

Semester: Summer 2010
MS Student: Christopher Utz
Chair: Dr. Amy McGovern
Thesis Title: LEARNING ENSEMBLES OF BAYESIAN NETWORK STRUCTURES
USING RANDOM FOREST TECHNIQUES

ABSTRACT:

We introduce an ensembled classifier of Bayesian Networks, EBNs, that automatically learn the structure of a set of Bayesian Networks. Our experiments demonstrate that EBNs have many of the same benefits of Random Forests when compared with the classification performance of traditional Bayesian Networks. Specifically, we highlight the success of EBNs as a model for classification when trained on small datasets. Similar to Random Forests, EBNs do not overfit. Additionally, EBNs outperform or perform as well as traditional Bayesian Networks when tasked with instance classification. EBNs also provide a measure of variable importance and can be generated efficiently on a variety of datasets. Using datasets from the standard Bayesian Network Repository we show that EBNs perform equal to or better than traditional Bayesian Networks when used for classification. We analyze a small case study in engineering education to illustrate the performance benefits of EBNs on small datasets and the ability of EBNs to identify important variables within datasets. We focus on predicting retention outcomes of engineering students and identifying factors related to attrition in engineering education.