Probabilistic Graphical Model

Bayesian Network with R package bnlearn

Nidhi Rajkotia Bipinbhai
MTECH CSE (2020-21)
IIIT Vadodara
Gandhinagar, India
202061007@iiitvadodara.ac.in

Dhameja Prashant
MTECH CSE (2020-21)
IIIT Vadodara
Gandhinagar, India
202061003@iiitvadodara.ac.in

Rathod Dharmraj Manojbhai
MTECH CSE (2020-21)
IIIT Vadodara
Gandhinagar, India
202061008@iiitvadodara.ac.in

***Abstract*—Main purpose of this tutorial is to understand probabilistic graphical model with implementation of Bayesian network in R using bnlearn package and will give some insight of learning structure and Conditional probability table(CPT).**

***Index Terms*—graphical models for inference under uncertainty,bayesian networks in R,naive Bayes classification,Modern AI.**

I. Introduction

In this uncertain world belief network is very important so, in recent years Bayesian networks have been used in many different fields like health,service research,cancer research, OLAP. high dimensional data lead to development of learning algorithms which are more focused on reducing the complexity of the problem with learning the correct network. bnlearn package in r language give free implementation of some of this learning structure using conditional independence test and some scores and that will construct the Bayesian network.

A. Bayesian networks

Bayesian networks which represent the random variables as nodes and conditional dependency as edge between them. The graphical structure G = (V, A), where V is vertex set and A is egde set. and Directed Acyclic graph defines the factorization of the joint probability distribution of those all random variables of set V. The form of the factorization is given by the Markov property of Bayesian networks which state that X_i (random variable) directly depends on its parents π_{xi}

$$P(X_1, ..., X_v) = [\prod_{i=1}^{v} P(X_i|\pi_{xi})] \tag{1}$$

$$P(X_1, ..., X_v) = [\prod_{i=1}^{v} f(X_i|\pi_{xi})] \tag{2}$$

(1)for discrete variables (2)for continuous variables an arbitrary triplet of disjoint subsets of V will tell about The correspondence between conditional independence of the random variables and graphical separation of the corresponding nodes of the graph that is the d-separation meaning that from direction-dependent separation of the nodes. first of all algorithms try to learn the graphical structure of the Bayesian network.

B. Structure learning algorithms

Structure learning algorithms of the Bayesian network can be grouped in two categories:

- *Constraint-based algorithms*: these kinds of algorithm learn the structure of the network by probability relations because of the Markov property of Bayesian networks and conditional independence and then construct the network which satisfy the d-separation statement.
- *Score-based algorithms* : These kinds algorithms gives a score to each possible Bayesian network and try to maximize it with some heuristic search algorithm. Greedy search algorithms such as hill-climbing or tabu search are a common choice, but almost any kind of search procedure can be used.
  for our problem statement we have mostly used score-based algorithm with different scores such as k2, aic(Akaike information criterion scores), bic(Bayesian information criterion scores)

II. Problem Statement

A. Table containing grades earned by students in respective courses is made available to you in (codes folder) 2020_bn_nb_data.txt.

Table contains the grades of 8 different courses and as the concept of the algorithm we have to understand the learning structure of the Bayesian network using the grades of the courses. we have used score-based algorithm for learning the structure of the Bayesian network that is hill climbing with different scores (bic, aic, k2).

figures of Bayesian network for different scores are given below.

