.uci.edu/ mlearn/MLRepository.html.

[2] M. Dorigo and T. Stützle. *Ant Colony Optimization*. The MIT Press, 2004.

[3] J. Furnkranz and P. Flach. ROC 'n Rule Learning - Towards a Better Understanding of Covering Algorithms. *Machine Learning,*, pages 39–77, 2005.

[4] F. Janssen and J. Furnkranz. On the quest for optimal rule learning heuristics. *Machine Learning*, pages 343–379, 2010.

[5] M. Medland and F. Otero. A Study of Different Quality Evaluation Functions in the cAnt-Miner(PB) Classification Algorithm. *4ht International Conference on Genetic and Evolutionary Computation Conference (GECCO 2012)*, pages 49–56, 2012.

[6] K. M. Salama and A. M. Abdelbar. Exploring Different Rule Quality Evaluation Functions in ACO-based Classification Algorithms. *IEEE Swarm Intelligence Symposium*, pages 1–8, 2011.

[7] K. M. Salama and A. A. Freitas. ABC-Miner: an Ant-based Bayesian Classification Algorithm. *International Conference on Swarm Intelligence (ANTS)*, pages 2677–2694, 2012.

[8] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 3rd edition, 2010.