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 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.

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 Naïve 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 \( Q(\text{best}) \) with the quality of the global best solution \( Q(\text{gbest}) \), in order to keep track of the best solution found along the entire search so far. This
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 experiments are Accuracy, F-measure, Sensitivity × Specificity and Jaccard Coefficient.