AIC_model_1:
$[EC100][EC160|EC100][IT101|EC100][MA101|EC100]$
$[PH100|EC100][IT161|IT101][HS101|IT101][PH160|HS101]$

BDLA_model_2:
$[IT161][PH160][IT101|IT161][MA101|IT101]$
$[HS101|IT101][EC100|MA101][EC160|EC100][PH100|EC100]$

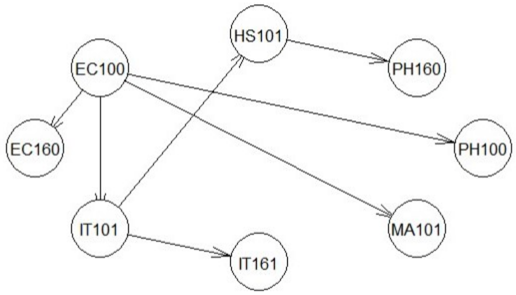


Fig. 1. AIC Model 1

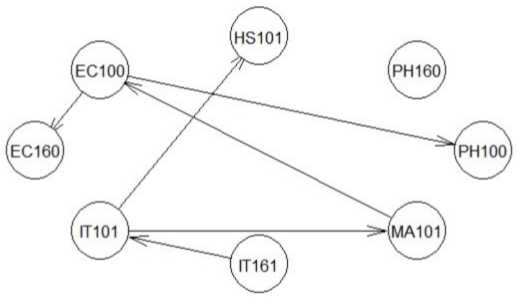


Fig. 2. BDLA Model 2

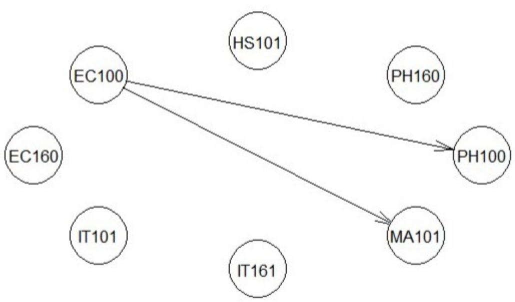


Fig. 3. BIC Model 3

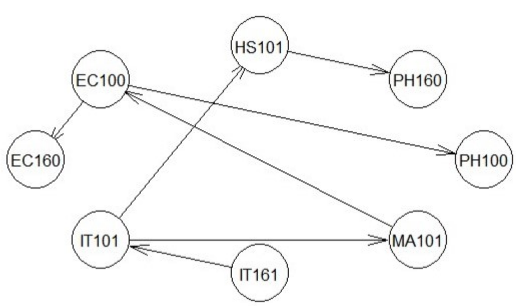
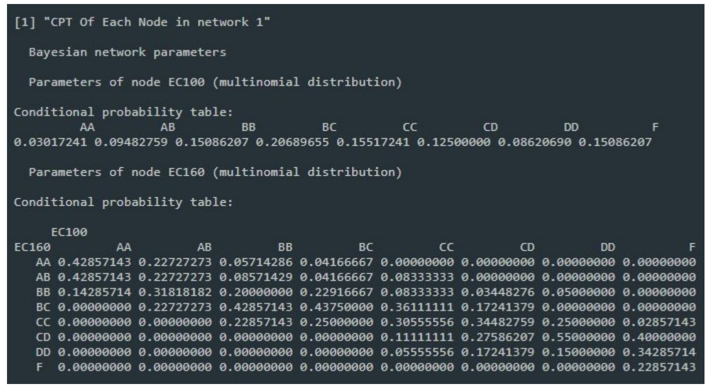


Fig. 4. k2 Model 4



```
[1] "CPT Of Each Node in network 1"

  Bayesian network parameters

  Parameters of node EC100 (multinomial distribution)

Conditional probability table:
        AA         AB         BB         BC         CC         CD         DD          F
0.03017241 0.09482759 0.15086207 0.20689655 0.15517241 0.12500000 0.08620690 0.15086207

  Parameters of node EC160 (multinomial distribution)

Conditional probability table:

     EC100
EC160         AA         AB         BB         BC         CC         CD         DD          F
   AA 0.42857143 0.22727273 0.05714286 0.04166667 0.00000000 0.00000000 0.00000000 0.00000000
   AB 0.42857143 0.22727273 0.08571429 0.04166667 0.08333333 0.00000000 0.00000000 0.00000000
   BB 0.14285714 0.31818182 0.20000000 0.22916667 0.08333333 0.03448276 0.05000000 0.00000000
   BC 0.00000000 0.22727273 0.42857143 0.43750000 0.36111111 0.17241379 0.00000000 0.00000000
   CC 0.00000000 0.00000000 0.22857143 0.25000000 0.30555556 0.34482759 0.25000000 0.02857143
   CD 0.00000000 0.00000000 0.00000000 0.00000000 0.11111111 0.27586207 0.55000000 0.40000000
   DD 0.00000000 0.00000000 0.00000000 0.00000000 0.05555556 0.17241379 0.15000000 0.34285714
   F  0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.22857143
```

Fig. 5. CPT of node EC100 and EC160

BIC_model_3:
$[EC100][EC160][IT101][IT161][PH160][HS101]$
$[MA101|EC100][PH100|EC100]$
k2_model_4:
$[IT161][IT101|IT161][MA101|IT101][HS101|IT101]$
$[EC100|MA101][PH160|HS101][EC160|EC100][PH100|EC100]$

B. Using the data, learn the CPTs for each course node.

By learning the structure now we can learn about the conditional probability Table(CPT) of each node in Bayesian network. Conditional independence test for data are the function of CPT. in bnlearn after fitting the network we can get the CPT of the each node. example of one of the CPT is as given Fig.5 for Model1.

