

Redes Bayesianas como Ferramentas para o Raciocínio Clínico

José Elvano Moraes

17 de março de 2021

Learning Medical Application of Bayesian Networks. Usando a package *bnlearn*

```
library(bnlearn)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.3.3      v purrr  0.3.4
## v tibble  3.0.6      v dplyr  1.0.4
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(haven)
library(Rgraphviz)

## Loading required package: graph
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
##
## The following objects are masked from 'package:parallel':
##
##   clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##   clusterExport, clusterMap, parApply, parCapply, parLapply,
##   parLapplyLB, parRapply, parSapply, parSapplyLB
##
## The following objects are masked from 'package:dplyr':
##
##   combine, intersect, setdiff, union
##
## The following object is masked from 'package:bnlearn':
##
##   score
##
## The following objects are masked from 'package:stats':
##
##   IQR, mad, sd, var, xtabs
##
## The following objects are masked from 'package:base':
##
##   anyDuplicated, append, as.data.frame, basename, cbind, colnames,
```

```
##      dirname, do.call, duplicated, eval, evalq, Filter, Find, get, grep,
##      grepl, intersect, is.unsorted, lapply, Map, mapply, match, mget,
##      order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##      rbind, Reduce, rownames, sapply, setdiff, sort, table, tapply,
##      union, unique, unsplit, which.max, which.min

##
## Attaching package: 'graph'

## The following object is masked from 'package:stringr':
##
##      boundary

## The following objects are masked from 'package:bnlearn':
##
##      degree, nodes, nodes<-

## Loading required package: grid

##
##
## setwd("~/Documents/books/bayes/cardio_(...)_")
##

cardio_train <- read_delim("cardio_train_copia.csv",
                          ";", escape_double = FALSE, col_types = cols(
                            gender = col_factor(levels = c("1",
                                                            "2")),
                            cholesterol = col_factor(levels = c("1",
                                                                "2", "3")),
                            gluc = col_factor(levels = c("1",
                                                         "2", "3")),
                            smoke = col_factor(levels = c("0",
                                                          "1")),
                            alco = col_factor(levels = c("0",
                                                         "1")),
                            active = col_factor(levels = c("0",
                                                           "1")),
                            cardio = col_factor(levels = c("0",
                                                           "1")),
                            trim_ws = TRUE)

set.seed(2)
cardio_train = filter(cardio_train, ap_hi>ap_lo, ap_lo > 40, ap_lo<200, ap_hi>60, ap_hi<300)

## Masculino
#
genero = '1'
#
c_ = filter(cardio_train, gender == genero)
#
t <- select(c_, age, height, weight, ap_lo, ap_hi)
t2 <- discretize(t, method = 'hartemink', breaks = 4)

dd_m <- cbind(select(c_, -age, -height, -weight, -ap_lo, -ap_hi), t2)
summary(dd_m)

##      id      gender  cholesterol gluc      smoke      alco      active
## Min.   :      1  1:44719   1:33031   1:37802  0:43925  0:43592  0: 8870
## 1st Qu.:25014   2:      0   2: 6253   2: 3345   1: 794   1: 1127  1:35849
## Median :50026                3: 5435   3: 3572
## Mean   :49913
## 3rd Qu.:74588
## Max.   :99999
## cardio      age      height      weight
```

```
## 0:22713 (10859,15532] : 4027 (55,155] : 8042 (21,57] : 4686
## 1:22006 (15532,18539.1]:11178 (155,164]:21341 (57,64] : 8224
## (18539.1,22402]:23701 (164,165]: 4124 (64,80] :21370
## (22402,23701] : 5813 (165,178]:11212 (80,115]:10439
##
##
## ap_lo ap_hi
## (45,70] : 9491 (70,110] : 9061
## (70,80] :22821 (110,120]:17816
## (80,90] : 9381 (120,140]:12208
## (90,100]: 3026 (140,180]: 5634
##
##
```

```
#
#
saveRDS(dd_m, 'dd_masculino_hartemink_breaks_4_seed_2.rds')
rm(t, t2, c_)
```

```
## Feminino
#
genero = '2'
#
c_ = filter(cardio_train, gender == genero)
