Bayesian Network Classifiers

Alessio Falai

University of Bologna
alessio.falai@studio.unibo.it

June 18, 2020

Abstract

This report explains the work done and the results achieved in the Knowledge Representation class project, specifically in the module about uncertainty and probabilistic reasoning, whose lectures were held by professor Paolo Torroni. In this work, we used the Adult dataset to classify world citizens that perceive a high income, based on different features. The classification task was performed using variuos Bayesian network structures and inference algorithms. Probabilistic inference results on a test set were then compared with ground-truth data to evaluate the accuracy of the dataset.

I. Introduction

In this work, we tested the capabilities of various Bayesian networks structures and inference algorithms combinations in a classification task, over the standard Adult dataset, which aims at separating people whose income is greater than 50 thousands dollars per year from the rest.

The first operation that needed to be done was data cleaning:

- Useless features, like fnlwgt, were removed
- Redundant features, like education-num, were removed too
- Rows containing null values were removed, since there were only a few

The second operation that needed to be done was data discretization, to simplify the following construction of the Bayesian networks structures:

- The age variable was divided into 4 bins (child, young adult, adult and senior)
- The hours-per-week variable was divided into 4 bins (part time, full time, over time and too much time)
- The capital-gain and capital-loss variables were binned according to different quantiles distributions

The output of this pre-processing step will be used as input for the following steps.

II. Bayesian networks

Different Bayesian networks structures were used to compare classification capabilities over the same dataset, so as to assess which structure could be more suitable to solve the presented problem.

Every tested structure is actually based on the Naive Bayes model and compared against it.

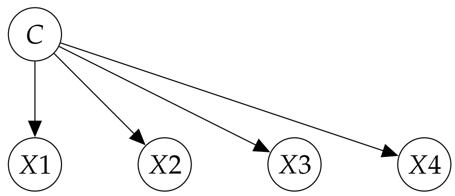
i. Naive Bayes (NB)

The Naive Bayes model has been extensively used in classification tasks, with good accuracy results, because of its simplicity.

It does not require any structural learning, since it has a fixed structure, where the classification variable is the parent node of every other feature variable.

The Naive Bayes model works by assuming full independence between each pair of variables, given the classification node.

Figure 1: *Example of a NB model*

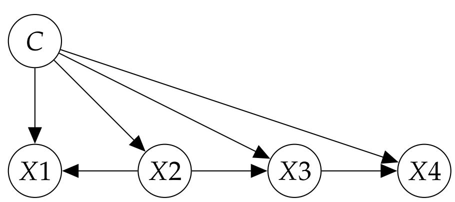


ii. Tree-Augmented Naive Bayes (TAN)

The TAN model is just like a Naive Bayes model, so there is a connection from the classification node to every other feature node, with the exception that, without the classification node and all its related edges, the Bayesian network becomes a tree.

The TAN model needs to be learned from the training data, by using a modified version of the Chow-Liu algorithm.

Figure 2: *Example of a TAN model*

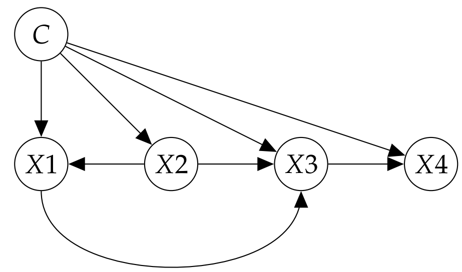


iii. BN-Augmented Naive Bayes (BAN)

The BAN model is just like a Naive Bayes model, so there is a connection from the classification node to every other feature node, with the exception that, without the classification node and all its related edges, the Bayesian network becomes a DAG (Directed Acyclic Graph).

The BAN model needs to be learned from the training data, by using a modified CBL2 algorithm.

Figure 3: *Example of a BAN model*

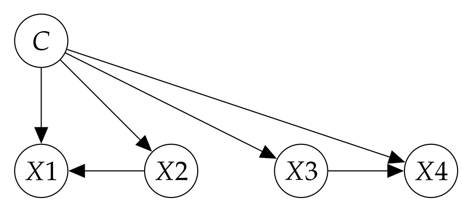


iv. Forest-Augmented Naive Bayes (FAN)

The FAN model is just like a Naive Bayes model, so there is a connection from the classification node to every other feature node, with the exception that, without the classification node and all its related edges, the Bayesian network becomes a forest.

The FAN model needs to be learned from the training data, by using a similar reasoning as the Chow-Liu algorithm.

Figure 4: *Example of a FAN model*



III. Inference algorithms

i. Variable elimination

ii. Other

IV. Results

Table 1: *Results summary*

	NB	TAN	BAN	FAN
Variable elimination				
Other				

References

[Chow, Liu, 1968] Chow, C.K. and Liu, C.N.. Approximating discrete probability distributions with dependence trees. *IEEE Trans. on Information Theory*, 14 (pp. 462-467).

[Cheng et al., 1997a] Cheng, J., Bell, D.A. and Liu, W.. An Algorithm for Bayesian Belief Network Construction from Data. *Proceedings of AI & STAT'97* (pp. 83-90), Florida.

[Cheng et al., 1997b] Cheng, J., Bell, D.A. and Liu, W.. Learning Belief Networks from Data: An Information Theory Based Approach. *Proceedings of ACM CIKM'97*.

[Cheng et al., 1999] Jie Cheng, Russell Greiner. Comparing Bayesian Network Classifiers. *UAI'99: Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence* (pp. 101-108).

[Friedman et al., 1997] Jie Cheng, Russell Greiner. Bayesian Network Classifiers. *Machine Language*, Volume 29, Issue 2-3.

[Jiang et al., 2005] Jiang, Zhang, Cai, Su. Learning Tree Augmented Naive Bayes for Ranking. *DASFAA'05: Proceedings of the 10th international conference on Database Systems for Advanced Applications* (pp. 688-698).