C. What grade will a student get in PH100 if he earns DD in EC100, CC in IT101 and CD in MA101.

In bnlearn package we have another function by which we can directly fire query to predict the values for that data.In this question we want predict grade for course PH100 given some evidences of the different courses. cpquery() function of the bnlearn will help us to find the probability that which grade will have the highest probability for PH100 given the evidences of different courses and from that we can predict the grade of PH100.

D. The last column in the data file indicates whether a student qualifies for an internship program or not. From the given data, take 70 percent data for training and build a naive Bayes classifier (considering that the grades earned in different courses are independent of each other) which takes in the student's performance and returns the qualification status with a probability. Test your classifier on the remaining 30 percent data. Repeat this experiment for 20 random selection of training and testing data. Report results about the accuracy of your classifier.

bnlearn package give us a very good functionality we can easily implement the Bayesian Classifier. first we have divide the dataset into training and test set in ratio of 7:3. we have made naive Bayes classifier with target feature

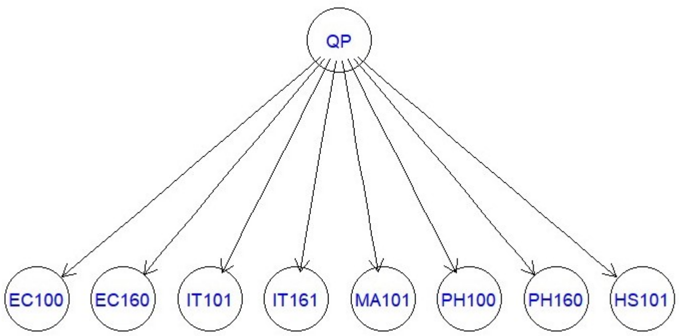


Fig. 6. Naive Bayes IID

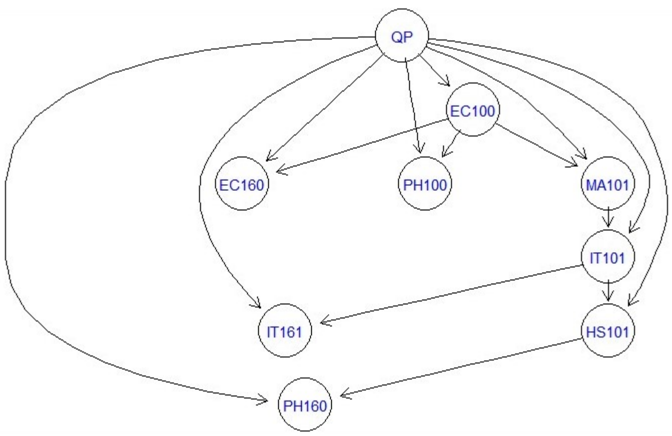


Fig. 7. Naive Bayes Dependent

QP(Qualification status) yes or no. bayesian network for naive Bayes classifier is as in Fig.6. now we can predict the QP for all the data which are present in the test data using Predict() function with or without the probability of the class yes and class no. as how much probability is there that this student will qualify for the placement or not.

so now we can get the confusion matrix using table() method in bnlearn and accuracy using accuracy() function.

we have run this experiment several times and then took the average for for all accuracy that will give us better idea for accuracy.

E. Repeat D, considering that the grades earned in different courses may be dependent.

same experiment we will conduct with dependent random variables and we will check that in this case accuracy is decreasing as compare to the IID random variables. bayesian network of the dependent random variables is as in Fig.7.

III. Implementation Repository

The implementation of the Bayesian network and naive Bayes classifier (Github)

References

Reference [1] [2].

References

[1] Mihaljević, Bojan. "Bayesian Networks with R," n.d., 59
[2] "Bnlearn - Bayesian Network Structure Learning." https://www.bnlearn.com/