#
t <- select(c_, age, height, weight, ap_lo, ap_hi)
t2 <- discretize(t, method = 'hartemink', breaks = 4)

dd_f <- cbind(select(c_, -age, -height, -weight, -ap_lo, -ap_hi), t2)
summary(dd_f)
```

```
## id gender cholesterol gluc smoke alco active
## Min. : 0 1: 0 1:18458 1:20574 0:18691 0:21400 0: 4630
## 1st Qu.:24991 2:23937 2: 3046 2: 1720 1: 5246 1: 2537 1:19307
## Median :49996 3: 2433 3: 1643
## Mean :50100
## 3rd Qu.:75429
## Max. :99996
## cardio age height weight
## 0:11979 (10798,15500] : 2396 (65,165] :6000 (11,65] :4841
## 1:11958 (15500,18384] : 6222 (165,170]:7660 (65,76] :8222
## (18384,19935.2]: 4787 (170,175]:5646 (76,90] :7252
## (19935.2,23713]:10532 (175,186]:4631 (90,119]:3622
##
##
## ap_lo ap_hi
## (50,75] : 3883 (80,115] : 3415
## (75,80] :12541 (115,120]:10093
## (80,90] : 5491 (120,140]: 6918
## (90,110]: 2022 (140,180]: 3511
##
##
```

```
#
#
saveRDS(dd_f, 'dd_feminino_hartemink_breaks_4_seed_2.rds')
```

```

dd_m <- readRDS('dd_masculino_hartemink_breaks_4_seed_2.rds')
bn_mas_iamb <- iamb(dd_m)

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> height <- weight is not applicable, because one or
## both arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> ap_hi <- weight is not applicable, because one or both
## arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cardio -> ap_hi <- weight is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure gluc -> cardio <- ap_lo is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure height -> weight <- ap_lo is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure age -> cardio <- ap_lo is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure age -> ap_hi <- ap_lo is not applicable, because one or both arcs are
## oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> cardio <- ap_lo is not applicable, because one or both
## arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> ap_hi <- ap_lo is not applicable, because one or both
## arcs are oriented in the opposite direction.

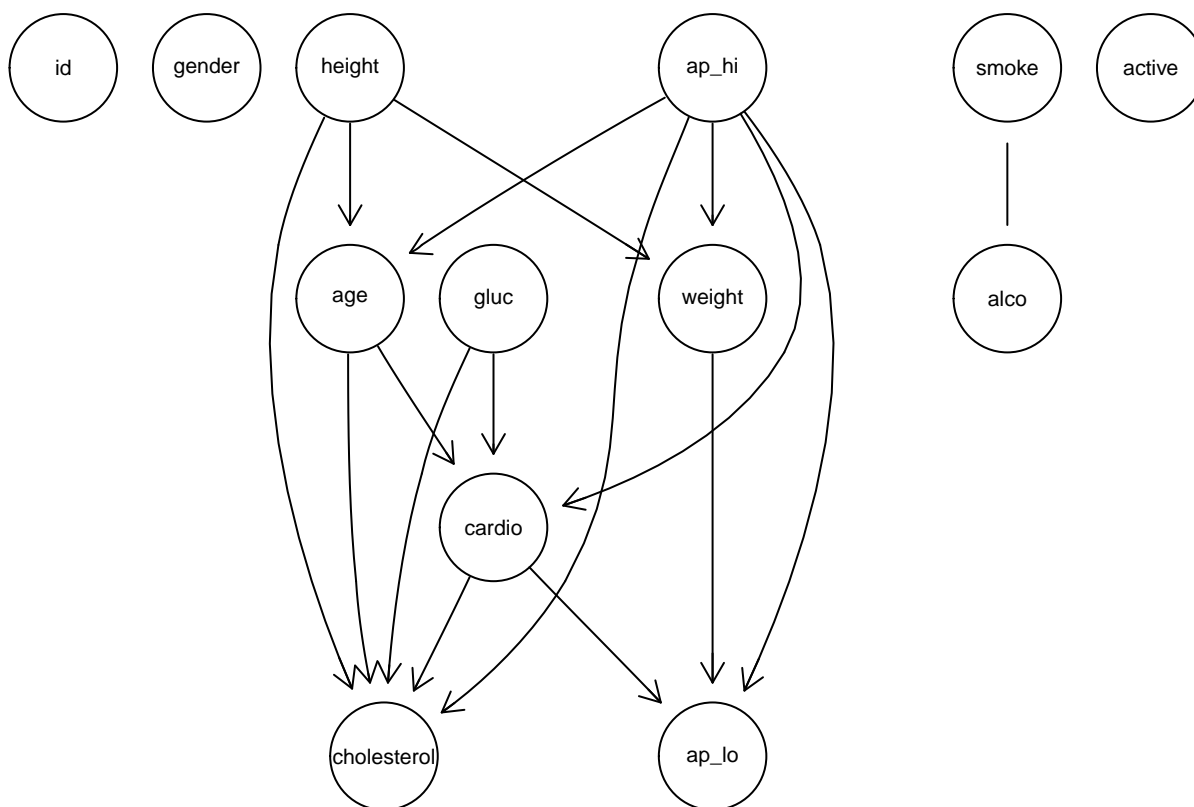
## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cardio -> age <- height is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure age -> height <- weight is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure age -> ap_hi <- weight is not applicable, because one or both arcs
## are oriented in the opposite direction.

graphviz.plot(bn_mas_iamb)

```



```
dd_f <- readRDS('dd_feminino_hartemink_breaks_4_seed_2.rds')
bn_fem_iamb <- iamb(dd_f)
```

```
## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure gluc -> cholesterol <- cardio is not applicable, because one or both
## arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure height -> weight <- ap_hi is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cardio -> cholesterol <- weight is not applicable, because one or
## both arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cardio -> ap_hi <- weight is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure gluc -> weight <- ap_hi is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> ap_hi <- ap_lo is not applicable, because one or both
## arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure smoke -> age <- height is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
```

```

## vstructure smoke -> alco <- active is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure smoke -> ap_hi <- cardio is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cardio -> age <- height is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure smoke -> age <- cardio is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> ap_hi <- smoke is not applicable, because one or both
## arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure smoke -> ap_hi <- weight is not applicable, because one or both arcs
## are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure cholesterol -> weight <- height is not applicable, because one or
## both arcs are oriented in the opposite direction.

## Warning in vstruct.apply(arcs = arcs, vs = vs, nodes = nodes, debug = debug):
## vstructure alco -> active <- cardio is not applicable, because one or both arcs
## are oriented in the opposite direction.

graphviz.plot(bn_fem_iamb)

```

